



**Universiteit Gent
Centrum voor Duurzame Ontwikkeling**

Gegevensbeheer inzake indicatoren voor duurzame ontwikkeling

**Standaardfiches en Informatiestromen
Een bijdrage aan het
InformatieSysteem voor Duurzame Ontwikkeling (ISDO)**

Studie in opdracht van het Federale Wetenschapsbeleid

**Promotor: Prof. Dr. R. Doom
Onderzoek: W. De Jonge, S. Deconinck**



Colofon

Gegevensbeheer inzake indicatoren voor duurzame ontwikkeling:
Standaardfiches en Informatiestromen;
Een bijdrage aan het Informatiesysteem voor Duurzame Ontwikkeling (ISDO)
<http://www.belspo.be/issd/>

Website van dit project: <http://cdonet.UGent.be/issd/>

Eindverslag, november 2003

Walter De Jonge
Stefan Deconinck

Centrum voor Duurzame Ontwikkeling (CDO)
Universiteit Gent
<http://Cdonet.UGent.be/>
Poel 16
B - 9000 Gent
België
tel: +32 (0)9 264.82.10
fax: +32 (0)9 264.83.90
email: cdo@UGent.be

Onderzoeksovereenkomst nummer AS/02/019, binnen het programma 'Ondersteunende Acties' bij het Plan voor Wetenschappelijke Ondersteuning van een beleid gericht op Duurzame Ontwikkeling (PODO-1);
Federaal Wetenschapsbeleid - <http://www.belspo.be/>

Gegevensbeheer inzake indicatoren voor duurzame ontwikkeling

**Standaardfiches en Informatiestromen
Een bijdrage aan het
InformatieSysteem voor Duurzame Ontwikkeling (ISDO)**

Gegevensbeheer inzake indicatoren voor duurzame ontwikkeling Inhoudsopgave

Samenvatting	1
I. Duurzame ontwikkeling, indicatoren, data en metadata	4
I.1 Het nut van (informatie-uitwisseling inzake) IDO	4
I.2 Het World Wide Web, Metadata en Duurzame Ontwikkeling	7
• Google, Yahoo en andere links	7
• Metadata betreffende Indicatoren voor Duurzame Ontwikkeling	10
• Metadata betreffende duurzame ontwikkeling in het algemeen	13
I.3 Enkele conclusies	18
II. De beschrijving van indicatoren voor duurzame ontwikkeling	20
II.1 Multicriteria-evaluatie	20
• Meting en bekommernis	20
• Compleetheid	22
• Boomstructuur	23
II.2 Standaardisering van informatie	24
II.3 Enkele conclusies	37
III. Metadata en Dublin Core Metadata Initiative	39
III.1 Inleiding	39
III.2 Dublin Core Metadata en Internet	40
• Dublin Core: een voorbeeld	40
• Metadata op het World Wide Web	42
III.3 Dublin Core: verfijning en uitbreiding	45
• Refinements	49
• Dublin Core Structured Values	51
• A-Core: Administrative Core	51
• EdNA: Education Network Australia	52
III.4 Evaluatie van Dublin Core	53
III.5 Enkele conclusies	55
IV. HTML, XML en XHTML	57
IV.1 Open Standaarden	57
• Open standaarden en metadata	58
• Open Source	59
• Officiële bronnen	59
IV.2 HTML: HyperText Markup Language	61
IV.3 XML: eXtensible Markup Language	63
IV.4 XHTML: eXtensible HyperText Markup Language	65

V. Web Accessibility Initiative (WAI)	73
V.1 Toegankelijkheid van informatie op het WWW	73
V.2 Enkele tips voor webmasters	75
V.3 Toegankelijkheid van informatie over (indicatoren voor) duurzame ontwikkeling	76
VI. Archivering van metadata	80
VI.1 Beheer met relationele database	80
• Regular expressions	81
• Management	84
VI.2 Harvest: oogsten van metadata	85
VI.3 Archiveren	90
VI.4 Open Source	91
• Open Source en duurzame ontwikkeling	93
• AMP: Apache-MySQL-PHP	94
VII. Producing XML- or XHTML-input through MS-Word	97
VIII. Using frames on an indicator website	114
IX. A simple DC-editor using client side Javascript	120
X. ISSD-metadata for indicators	129
XI. Slideshow: Indicators for Sustainable Development (ISD) metadata: ambitions	138

Samenvatting

Dit rapport brengt verslag van een (lange) zoektocht naar een informatiesysteem dat de opsporing en uitwisseling van gegevens inzake Indicatoren voor Duurzame Ontwikkeling (IDO) moet vergemakkelijken. In het eerste hoofdstuk wordt het gebruik van Indicatoren voor Duurzame Ontwikkeling kort toegelicht. Het belang dat aan IDO wordt gehecht en de vele IDO-projecten die wereldwijd werden en worden ontwikkeld, leidt tot de conclusie dat het in elk geval de moeite loont om te investeren in een informatiesysteem dat het opsporen van informatie over IDO en IDO-projecten vergemakkelijkt. Er worden enkele initiatieven belicht die daarin hebben geïnvesteerd. Het gaat dan om archieven (databanken) waarin gegevens over verschillende IDO-projecten zijn verzameld, ofwel archieven waarin men zijn weg kan zoeken naar informatie en projecten omtrent duurzame ontwikkeling in het algemeen (en niet alleen indicatoren).

Men kan daarbij vaststellen dat men in eerste instantie focust op de productie van metadata: gegevens die de oorspronkelijke gegevens of data (kort) beschrijven. Bij deze metadata behoort dan een link naar de data die worden beschreven. Meestal construeert men een database en elke record bevat dan een aantal metadata-velden over een indicator, indicator-project of andere informatie inzake duurzame ontwikkeling. In feite gaat men er dan steeds vanuit dat iemand manueel de database aanvult. We vonden alsnog geen projecten die voorzien in een procedure die 'automatisch' metadata verzamelt uit de oorspronkelijk gegevens.

Omdat de meeste data op het World Wide Web worden gepubliceerd is de geautomatiseerde inzameling van gegevens nochtans een optie die men kan/moet overwegen. Om dat te realiseren moet men evenwel tot afspraken komen over de manier waarop de data worden aangeboden op het web. Men moet bepaalde procedures standaardiseren. In dit rapport wordt een dergelijk procedure ontwikkeld, een procedure die aan de toeleveranciers een grote vrijheid geeft qua inhoudelijke samenstelling van hun gegevens, een procedure die bovendien nauwelijks beperkingen oplegt aan de metadata die men binnen een bepaalde context wil uitwisselen.

Indien men de informatie over IDO beschouwt, dan is de meest voor de hand liggende vraag of men het eens kan raken over een standaard fiche voor de beschrijving van een indicator. Met ander woorden: kan men het met de vele betrokken actoren wereldwijd eens raken over de inhoudelijke structuur van een document die een indicator beschrijft.

In het tweede hoofdstuk wordt aan de hand van een aantal voorbeelden gedemonstreerd dat dit moeilijk haalbaar is en wellicht zelfs onmogelijk. De inhoud en structuur van indicator-beschrijvingen varieert sterk als men meerdere projecten vergelijkt en men kan daar rationele verklaringen voor geven. Bovendien worden IDO zelden als losse entiteiten gepresenteerd. De IDO van een project zijn complementair aan elkaar en hebben een onderlinge samenhang. Meestal wordt 'duurzaamheid' beoordeeld via een multicriteria-evaluatie en de betekenis van een bepaalde indicator moet in het grotere geheel worden gezien.

Niet alleen de beschrijvingen van indicatoren, maar ook de grotere constructies waarvan ze deel uit maken, leiden tot een grote inhoudelijke en structurele variatie in het data-aanbod.

Indien men geen inhoudelijke standaard voor de data zelf kan ontwikkelen, dan kan men evenwel nog steeds streven naar een standaard voor metadata. Terzake is dan ook eerst een evaluatie gemaakt van een populaire metadata-standaard - met name de Dublin Core Metadata. In het derde hoofdstuk wordt de Dublin Core doorgelicht.

Het Dublin Core Metadata Initiatief (DCMI) was al eerder belicht in een ander project van het DWTC, waarin een 'informatiesysteem duurzame ontwikkeling' werd ontwikkeld (ISDO). Toen vond men de Dublin Core (DC) ontoereikend en men besliste een eigen metadata-systeem te ontwikkelen voor de beschrijving van projecten, organisaties, publicaties, personen en databases. Destijds had men blijkbaar voldoende argumenten om de Dublin Core niet toe te passen. Maar tijden veranderen: de Dublin Core is de laatste jaren enorm gestegen in populariteit en wordt blijkbaar een de facto standaard op het World Wide Web. Men kan de Dublin Core nog moeilijk negeren. Het is niet onwaarschijnlijk dat de meeste actoren op termijn liefst eerst Dublin Core metadata zullen produceren, omdat ze die metadata in vele metadata-archieven kunnen invoeren. Men stelt ook vast dat initiatieven die weliswaar hun eigen metadata ontwikkelen, daarnaast ook een DC-variante voorzien. Bovendien is de Dublin Core ook verder

ontwikkeld, waardoor deze tegemoet kan komen aan veel meer noden en zich flexibel opstelt voor het toevoegen van complementaire metadata. Te vaak wordt gedacht dat men moet kiezen tussen de Dublin Core enerzijds en metadata die afgestemd zijn op lokale behoeften anderzijds. Echter sluit het ene het andere niet uit. Men kan beide combineren.

Dat laatste leidt evenwel tot de vraag hoe men dat softwarematig implementeert. Toen het project startte lag het voor de hand om te bekijken hoe men het metadata-beheer en de uitwisseling van metadata kon baseren op het gebruik van XML of eXtended Markup Language. Deze internationale standaard voor informatie-opslag en uitwisseling was bij de aanvang van het project nog jong, maar veelbelovend. De laatste jaren is bewezen dat het dé standaard wordt voor databeheer en uitwisseling via het Internet. XML zelf is niet ingewikkeld, maar anders ligt het met het vruchtbaar - efficiënt en effectief - gebruik daarvan. Een ingewijde heeft geen gebrek aan gestandaardiseerde instrumenten om data in XML-format te manipuleren (XSLT, DOM en andere standaarden die het mogelijk maken om de inhoud van XML-documenten te verwerken). Maar een doorsnee organisatie die actief is in het domein duurzame ontwikkeling beheerst die expertise niet.

XML en gerelateerde standaarden (XSLT en DOM) bieden echter zoveel mogelijkheden, dat men er ook in verzuipt, zo is alvast tijdens dit project gebleken. We hebben het lange tijd veel ingewikkelder gemaakt dan nodig. In het vierde hoofdstuk wordt het standaardformaat XML belicht, en de oplossing die we bedachten om ook leken snel te laten genieten van deze www-technologie. De methode die we naar voor schuiven is uiterst eenvoudig en vereist bijna geen expertise inzake XML. Een organisatie die data produceert volgens ons XML-format heeft er overigens zelf direct baat bij: ze kunnen die data op hun website publiceren en elke browser zal die probleemloos interpreteren. We gebruiken immers XHTML en hebben dan zowel XML die geschikt is voor de meest uiteenlopende manipulaties met XML-software, als HTML die door elke browser correct weergegeven wordt, ook oudere browsers. Bovendien zijn de data toegankelijk voor mensen die blind of slechtziende zijn. De richtlijnen van het WAI (Web Accessibility Initiative) worden toegepast (zie hoofdstuk 5). Blinden die webpagina's laten voorlezen met 'text to voice'-software, of anderen die ze laten omzetten in braille kunnen probleemloos de data lezen.

Maar wat meer is: onze format en basisstructuur van de XHTML-documenten laat toe in de oorspronkelijke data de nodige metadata aanduidingen te geven (zowel Dublin Core als andere, al dan niet gecombineerd). Eigenlijk wordt er bijna geen enkele restrictie opgelegd aan de omvang en inhoud van data en metadata die in hetzelfde document liggen opgeslagen. We besparen de toeleveranciers van (meta)data enorm veel knip- en plakwerk. De metadata kunnen automatisch verzameld worden in een centraal archief.

Bovendien hebben we een eenvoudige 'truuk' gevonden om deze XHTML-documenten te laten invullen door leken die de XHTML-standaard zelf niet beheersen. Overigens op zo'n manier dat men het document kan laten circuleren tussen verschillende auteurs, waarbij de ene de bijdrage van de andere corrigeert of aanvult, een praktijk die heel veel voorkomt in het kader van duurzame ontwikkeling en bij de ontwikkeling van indicatoren. In dit rapport beschrijven we de methode zoals die toegepast kan worden op PC's die MS Windows draaien. Dat is alsnog een meerderheid in ons land. Let wel: we produceren een XHTML-document dat gemakkelijk kan overgedragen worden op andere Operating Systems zoals Linux en anticipeerden dus op aansluiting bij Open Source initiatieven. Maar we focusten dus wel in eerste instantie op het grootste cliënteel dat alsnog verknocht is aan Microsoft.

In feite bekomen we een methode die het voor een organisatie mogelijk maakt langzaam maar zeker aan te sluiten bij de XML-toepassingen, al dan niet in de Open Source-wereld. De organisatie produceert zijn data en voorziet de documenten van de nodige codes die duiden wat Dublin Core en andere metadata zijn. Deze data kunnen op de eigen website worden geplaatst en zijn toegankelijk voor mensen met een handicap.

Hoofdstuk 6 belicht een optie voor het geautomatiseerd verzamelen van de metadata voor opslag in gecentraliseerde archieven waar zoekfuncties en andere diensten kunnen worden aangeboden. Er zijn twee belangrijke aspecten behandeld: enerzijds het feit dat men in het kader van duurzame ontwikkeling in het algemeen, en IDO in het bijzonder moet opletten voor de-contextualisering. Een bepaalde databron (een beschrijving van een indicator bijvoorbeeld) moet - voor een goed begrip - vaak in een bredere context van informatie worden gezien. Daarom hebben we een elementaire procedure ontwikkeld waarmee men een archief kan informeren over de samenhang tussen verschillende databronnen. In het zesde hoofdstuk belichten we eveneens het management van de (meta-)data die bij verschillende organisaties worden ingezameld in XHTML-format. Daarbij dient opgemerkt dat eerst geprobeerd is om dit systeem enkel en alleen met XML-standaarden te laten draaien (DOM, XSLT e.d.). Echter hebben we

moeten inzien dat de kans groot is dat de XHTML-documenten die organisaties aanbieden niet altijd perfect zullen zijn (*valid* volgens XHTML-richtlijn). Men kan deze documenten weliswaar weigeren, maar dat is geen gebruikersvriendelijk optie.

Omdat de elementaire operaties die men op de ingezamelde gegevens wil toepassen voorzien zijn in een doorsnee relationeel database management systeem (RDBMS) hebben we uiteindelijk geopteerd voor het gebruik van dergelijke software die nu ook in Open Source distributies beschikbaar is (AMP-distributies: Apache als server, MySQL als database en PHP of Perl als scripttaal). Het gebruik van een RDBMS heeft ons overigens veel minder beperkingen opgelegd dan we lange tijd dachten. Het feit dat de XHTML-bronnen kunnen verschillen in lengte en samenstelling bleek uiteindelijk geen hinderpaal te zijn –mits men de 'records' in de database anders samenstelt dan gewoonlijk (één record per metadata-element in plaats van een record per metadata-set). Door het invoeren van de gegevens in een RDBMS komt men sneller tot een robuust informatiesysteem dan met 'zuivere' XML software die toch wel discipline vraagt van de toeleveranciers van de data (terzake hadden we enorm veel problemen met de 'encoding' die door toeleveranciers van data gemakkelijk fout aangeduid wordt). Let wel: een overstap naar 'zuivere' XML-data-managementsystemen is mogelijk.

De rest van dit rapport bestaat uit een reeks bijlagen met 'nevenproducten' van het onderzoek. In elk geval nodigen we de lezer uit onze website te bezoeken waar onder het onderwerp data- en metadatabeheer eventuele verdere ontwikkelingen worden belicht en gedemonstreerd.

Zie: <http://cdonet.UGent.be/issd/> en/of <http://www.belspo.be/issd/>

I. Duurzame ontwikkeling, indicatoren, data en metadata

I.1 Het nut van (informatie-uitwisseling inzake) IDO

In dit rapport gaat de aandacht naar gegevensbeheer en -uitwisseling inzake Indicatoren voor Duurzame Ontwikkeling (IDO).

IDO zijn meetbare -observeerbare - parameters die moeten toelaten een antwoord te geven op vragen als:

- Wordt de wereld, ons land, onze stad of ons bedrijf duurzamer of niet?
- Is dit product of productieproces duurzaam?
- Is deze belegging duurzaam?

'Duurzaam' is een eigenschap of kwaliteit die men nastreeft, maar die kwaliteit kan niet direct gemeten worden. Daarvoor is ze te complex en men zal ze dan ook ontrafelen in deelaspecten (deelkwaliteiten) die wel direct gemeten kunnen worden. Zo kan men bijvoorbeeld afspreken dat de duurzaamheid van een stad wordt gemeten middels een aantal waarnemingen inzake de kwaliteit van lucht, water en bodem (ecologische indicatoren), een aantal waarnemingen inzake tewerkstelling en inkomen (economische indicatoren) en een aantal waarnemingen inzake huisvesting, sport- en cultuurbeleving, *et cetera* (sociaal-culturele indicatoren). IDO worden dus steeds in groep gebruikt. Elke indicator op zich zegt weinig maar een collectie van indicatoren kan wel een betrouwbaar antwoord geven op de vraag: Is X duurzaam? Dat kan overigens nog steeds een genuanceerd antwoord zijn zoals: Op ecologisch en economisch vlak scoort de stad vrij goed, maar op sociaal vlak minder?

Indien men uit alle observaties één eindconclusie wil distilleren dan zal men wellicht gebruik maken van technieken die bekend staan als multicriteria-evaluatie, (evaluatie op basis van meerdere criteria of deelaspecten of deelkwaliteiten).

Wereldwijd vindt men talloze projecten die focussen op de ontwikkeling van IDO. De voornaamste argumenten om IDO te ontwikkelen zijn:

- IDO helpen bij het duiden van het concept 'Duurzame Ontwikkeling'. Door de selectie van indicatoren komt men tot een concreter invulling van het vage begrip 'duurzame ontwikkeling'. Indicatoren liggen dus in het verlengde van visievorming¹.
- Duurzame Ontwikkeling vereist een monitoring- of evaluatie-instrument. Men moet op regelmatige tijdstippen kunnen beoordelen of men de ambities realiseert. IDO kunnen de besluitvorming ondersteunen. Uit observaties kan men eventueel afleiden waar beleidsprioriteiten moeten worden gelegd.
- Middels IDO kan men nagaan welke acties het beoogde resultaat opleveren, welke beleidsdaden al dan niet succesvol waren. IDO kunnen dus ook leerprocessen ondersteunen.

¹ Meadows D. (1998); *Indicators and Information Systems for Sustainable Development - A report to the Baltaon Group*; The Sustainability Institute

Gouzee N., Mazijn B., BillHarz S. (eds.) (1995); *Indicators for Sustainable Development - Report of the Workshop of Ghent, Belgium, 9-11 January 1995, submitted to the UN Commission on Sustainable Development*; Federal Planning Office of Belgium, Brussels

Bruyninckx H. (1998), *Duurzame ontwikkeling: op zoek naar operationele definities en indicatoren - DWTC Working Paper Nr 1*, HIVA-KUL (leuven) / DWTC (Brussel)

Kuik O., Verbruggen H. (eds.) (1991), *In search of Indicators of Sustainable Development*, Kluwer Academic Publishers, Boston

"Indicators of sustainable development need to be developed to provide solid bases for decision-making at all levels and to contribute to a self-regulating sustainability of integrated environment and development systems". (VN, Agenda 21)²

IDO worden 'gepromoot' door internationale organisaties zoals de Verenigde Naties (UN) en de Organisatie voor Economische Samenwerking en Ontwikkeling (OECD).

Onderstaande voorbeelden zijn exemplarisch. Het is niet de bedoeling hier een 'representatieve' lijst te geven, en evenmin om de voornaamste of belangrijkste initiatieven te duiden.

- **CSD: UN-Commission on Sustainable Development (CSD)**

<http://www.un.org/esa/sustdev/isd.htm>

"Indicators for monitoring progress towards sustainable development are needed in order to assist decision-makers and policy-makers at all levels and to increase focus on sustainable development. Beyond the commonly used economic indicators of well-being, however, social, environmental and institutional indicators have to be taken into account as well to arrive at a broader, more complete picture of societal development."

- **Millennium Development Goals**

<http://www.developmentgoals.org/>

"At the Millennium Summit in September 2000 the states of the United Nations reaffirmed their commitment to working toward a world in which sustaining development and eliminating poverty would have the highest priority. The Millennium Development Goals grew out of the agreements and resolutions of world conferences organized by the United Nations in the past decade. The goals have been commonly accepted as a framework for measuring development progress."

- **Measuring Development Progress** (OECD's Development Assistance Committee, DAC)

<http://www1.oecd.org/dac/indicators/>³

"A ground-breaking collaborative process, jointly led by bilateral donors through the DAC, in partnership with the UN, the World Bank and the IMF and with participation from developing countries, has agreed a core set of indicators for measuring progress towards the international development goals. This set was discussed in detail at indicator meetings held at OECD Paris in 1997 and 1998, together with working group meetings. A forum in Paris in March 2000 refined and adopted the set as shown overleaf."

- **OESO: Organisatie voor Economische Samenwerking en Ontwikkeling**

<http://www.oecd.org/>

Op deze site kan men op de homepage het thema 'sustainable development' selecteren. Daar vindt men verschillende deelthema's waar vaak indicatoren worden ontwikkeld.

In 2001 verscheen een document waarin de OECD haar plannen met betrekking tot IDO bondig beschrijft: *"Policies to Enhance Sustainable Development"* (OECD, Paris, 2001)

<http://www.oecd.org/pdf/M00008000/M00008866.pdf>

Vele regio's, landen en steden hebben hun IDO ontwikkeld.

Zie onder andere:

² Zie hoofdstuk 40 ('Information for Decision-Making') van Agenda 21:

<http://www.un.org/esa/sustdev/documents/agenda21/english/agenda21toc.htm>

Een Nederlandse vertaling van Agenda 21: <http://www.ddh.nl/agenda21/rioverklaring/>

Een Franse vertaling: http://www.agora21.org/rio92/A21_html/A21_1.html

³ Deze website is recent verwijderd en de bezoeker wordt doorverwezen naar de Millennium Development Goals

- **Baltic 21**

<http://www.ee/baltic21/>

"In 1996, the Prime Ministers of the Baltic Sea Region took the initiative to develop an Agenda 21 for the Baltic Sea Region....In order to monitor the development in the Baltic Sea Region, a set of core indicators has been elaborated for each goal, and statistics compiled."

- **United Kingdom - Indicators of Sustainable Development**

<http://www.sustainable-development.gov.uk/indicators/index.htm>

"Indicators are central to the monitoring and reporting of progress towards sustainable development. They are also powerful tools which can help focus public attention on what sustainable development means and to give a broad overview of whether we are achieving a better quality of life for everyone, now and for generations to come....This site covers the indicators that have been developed in the United Kingdom both at national, regional and local levels. It also includes links for reference to various key international initiatives and organisations."

- **Duurzaamheidsbalans Noord-Brabant Nederland**

<http://fsw.kub.nl/telos/documenten/duurzaam.pdf>

- **Stadsmonitor voor leefbare en duurzame Vlaamse steden**

<http://www.thuisindestad.be/html/monitor/index.html>

- **Duurzaamheidsbarometer regio Kortrijk**

<http://www.leiedal.be/duurzaamheidsbarometer/>

Ook in de NGO's en wetenschappelijke onderzoekswereld wordt het instrument gewaardeerd. Zie onder andere:

- **Monitoring and Evaluation Initiative of IUCN**

<http://www.iucn.org/themes/eval/sustassess.htm>

Wat dit initiatief betreft vindt men een goede uiteenzetting in volgend document: "Assessing Progress Toward Sustainability: Workshop on System Assessment Method" (Islamabad, Pakistan, 1999)

<http://www.iucn.org/themes/eval/documents/islamabad.pdf>

- **Dashboard of Sustainability**

http://iisd.org/cgsdi/dashboard_dsply.htm

<http://esl.jrc.it/dc/index.htm>

"The Dashboard of Sustainability ... is an instrument panel designed to inform decision-makers, the media and the general public on the status of a nation's progress toward, or away from, sustainability. This concept grew out of the work of the Consultative Group on Sustainable Development Indicators (<http://iisd1.iisd.ca/cgsdi/members.htm>), which has been engaged in critical assessments and design discussions since 1996. The Dashboard project is part of the sustainability indicator initiative of the Bellagio Forum for Sustainable Development (<http://www.bfsd.org/>), one of the main funders of the work".

Om ieder misverstand weg te werken: dit rapport gaat niet verder in op het nut, het belang of de rol die men hecht aan IDO. Daarover zijn reeds vele publicaties verschenen.

Anderzijds dient wel de aandacht gevestigd op het feit dat men Indicatoren voor Duurzame Ontwikkeling niet los mag zien van de bredere context van visievorming en/of besluitvorming. IDO zijn een instrument en geen finaliteit op zich. Toch blijkt er een brede consensus te bestaan inzake het belang van dit instrument.

Het is dan ook bijna vanzelfsprekend om initiatieven te ontwikkelen inzake de uitwisseling van gegevens terzake. De opdrachtgever van dit project, het DWTC, had reeds eerder een project afgerond waarin een database werd aangelegd met gegevens inzake personen, organisaties, projecten, documenten en databanken die handelen over Duurzame Ontwikkeling. Zie:

<http://www.belspo.be/issd/>

Dit Informatiesysteem voor Duurzame Ontwikkeling (ISDO) kan worden uitgebreid met een databank 'Indicatoren voor Duurzame Ontwikkeling', zo luidde de ambitie bij aanvang van dit project.

Voor de hand liggende vraag is dan wat de inhoud en structuur van de gegevens over de IDO moet zijn. Kan men een standaardfiche of steekkaart ontwikkelen die tegemoet komt aan de noden of wensen van de talloze eindgebruikers. Moet men een steekkaart ontwikkelen voor de beschrijving van één indicator en/of een collectie van indicatoren?

Vooreerst maken we hier een korte evaluatie van de stand van zaken met betrekking tot informatie-uitwisseling omtrent IDO. Het is een 'state of the art' zoals geregistreerd in de periode 2001-2003. Het is best mogelijk dat er nadien – ergens op de wereld - initiatieven werden genomen die hier niet vermeld worden.

I.2 Het World Wide Web, Metadata en Duurzame Ontwikkeling

De ruime belangstelling voor IDO, moet als verklaring dienen voor de belangstelling die in dit rapport gehecht wordt aan informatie-uitwisseling terzake.

Het Internet en het World Wide Web verleent toegang tot een groot deel van de informatie met betrekking tot duurzame ontwikkeling in het algemeen en IDO in het bijzonder. Dit is geen toeval, maar wel degelijk onderdeel van het beleid. Openbaarheid van informatie wordt gezien als een belangrijk aspect van duurzame ontwikkeling.

Evenwel blijft het moeilijk om op efficiënte en effectieve wijze die informatie te selecteren die men zoekt. Nog moeilijker is het zeker te zijn dat men alles gevonden heeft wat relevant is, gegeven een bepaalde informatiebehoefte. Toegankelijkheid of beschikbaarheid van de gegevens is één zaak, men moet ze ook kunnen opsporen en vinden als men ze nodig heeft.

Er zijn verschillende manieren om het opsporen van informatie te ondersteunen. Hieronder lichten we enkele praktijken toe, en focussen daarbij op het gebruik van metadata die de efficiëntie en effectiviteit van het zoeken naar informatie moeten vergroten.

Google, Yahoo en andere links

Google, Yahoo, Altavista en andere webrobots zijn computerprogramma's die webpagina's bezoeken en woorden uit die pagina's opslaan in een trefwoordenbank. De webrobots kunnen ook de links in de webpagina's volgen om automatisch andere webpagina's te ontdekken.

Vele gebruikers raadplegen de trefwoordenbank om informatie te zoeken. De voor en nadelen zijn gekend. Sommige trefwoorden leveren vele honderden 'links' en het is niet vanzelfsprekend dat de meest relevante informatiebronnen in de ellenlange lijsten met links verschijnen. Dit fenomeen is gevolg van het feit dat de webrobots uit de webpagina's trefwoorden selecteren op basis van criteria die niet altijd een goede oogst opleveren. Zo zal bijvoorbeeld een bepaald woord als trefwoord gecatalogeerd worden omdat het bovenaan op een pagina voorkomt of vaak in het document voorkomt. Het spreekt vanzelf dat robots op die manier gemakkelijk 'misleid' kunnen worden. Een webrobot kan in de meeste gevallen niet met zekerheid bepalen wat de titel is van een document, of de auteur, en dergelijke.

Enkel wanneer de producent van de informatie terzake expliciet informatie toevoegt aan het document -op een gestandaardiseerde manier - zal de robot daar eventueel gebruik van maken en die trefwoorden oogsten die door de auteur expliciet als trefwoord werden vermeld. Hier melden we enkel dat men dan metadata toevoegt aan de data (gegevens) die men op het net publiceert. Die metadata - gegevens over de gegevens - kunnen een lijst van trefwoorden zijn, een korte samenvatting, etc. Die metadata zijn niet zichtbaar voor de doorsnee bezoeker van een website –tenzij die naar de zogenaamde broncode van webpagina's kijkt- , maar ze zijn wel 'zichtbaar' voor zoekrobots. Ondanks het feit dat er vandaag een brede consensus is gegroeid over de manier waarop dergelijke metadata moeten worden aangeboden in webpagina's, blijkt alsnog slechts een minderheid dit in praktijk te brengen. Deels allicht omdat deze techniek noch onvoldoende bekend is, maar mogelijks ook omdat het enige discipline en tijd vraagt.

Wie meer wil vernemen over het gedrag van webrobots, kan o.a. terecht bij:

<http://www.zoekprof.nl/> en meer bepaald: <http://www.zoekprof.nl/achtergronden/hoewerkt0.html>
<http://www.masterdiz.com/theturbomechanic/>
http://www.kuleuven.ac.be/facdep/social/onderwijs/werkcollege/zoeken/www_zoekinstr.htm
<http://www.esat.kuleuven.ac.be/teo/docarch/iv/2002/jun02/im162.htm#5>

Yahoo, Google e.a. leveren ook een service waar de links thematische geordend zijn. De links zijn gecatalogeerd in een hiërarchische structuur van directories en subdirectories (thematische folders en deelfolders).

De 'root' van deze structuur vindt men bij Yahoo op <http://dir.yahoo.com/>

Het thema 'duurzame ontwikkeling' komt meerdere malen voor in de hiërarchie. Zo is het een deelthema (sub-directory) van 'Business_and_Economy', maar eveneens een deelthema (sub-directory) van 'Society_and_Culture'.

http://dir.yahoo.com/Business_and_Economy/Business_to_Business/Environment/Sustainable_Development/

http://dir.yahoo.com/Society_and_Culture/Environment_and_Nature/Sustainable_Development/

Een gelijkaardige service vindt men ook bij Google, Altavista e.a.

<http://dir.google.com>
<http://dir.altavista.com>

Hoe dergelijke service kan georganiseerd worden kan men onder andere vernemen bij het Open Directory Project: <http://dmoz.org/>

"The Open Directory Project is a large public directory managed by Netscape. The ODP is maintained by a group of volunteer editors from around the world who evaluate sites for inclusion in the directory. The web pages selected by these editors are organized into a number of broad categories under which are many more specific subcategories. Google uses this hierarchy as the basis for its directory".

Google neemt de gegevens van ODP over.

Er is ook een Belgisch Open Directory Project: <http://dmoz.be>

Het principe is in feite eenvoudig. Het domein waarover men informatie - of links naar informatie - wil verzamelen wordt opgedeeld in deelthema's en kandidaat redacteurs kunnen dan het beheer van een deelthema - een directory - op zich nemen. Deelthema's kunnen vanzelfsprekend verder opgedeeld worden in – subdirectories - waarbij een andere redacteur deze 'deeltaak' op zich neemt.

Hier is dus niet een 'domme' webrobot actief, maar een 'intelligente' mens die niet alleen woorden kan zien en tellen, maar ook begrijpen en interpreteren en dus ook met verstand van zaken kan oordelen of een trefwoord wel degelijk representatief is voor de inhoud.

Bedoeling is dus dat verschillende redacteurs - iedereen die zich geroepen voelt - gaan samenwerken en een grote collectie van links verzamelen, netjes geordend in thematische folders.

Merk op dat het Open Directory Project meer is dan alleen maar een collectie links. Zo kan men in iedere directory een beschrijving van het thema raadplegen; men kan met trefwoorden gaan zoeken in de collectie; men kan zijn eigen site toevoegen (melden aan de redacteur), etc.

Nogal wat onderzoekers, onderzoeksgroepen, ngo's of netwerken die actief zijn in het domein duurzame ontwikkeling produceren en beheren hun eigen pagina's met links. Ook deze kunnen thematisch opgedeeld zijn, en men hoeft zich dan ook niet te houden aan de directorystructuur van het ODP. Overigens kan men nog steeds zijn eigen service - als geheel - aanmelden in het ODP. Zie o.a.

- **WWW Vitual Library Sustainable Development**

<http://www.ulb.ac.be/ceese/meta/sustvl.html>

- **Sustainable Development and the Americas Index**

<http://www.txinfinet.com/mader/ecotravel/sustain.html>

- **Duoweb**

<http://www.ecomarkt.nl/ecoduo/>

Search engine on sustainable development (Zoekmachine voor duurzame ontwikkeling)

- **Sustainable Development Online**

<http://sd-online.ewindows.eu.org/>

- **Sustainable Production and Consumption**

<http://www.iisd.ca/linkages/consume/other.html>

Al deze initiatieven focussen in eerste instantie op het aanbod van webadressen of URL's (Uniform Resource Locators), waarbij slechts een minimum aan informatie wordt voorzien: men vermeldt de URL (<http://...>), een titel en eventueel een korte beschrijving of enkele trefwoorden.

De rangschikking in thematische folders is een laatste hint op basis waarvan de gebruiker kan oordelen of de link voor hem/haar een interessante informatiebron kan zijn.

Deze projecten zijn gebaseerd op een minimum aan metadata: voor elke informatiebron houdt men het vaak op maximum vier metadata-elementen:

- URL (web-adres)
- titel
- beschrijving
- trefwoorden.

Ook in zogenaamde **gateways** werkt men vaak met een beperkt aantal metadata-elementen:

- **Sustainable Development Gateway - SD Gateway**

<http://sdgateway.net/>

- **Eldis Gateway to Development Information**

<http://www.eldis.org/>

Metadata zijn data of gegevens die andere data of gegevens beschrijven. In bovenstaande voorbeelden zitten de data dus op de websites waarnaar men verwijst, en de metadata zijn de gegevens die opgeslagen liggen in de 'Open Directories' of de pagina's met links.

Zoals in vermelde voorbeelden worden metadata meestal ontwikkeld voor de ondersteuning van informatie-opsporing. De metadata worden dan opgeslagen in een databank waarin men de data kan zoeken en selecteren door het gebruik van (combinaties van) trefwoorden.

Het gebruik van metadata heeft als belangrijk voordeel dat men tot afspraken kan komen over de aard en hoeveelheid van metadata en de manier waarop deze worden uitgewisseld, wat niet impliceert dat men het eens moet zijn over de aard en vorm van de data zelf. Met één metadatastandaard kan men zeer uiteenlopende informatiebronnen beschrijven (CD's én boeken én websites bijvoorbeeld).

Vandaag wordt informatie over duurzame ontwikkeling en IDO op velerlei manieren gepubliceerd. Men produceert Word-documenten, PDF-documenten, websites en databases.

Het is praktisch onmogelijk om zoekrobots te ontwikkelen die op efficiënte wijze in dit diverse aanbod van informatiebronnen gaan zoeken naar specifieke trefwoorden. Het kan, maar bij een enorm groot aanbod van informatie - zoals dat vandaag het geval is in het domein duurzame ontwikkeling - zou dat een tijdrovende onderneming worden.

Men zou deze processen kunnen versnellen door afspraken te maken over de manier waarop alle data worden opgeslagen, maar dit zou een ingrijpende wijziging in de praktijk van vandaag vergen. Gemakkelijker is om de ondersteuning van informatie-opsporing te baseren op afspraken over het gebruik en de uitwisseling van metadata, waarbij de gangbare praktijken inzake databeheer en uitwisseling (voorlopig) gehandhaafd kunnen blijven.

Metadata betreffende Indicatoren voor Duurzame Ontwikkeling

Er bestaan reeds enkele projecten die middels het gebruik van metadata de 'opsporing' van IDO of IDO-projecten ondersteunen. We vestigen hier de aandacht op twee zulke projecten die een verschillende aanpak demonstreren. Deze voorbeelden worden hier al belicht omdat ze het gebruik van metadata demonstreren, een informatie-instrument dat verder in dit rapport nog uitgebreid aan bod komt. Het was niet de bedoeling om hier alle bestaande initiatieven terzake op te sommen. Weliswaar kan worden gesteld dat inzake IDO slechts enkele initiatieven konden worden opgespoord. We vermelden er twee, voornamelijk om aan te geven dat er verschillende opties mogelijk zijn, zowel voor wat betreft de inhoud als de presentatievorm en de daaraan gekoppelde zoekfaciliteiten.

- **Database van Sustainable Measures (Maureen Hart)**

<http://www.sustainablemeasures.com/Database/index.html>

Deze database laat toe om middels trefwoorden een specifieke indicator op te sporen. Daarbij wordt dan vermeld wie deze indicator gebruikt of waar men terzake meer informatie kan vinden.

Indien men evenwel een trefwoord als 'city' gebruikt, met de bedoeling IDO-projecten op stedelijk niveau op te sporen, dan krijgt men een lijst van indicatoren inzake 'electri-city', 'capa-city', e.d. Een indicator met de naam 'Percent of people who live and work in city' verschijnt ook, maar de geproduceerde lijst focust dus niet uitsluitend op stedelijke initiatieven.

Hoewel er in deze database gegevens zitten van Minnesota zullen deze niet te voorschijn komen als men het trefwoord 'Minnesota' intikt.

De zoekfunctie kan gebaseerd zijn op trefwoorden die in de naam van een indicator moeten voorkomen. Daarnaast worden thematische lijsten van indicatoren aangeboden.

Indien men bijvoorbeeld het trefwoord 'crime' intikt dan verschijnt een lijst van indicatoren met het trefwoord in hun naam, bijvoorbeeld:

'Crime victims as percent of population'

Daar wordt dan bij vermeld wie deze indicator gebruikt en wat hun gegevensbron is.

Wordt een gebruiker van een indicator aangeklikt dan verschijnen bijvoorbeeld volgende gegevens:

- Jacksonville Community Council Inc.
- 2434 Atlantic Boulevard, Suite 100, Jacksonville, Florida 32207
- Phone: (904) 396-3052
- E-mail: jcci2@leading.net
- Web site: www.jcci.org
- *Quality of Life in Jacksonville: Indicators for Progress* includes graphs of trends for most indicators for most years since 1983. One of the longest running community indicator projects, the Jacksonville report has been an inspiration for many other communities. The executive summary of the report, which is intended for general distribution, does not include detail on data sources or precise indicator definitions although the titles are quite descriptive. The full report, which is designed for planners, decision makers, researchers, etc. is very informative. Both documents are made available free locally.

Kortom:

- de naam van de initiatiefnemer
- adres, telefoon en email, contact
- URL van de website
- korte beschrijving van het project

Hier vinden we dus reeds meer metadata-elementen dan gebruikelijk in de pagina's met links of de Open Directory Projecten die eerder zijn belicht.

Zoals bovenstaand voorbeeld aangeeft zijn metadata enerzijds een soort 'abstract', een korte beschrijving van de data, maar tegelijk ook bedoeld als een gids die aangeeft waar men de data vindt, wie ze heeft gemaakt of beheerd.

In heel veel gevallen zijn metadata zodanig samengesteld dat ze met een minimum aan gegevens een maximale service kunnen geven aan de personen die data opzoeken. Tegelijk wil men meestal ook voldoende informatie geven, zodat de lezer kan beslissen of de data datgene te bieden hebben wat hem/haar interesseert.

We bekijken nu een tweede voorbeeld waar ook weer metadata over IDO-projecten zijn ingezameld. De Engelse paragrafen zijn letterlijk overgenomen van de site.

- **Compendium of Sustainable Development Indicator Initiatives**

<http://www.iisd.org/measure/compindex.asp>

"The compendium provides a comprehensive and up-to-date information base of sustainable development indicator initiatives being carried out at the international, national and provincial/territorial/state levels. Some projects at the regional/local and community levels are included. The compendium also contains an annotated bibliography of indicator-related literature and a collection of real-life indicator examples."

Het compendium is ontwikkeld door het International Institute for Sustainable Development, Environment Canada, Redefining Progress, de World Bank en de United Nations Division for Sustainable Development.

"The compendium seeks to:

- *Improve communication among the various stakeholders in sustainable development to promote the sharing of experiences, methods and approaches on indicator development and use for mutual benefit;*
- *Facilitate the harmonization of indicator development approaches and indicator sets;*
- *Avoid duplication of efforts and facilitate the integration of monitoring, data analysis and reporting activities;*
- *Provide governments, NGOs, the private sector and the public with access to a pool of experts working on indicator development;*
- *Help identify areas of future research where indicator work is required;*

- *Provide information on a wide range of publications related to developing indicators and indices for sustainable development".*

De fiche waarmee een IDO-project wordt beschreven bevat volgende metadata-elementen:

Name of organization**Contact person****Title****Address****Fax****Phone****E-mail****Internet home page address****Title of initiative**

Geographic Scope : *The core geographic area that the indicator set applies to. This may be either a spatial unit with jurisdictional boundaries, such as a state or city, or an area with physical boundaries, such as an ecosystem or a watershed. This definition should hold even if some of the indicators report on issues outside the scope of the majority of the indicators.*

Hier kiest men een term uit een beperkt aanbod:

- *corporate*
- *global*
- *municipal*
- *national*
- *regional*
- *site/specific*
- *state/provincial*

Project objectives: *What do the indicators measure in the broad sense; what is the key purpose of reporting?*

Framework for indicator set: *Clarification of which conceptual model or framework is used in the identification and organization of indicators. Among several others, the following frameworks may be referred to: pressure-state-response, environmental media based (e.g. air, land, biota, etc.), human made, social-human capital, environmental-economic-social, economic sector base, etc.*

Partnerships and stakeholder consultations: *Which stakeholders and partners are involved in the development process?*

Progress to date/future areas of work: *Summary of which indicator(s) have been developed so far, brief descriptions of the indicator(s), the key methodological aspects and conclusions including proposed next steps.*

Timeframe: *Start and completion dates (if applicable), frequency of updates*

Presentation of the information and indicators: *Description of how the information is presented (e.g., graphics, tables, figures); in which format the product is available (e.g., diskettes, bulletins, publications); and identification of the software in which the information has been developed (e.g., Access, Lotus, dBase, ASCII, Excel)*

Publications: *Full bibliographic reference list of relevant publications.*

Initiative Location: hier kunnen één of meerdere locaties gekozen worden uit volgende lijst:

- *Australia, New Zealand and Pacific Islands*
- *Caribbean*
- *Central America*
- *Central and Eastern Europe*
- *East Africa*
- *East Asia*

- Middle East
- North Africa
- North America
- Russian Federation
- Scandinavia
- South America
- South Asia
- Southeast Asia
- Southern Africa
- West Africa
- Western Europe
- Worldwide

Het aantal metadata-elementen is uitgebreider dan in het project van Maureen Hart dat eerder werd toegelicht. De bijhorende zoekfaciliteiten laten nu toe om gericht te zoeken naar projecten die bijvoorbeeld in 'Scandinavië' zijn uitgevoerd én zich richten op het niveau 'municipal'.

Wat het compendium betreft dient overigens ook opgemerkt dat elke bezoeker van de site zijn of haar gegevens zelf kan invoeren (of wijzigen), wat in het eerste voorbeeld blijkbaar niet het geval is.

Wanneer men in het compendium zoekt met het trefwoord 'crime' dan vindt men ook daar een reeks projecten die blijkbaar indicatoren voor dit onderwerp hebben gebruikt.

Dit project is ondertussen 'verhuist' naar <http://www.sustainabilityindicators.org>, de website van ISIN

"The International Sustainability Indicators Network is a member-driven organization that provides people working on sustainability indicators with a method of communicating with and learning from each other. Through listserv discussions, virtual and in-person meetings, and special programs and trainings, the Network facilitates shared learning and development among sustainability indicators practitioners and others."

Metadata betreffende duurzame ontwikkeling in het algemeen

Naast deze initiatieven die focussen op IDO, zijn er andere die meer gedetailleerde metadata ontwikkelen voor de beschrijving van informatiebronnen inzake 'duurzame ontwikkeling' in het algemeen. Opnieuw vermelden we er slechts twee, voornamelijk om aan te geven dat er zeer uiteenlopende manieren van aanpak mogelijk zijn.

- **Informatie Systeem voor Duurzame Ontwikkeling**

<http://www.belspo.be/issd>

"Het Informatiesysteem voor Duurzame Ontwikkeling (ISDO) is een initiatief van de DWTC (Belgische Federale Diensten voor Wetenschappelijke, Technische en Culturele Aangelegenheden) in het kader van de Ondersteunende Acties bij het "Plan voor wetenschappelijke ondersteuning van een beleid gericht op duurzame ontwikkeling" (PODO).

*Het ISDO is in de eerste plaats een **zoekinstrument** dat aan eenieder de mogelijkheid wil bieden op een snelle en eenvoudige wijze bruikbare gegevens en gegevensbronnen inzake duurzame ontwikkeling te identificeren en te lokaliseren. In bepaalde gevallen laat het instrument rechtstreekse consultatie van de gegevens toe via internet.*

*Daarnaast is het een **communicatiekanaal** voor de diverse actoren in het domein van duurzame ontwikkeling (overheden, NGO's, de wetenschappelijke wereld,...) om hun activiteiten en de informatie waarover zij beschikken op ruime schaal bekend te kunnen maken".*

Het ISDO van de Belgische Dienst voor Wetenschappen, is een database met records (steekkaarten, fiches) die specifiek ontworpen zijn voor de beschrijving van organisaties, projecten, publicaties, of personen.

Onderstaand voorbeeld is een voorbeeld van een steekkaart met metadata die een organisatie beschrijven.

ISDO-Metadata van een organisatie:

Organisatie:	Universiteit Gent - Centrum voor Duurzame Ontwikkeling
Type:	Academisch, wetenschappelijk
Naam_fr:	Université Gand - Centre de Développement Durable
Naam_nl:	Universiteit Gent - Centrum voor Duurzame Ontwikkeling
Naam_en:	Ghent University - Centre for Sustainable Development
Acroniem:	CDO
Acroniem_fr:	CDD
Acroniem_nl:	CDO
Acroniem_en:	CSD
Algemene beschrijving_nl:	<p>Het Centrum voor Duurzame Ontwikkeling werkt met kennis en vaardigheden vanuit vijf vakgroepen van de universiteit Gent. Dit zijn de vakgroep Wijsbegeerte en Moraalwetenschappen, Studie van de Derde Wereld, Toegepaste ecologie en Milieubiologie, Burgerlijk Recht en Algemene Economie.</p> <p>Het Centrum voor Duurzame Ontwikkeling heeft tot doel op een multidisciplinaire wijze duurzame ontwikkeling als beleidskader verder uit te bouwen.</p> <p>De werking is gericht op onderzoek naar en ondersteuning van rationele besluitvorming en doelmatigheid van beleidsinstrumenten, in functie van lange termijndoelstellingen, in een geïntegreerd perspectief met maximale aandacht voor participatie en communicatie.</p> <p>In de multidisciplinaire aanpak dienen ondermeer volgende vijf hoofdlijnen aanwezig te zijn : ecologische, sociale, economische, institutionele en ethische facetten.</p> <p>De belangrijkste onderzoeklijnen zijn:</p> <ol style="list-style-type: none"> 1. Duurzame ontwikkeling op elk beleidsniveau, met bijzondere aandacht voor indicatoren 2. Milieugebruiksruimte, voorraadbeheer en grondstoffen 3. Product- en procesevaluatie 4. Mobiliteit en verkeersveiligheid 5. Ontwikkelingssamenwerking en duurzame ontwikkeling in de Derde Wereld
Algemene beschrijving_en:	<p>The Centre for Sustainable Development (CSD) is an interdisciplinary research group drawing on the knowledge and expertise of nine departments at Ghent University.</p> <p>The aim of the centre is to use an interdisciplinary methodology for the elaboration on the sustainable development as a social framework. This approach includes, but is not limited to, research in the following five keyfields: social, economic, ecological, institutional and ethical.</p> <p>The CSD undertakes three kinds of tasks in order to achieve its aims: research, education and provision of a social service.</p> <p>Within the framework of the research task there are three topics in which basic and applied research is carried out.</p> <p>* Education is not restricted to our own lectures, but is broadened to introduce and integrate concepts concerning sustainable development into existing lectures and courses, furthermore an effort is made to inform and sensitize social target groups of all descriptions (from outside the university).</p> <p>* Within the framework of the provision of a social service, studies for the preparation of policy and short-term advisory duties are carried out for different levels of government (from international to local).</p>
Sleutelwoorden GEMET:	energy, sustainability, prevention, labelling, ecolabel, documentation centre, environmental data, environmental legislation, Agenda 21, building material, energy resources, social participation, pedagogy, wastes, association, social movement, social indicators
Vrije sleutelwoorden:	carrying capacity, speed management, road security, chain management, prevention, stakeholders, factor 4, environmental utilisation space
Relatie met Agenda21:	internationale samenwerking, rationeel energieverbruik, afvalpreventie, verkeer en vervoer, communicatie
Administratieve	België

geografische dekking:	
Taal(talen) van de organisatie:	Nederlands
Adres:	Poel 16
Postcode:	9000
Gemeente:	Gent
Land:	België
Telefoon 1:	+32 (0) 9 264 82 10
Fax:	+32 (0) 9 264 83 90
E-mail:	cdo@UGent.be
URL:	http://cdonet.UGent.be
Type adres:	Hoofdadres
Contactpersoon:	Johan De Mol
Relaties met andere gegevens:	
Organisaties	Universiteit Gent - Centrum voor Milieueconomie en Milieumanagement
Personen	Doom Ruddy; onderzoekspromotor
Projecten	Onderzoek van de mogelijkheden en beperkingen van het concept 'milieugebruiksruimte'
Projecten	Een integrale benadering van de ketenanalyse ten behoeve van ketenbeheer door bedrijven

De laatste gegevens zijn 'links' naar andere steekkaarten in het ISDO-systeem die gerelateerd zijn aan deze steekkaart. Het ISDO wordt beheerd met een relationele databank.

De andere steekkaarten hebben andere metadata-elementen of velden. Toch blijft het systeem vrij eenvoudig van opzet en een bezoeker van de ISDO-site vindt dan ook snel zijn weg in het systeem.

Men kan in ISDO zoeken middels het opgeven van trefwoorden en daarbij kiezen:

- ofwel zoekt men het trefwoord in alle velden of metadata-elementen van alle beschikbare steekkaarten
- ofwel zoekt men het trefwoord in slechts enkele velden of metadata-elementen van een beperkt type fiches

Een CDO-medewerker stelde vast dat de gegevens die opgeslagen zijn bij ISDO niet geheel in overeenstemming zijn met datgene wat het CDO over zichzelf vertelt op haar eigen website (<http://cdonet.UGent.be>). Het demonstreert alvast één punt: metadata kunnen een eigen leven gaan leiden en kunnen achterhaald raken wanneer zij het resultaat zijn van een project dat niet of moeilijk gecontinueerd wordt. Het feit dat de metadata bij ISDO niet geheel actueel zijn kan in dit geval gevolg zijn van onoplettendheid van het CDO zelf of van de beheerder van de centrale database. De schuldvraag moet hier niet worden opgelost. Maar de vaststelling dat metadata best voorwerp zijn van een continu 'onderhoud' is wel belangrijk.

Overigens zijn de metadata in ISDO wel voldoende update als men ze gebruikt als gegevens die het zoeken naar (bijvoorbeeld) organisaties ondersteunen. Via deze metadata kan het CDO nog steeds 'gevonden' worden. Maar de eindgebruiker moet wel naar de site van het CDO als hij/zij de meeste recente gegevens over het CDO wenst.

Heel veel metadata-initiatieven struikelen over het probleem actualisering omdat bij de startfase wel de nodige budgetten waren voorzien, maar na afloop van het project niet alleen de middelen op zijn, maar eventueel ook het betrokken personeel verdwijnt. Wie in metadata-archieven herhaaldelijk op links klikt die blijkbaar nergens meer naartoe leiden, kent het fenomeen en wellicht ook de frustraties. Vaak worden gegevens op het internet verplaatst naar andere servers, of websites worden geheel herschikt, met als gevolg dat metadata die elders opgeslagen liggen foute aanwijzingen geven.

Thesaurus: Gemet

Bij het beschrijven van gegevensbronnen kunnen deze getypeerd worden aan de hand van sleutelwoorden uit een vooraf gedefinieerde thesaurus of 'gecontroleerde woordenschat'.

In het kader van ISDO wordt een thesaurus van sleutelwoorden gebruikt (GEMET) om gegevens(bronnen) te beschrijven en om de metagegevens te consulteren.

GEMET (General European Multilingual Environment Thesaurus) werd ontwikkeld in opdracht van het Europees Milieu Agentschap ten behoeve van de Catalogue of Data Sources. Een thesaurus voor het

thema duurzame ontwikkeling was onbestaande. Hoewel de naam van de thesaurus GEMET suggereert dat het een milieuthesaurus betreft, komen een aantal andere dimensies van duurzame ontwikkeling (sociale, economische, institutionele aspecten) eveneens aan bod.

Zie:

- http://www.mu.niedersachsen.de/cds/etc-cds_neu/library/select.html
- <http://www.mu.niedersachsen.de/system/cds/>

In de eerste plaats is het de bedoeling met behulp van een thesaurus het taalgebruik te harmoniseren, ze meldt de ISDO-site. Dit vocabularium kan ook aangeboden worden aan de gebruiker die de metagegevens raadpleegt (zoeken via sleutelwoorden).

Tevens kan dus een eenduidig verband worden gelegd tussen de termen in verschillende talen. Iemand die dan zoekt op basis van een Engels trefwoord kan (eventueel) aanstippen dat hij/zij ook alle informatie wenst met trefwoorden die een vertaling zijn van het Engelse begrip.

Overigens kan men een bepaalde term ook koppelen aan 'verwante' begrippen, zodat het zoeken op het begrip 'zure regen' bijvoorbeeld ook de databronnen opspoort met het trefwoord 'verzuring'.

Daarnaast laten de systematische en/of thematische structuur van de thesaurus het toe de informatie te rangschikken en te ordenen.

Het raadplegen van de metadatabank met een thesaurus van sleutelwoorden is vaak complementair aan andere zoekfuncties.

Zoals bovenstaande projecten aantonen kan men de informatie-uitwisseling niet alleen ondersteunen met het gebruik van een al dan niet uitgebreid aantal metadata-elementen, maar ook met een thesaurus (gecontroleerde woordenschat).

Voorts dient opgemerkt dat zoals bij ISDO de metadata kunnen opgeslagen worden in een informatiesysteem waarin relaties worden gelegd tussen verschillende steekkaarten of fiches met metadata en dus ook tussen de respectieve bronnen die deze metadata beschrijven.

Dergelijke informatiesystemen kunnen zeer uiteenlopende ambities hebben. Een bijzondere toepassing vindt men bij het Massachusetts Institute for Technology - MIT:

- **Global System for Sustainable Development (GSSD)**

<http://gssd.mit.edu/>

Ook hier gaat het om metadata die informatiebronnen inzake duurzame ontwikkeling beschrijven. Maar de metadata zijn op een bijzondere manier opgeslagen in thematische deelgebieden die men 'slices' noemt. Men onderscheid veertien deelgebieden:

"The sustainability issue areas are as follows:

- *Energy Use & Resources*
- *Trade & Finance*
- *Industry & Manufacturing*
- *Mobility*
- *Agriculture*
- *Forest & Land use*
- *Water Use & Sources*
- *Conflicts & Wars*
- *Urbanisation*
- *Consumption*
- *Unmet Basic Needs*
- *Population*
- *Migration and Dislocation*
- *Governance & Institutions."*

Alle deelgebieden worden doorkruist door zes 'ringen' die ofwel focussen op probleembeschrijvingen of op oplossingen terzake.

"Problem/Solution or 'Ring' refers to the hierarchies of actions, problems, and solutions across all issue areas:

1. *core: holdings which span the breadth of the sustainability domain*
2. *activities & conditions: specific domain implied by the concept;*
3. *problems types: dysfunctions traced to human activities*
4. *technical & scientific solutions: structured response type to problems;*
5. *economic, regulatory solutions social solutions and related response modes*
6. *coordinated international actions transnational responses and global accords"*

Op het snijpunt van een 'slice' en een 'ring' heeft men een "concept-cell" die op haar beurt onderverdeeld is in concepten (deelaspecten binnen de kruising slice/ring).

Het komt er dus op neer dat men kan opzoeken wie binnen een bepaald deelgebied (slice) bepaalde specifieke problemen behandelt of er technische of andere oplossingen voor aanbiedt.

In de metadata wordt ook nog onderscheid gemaakt tussen verschillende datatypes:

- *agreements*
- *bibliography*
- *case studies*
- *definitions*
- *events*
- *indicators*
- *models*
- *organisations*

Er worden ook gegevens opgeslagen betreffende de geografische locatie waarop de informatie betrekking heeft, enzovoort.

De ambitie van MIT en partners is een kennisbank te ontwikkelen waarmee men informatie betreffende problemen en oplossingen uitwisselt. Het systeem gebruikt slechts een zeer beperkt aantal metadata-elementen. Over het algemeen wordt niet veel meer opgevraagd dan een titel, een abstract (korte beschrijving), de geografische locatie en het webadres, naast enkele administratieve gegevens (naam en e-mail van metadata-leverancier).

Daarnaast is er dus nog de plaatsing van de informatiebronnen in de 'rings' en 'slices'.

Finaal zou men kunnen stellen dat MIT en partners een geavanceerde versie op het Open Directory Project ontwikkelden, omdat men er ook een structuur van folders- en subfolders kan in terugvinden. Maar men moet wel vaststellen dat eenzelfde informatiebron in meerdere folders kan opduiken.

De vraag hoe men nu best een informatie- of kennissysteem in elkaar steekt is al vele jaren onderwerp van onderzoek. Sommigen houden het bij systemen die gebaseerd zijn op het gebruik van relationele databanken, zoals dat ook bij ISDO het geval is. Op zich is dat al een krachtig instrument dat in het verleden in veel domeinen zijn diensten heeft bewezen.

Er zijn andere opties en het meest besproken alternatief dat de laatste jaren aandacht trekt is gebaseerd op de zogenaamde object georiënteerde benadering.

Wie op geavanceerde kennissystemen broedt, kan ook inspiratie zoeken bij wat men in de informatica 'ontologies' noemt. We vermelden dit enkel om aan te geven dat men inzake kennissystemen de lat zeer hoog kan leggen en zowel inzake aanbod en vraag naar data kan anticiperen op alle mogelijke wensen en eisen die kunnen worden verwacht. Daarmee beland men echter in een piste die vele leken kan afschrikken. Het is technisch vrij complexe materie waarbij voer wordt geleverd door verschillende disciplines: informatica, artificial intelligence, kennistechnologie...enzovoort.

Meer hierover:

<http://www.semanticweb.org/knowmarkup.html>

<http://www.cs.utexas.edu/users/mfkb/related.html>

<http://ksl-web.stanford.edu/kst/ontology-sources.html>

<http://ontology.buffalo.edu/>

<http://www.ontology.org/>
<http://members.iinet.net.au/~lonsdale/bm/bm06.htm#ontology>

Als men zich wil wagen aan een experiment met ontology-software kunnen we Protégé aanbevelen (let wel, het duurt minstens een dag vooraleer men de basis onder de knie heeft):

<http://protege.stanford.edu/>

De laatste jaren wordt de ontwikkeling van kennissystemen vanzelfsprekend gekoppeld aan het gebruik van het Internet en het World Wide Web. Hier duikt dan het concept 'semantic web' op, waarbij men effectief de ambitie heeft een alternatief WWW te ontwikkelen waar men wel degelijk vindt wat men zoekt.

Zie o.a:

<http://www.semanticweb.org/>
<http://www.w3.org/2000/01/sw/>

Een artikel in Scientific American beschrijft de ambities en krachtlijnen achter het Semantic Web:
<http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>

Zoals dat met 'ontologies' het geval is vermelden we het 'semantic web' hier slechts als hint voor diegenen die inzake kennismanagement eventueel tot het uiterste willen gaan of althans de ambitie hebben een geavanceerd systeem te ontwikkelen. Dit is evenwel niet de ambitie van dit onderzoek. Wel durven we beweren dat wat we verder aan bod laten komen betreffende metadata en metadata-beheer, kan gezien worden als wegbereiding naar dergelijk meer geavanceerde informatiesystemen. Het is uiterst belangrijk om in te zien dat in de meeste gevallen een kennissysteem de ontwikkeling en het beheer van metadata inhoudt. met ander woorden: metadata zijn heel vaak een cruciaal element.

Wat duurzame ontwikkeling betreft, een domein waar zowel qua inhoud als vorm een zeer grote verscheidenheid aan informatie wordt aangeboden, kan men onzes inziens geen efficiënte en effectieve kennissystemen ontwikkelen zonder metadata.

I.3 Enkele conclusies

Bij de ontwikkeling van kennissystemen heeft men vele opties. Punt is echter dat het een technisch vrij complexe materie betreft die niet in kort bestek kan worden uiteengezet. Maar kan men overwegen om zich vandaag vooral te laten leiden door de ontwikkelingen die internationaal onder de noemer 'semantic web' plaatsvinden. In een algemene krachtlijn wordt met *semantic web* bedoeld: een 'beter' World Wide Web waarin men -met ondersteuning van software - op efficiënte en effectieve manier die informatie kan opsporen die men zoekt.

Men hoeft terzake echter niet te kiezen tussen alles of niets. Met ander woorden de vraag is niet noodzakelijk of we vandaag al dan niet een semantisch web voor duurzame ontwikkeling opstarten. De technische middelen zijn bij vele betrokkenen in dit domein daarvoor nog onvoldoende of zelfs helemaal niet aanwezig, en ook niet snel implementeerbaar.

Wel mag worden vermeld dat de auteurs van dit rapport de ontwikkelingen terzake hebben bekeken en volgende conclusie durven trekken:

Elk project dat bijdraagt tot een beter gebruik van metadata kan op termijn bijdragen tot de ontwikkeling van geavanceerde informatie- of kennissystemen, die de toegankelijkheid en opspoorbaarheid van data kunnen vergroten.

Met andere woorden: men kan voorlopig op de productie van metadata focussen, en dit voorlopig loskoppelen van de vraag hoe deze metadata moeten functioneren in een kennissysteem. Meer nog, men kan er naar streven dat dezelfde metadata als input kunnen dienen voor meerdere informatiesystemen die dan mogelijks hun geëigende diensten aanbieden.

Anderzijds is het wel belangrijk te weten welke de voornaamste rol van die metadata in kennissystemen zal zijn. Dat aspect is in deze inleiding al meerdere malen aangekaart: metadata dienen vooral om het opsporen van informatie te vergemakkelijken. Terzake volgen we de hoofdlijnen

die al in het ISDO-project van de DWTC zijn getrokken (opsporen van informatie vergemakkelijken en communicatie tussen betrokken actoren bevorderen).

Daarnaast kan men ook belang hechten aan het gebruik van metadata om informatie te rangschikken volgens thema's, datatypes en dergelijke. Deze praktijk kan ondersteund worden door gebruik van een thesaurus (gecontroleerde woordenschat). Onzes inziens stelt zich wat dat betreft in het domein duurzame ontwikkeling een lastig probleem. Omdat informatie inzake duurzame ontwikkeling uit praktisch alle disciplines kan komen, zou men een 'woordenschat' moeten hebben waarmee men zowel milieu, economie, sociologie en andere terreinen kan vatten. Zo komt men wellicht tot een woordenschat die zeer omvangrijk wordt. Het praktische nut daarvan kan worden betwijfeld. Wellicht moet als belangrijkste richtlijn bij het ontwikkelen of gebruik van thesauri rekening gehouden worden met de belangrijkste functie van een gecontroleerde woordenschat, vooral het indexeren of catalogeren van informatie in deelthemas, types, doelgroepen en dergelijke. Daarbij zou men zich kunnen laten inspireren door thema-opdelingen die in het domein duurzame ontwikkeling zeer 'populair' zijn, bijvoorbeeld de opdeling 'milieu, economie, sociaal', waarbij milieu dan eventueel onderverdeeld wordt in 'lucht, water, bodem' en/of de milieuthema's of impactcategorieën die gebruikelijk werden in levenscyclus-evaluatie en milieu effect rapportering (verzuring, klimaatsverandering, vermesting, ...).

Het is belangrijk om bij de introductie van een gecontroleerde woordenschat steeds te vermelden dat dit niet de enige woorden zijn waaruit men trefwoorden kan putten, en evenmin de enige woorden zijn die men in een zoekdienst kan opgeven.

Naast een niet-beperkte lijst van trefwoorden, kan men dus ook trefwoorden geven die moeten dienen voor het thematisch ordenen van het informatieaanbod, bijvoorbeeld in een Open Directory.

In dit rapport wordt de ontwikkeling van een gecontroleerde woordenschat geen hoofdbekommernis. Wel wordt er voldoende geanticipeerd op het gebruik daarvan, met ander woorden een opening voorzien die een flexibele toepassing van thesauri toelaat.

II. De beschrijving van Indicatoren voor Duurzame Ontwikkeling

II.1 Multicriteria-evaluatie

Als men zich afvraagt of de wereld, een land, een stad of bedrijf duurzaam is of duurzamer wordt, dan is de eerste vraag vanzelfsprekend: hoe kunnen we de toestand of evolutie meten en evalueren. Over het algemeen zal men de complexe eigenschap 'duurzaamheid' dan onttrafen in een serie deelaspecten, bijvoorbeeld door aan te nemen dat duurzaamheid een gunstige evolutie veronderstelt zowel op economisch, sociaal als ecologisch vlak. Voor het ecologische kan men vervolgens opnieuw een verdere verdeling maken in deelaspecten, bijvoorbeeld door een opdeling te maken in lucht, water en bodem om dan te eindigen bij de selectie van een aantal meetbare parameters die een beeld geven van de (evolutie van) de kwaliteit van die milieucompartimenten. Een dergelijk meetbare parameter is dan bijvoorbeeld de concentratie van zware metalen in het rivierwater.

Het sociale aspect van duurzame ontwikkeling kan eventueel opgedeeld worden in deelaspecten als onderwijs, gezondheid, armoede, ..., waarbij dan opnieuw een aantal meetbare parameters worden geselecteerd die het mogelijk maken om op basis van waarnemingen tot een uitspraak te komen betreffende die deelaspecten. Een meetbare parameter voor gezondheid is bijvoorbeeld de gemiddelde leeftijd waarop de mensen overlijden.

De meetbare parameters noemt men 'indicatoren', gegevens die een indicatie zijn voor iets anders dan datgene wat ze op zich betekenen. Het gemiddelde inkomen van de mensen heeft een betekenis op zich, maar het kan ook gebruikt worden als een indicatie voor het al dan niet voorkomen van armoede. Armoede wordt dus indirect gemeten en het is zeer waarschijnlijk dat het gemiddelde inkomen op zich geen voldoende informatie geeft om een betrouwbare uitspraak te doen over armoede. Het gemiddelde inkomen zegt 'iets', maar niet alles of niet voldoende over armoede. Zo is ook de concentratie van zware metalen slechts een 'indicatie' voor waterkwaliteit, een teken dat in combinatie met andere indicatoren toelaat een redelijk betrouwbare uitspraak te formuleren met betrekking tot de waterkwaliteit.

Drie zaken dienen hier beklemtoond:

1. Een indicator - datgene wat gemeten wordt - staat niet op zich en wordt in verband gebracht met iets anders: de bekommernis die aanleiding gaf tot de meting. Men meet het gemiddelde inkomen omdat men bekommerd is over armoede, zoals een arts een bloeddruk meet omdat hij/zij bekommerd is over de gezondheid van de patiënt.
2. Zelden kan één indicator - één meetbare parameter - voldoende informatie geven om tot een betrouwbare uitspraak te komen betreffende de bekommernis die tot de selectie van de indicator leidde. Men werkt heel vaak met een combinatie van indicatoren om een betrouwbaar eindoordeel te kunnen vormen met betrekking tot een bekommernis. De arts meet niet alleen bloeddruk, maar eventueel ook lichaamstemperatuur.
3. Een bekommernis zoals 'duurzame ontwikkeling' wordt vaak opgedeeld in deelbekommernissen zoals sociale, ecologische en economische ontwikkeling.

Meting en bekommernis

Indicatoren zijn dus meetbare parameters, die geobserveerd worden om uitspraken te doen over een zaak die men belangrijk vindt of een doel dat men nastreeft.

De meetbare parameters zullen we voortaan 'meting' noemen (**measurement**), terwijl het belang of doel voortaan met de term 'bekommernis' (**concern**) wordt gedeut. Het begrip doel kan immers een dubbelzinnige betekenis krijgen, omdat men voor datgene wat men meet ook doelstellingen kan afspreken, m.a.w. men kan stellen dat de meting op een gegeven moment gelijk of groter of kleiner moet zijn dan een bepaalde waarde. **Doel (target)** wordt hier dus gebruikt in de betekenis van 'norm'.

Er is overigens nog een andere reden waarom we de term bekommernis verkiezen boven termen als belang of doel. Datgene dat we onrechtreeks willen meten is niet noodzakelijk een gunstige zaak. Artsen observeren bijvoorbeeld ziektesymptomen en via metingen (bloeddruk en dergelijke) komen ze tot een uitspraak over hun bekommernis, een mogelijke ziekte die men moeilijk een doel kan noemen. De term bekommernis is neutraal en kan dus ingevuld worden met zowel gunstige (belang, doel,...)

als ongunstige (ziekte, symptoom, ...) aspecten. Het spreekt vanzelf dat men vaak een ongunstige bekommernis kan hertalen in een gunstige bekommernis. Als de meting van ziekte(symptoom) een negatief resultaat heeft, dan kan men daaruit wellicht afleiden dat men inzake gezondheid positief scoort. (hoewel: uit het feit dat patiënt X geen griep heeft, volgt niet automatisch dat x gezond is. X kan nog steeds een terminale patiënt zijn).

De definitie van de begrippen bekommernis, meting en doel zoals hierboven beschreven wordt duidelijker met een voorbeeld:

Stel bijvoorbeeld dat we willen bepalen of de welvaart op voldoende peil is:

Bekommernis: welvaart

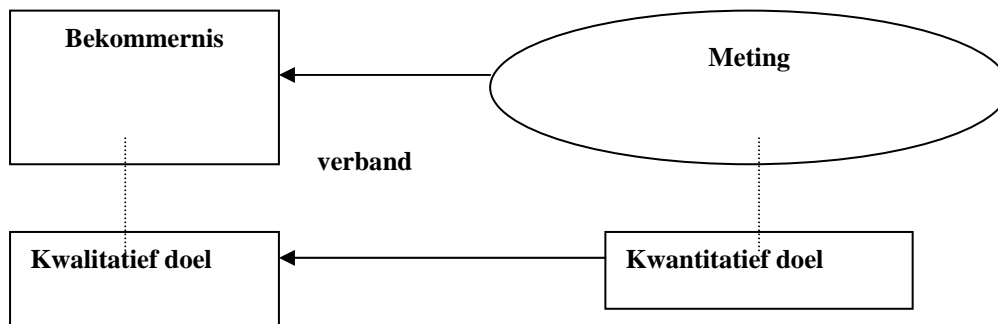
Meting: gemiddeld gezinsinkomen.

Nu kunnen we zowel inzake bekommernis als meting een doel formuleren:

Doel met betrekking tot bekommernis: geen armoede

Doel met betrekking tot meting: gemiddeld gezinsinkomen is groter dan €1000 per maand

Merk op dat de doelstelling met betrekking tot de bekommernis hier in kwalitatieve termen wordt beschreven en dat we die moeten kwantificeren als we deze doelstellingen willen koppelen aan de meting. Overigens is ook het koppelen van metingen aan bekommernissen een kwantificering - operatie.



Wat hoort nu bij de beschrijving van een indicator? Is dat een beschrijving van datgene wat gemeten wordt, of een beschrijving van de meting én de bekommernis en hun onderling verband. Moeten daar ook de kwalitatieve en kwantitatieve doelstellingen bij vermeld worden?

Als men meerdere indicatorprojecten bekijkt dan stelt men vast dat met betrekking tot IDO zelden alleen een beschrijving wordt gegeven van datgene wat men meet. In het algemeen wordt ook de bekommernis besproken en de relatie tussen meting en bekommernis. Men beschrijft dus meting en de interpretatie van die meting in functie van een bekommernis.

Daarbij is het best mogelijk dat de relevante gegevens worden verdeeld over verschillende documenten: een document dat informatie verschaft over de meting en een ander dat de bekommernis beschrijft.

Overigens kunnen ook de doelen (normen) apart behandeld worden. Vaak wordt gesproken over IDO, daar waar men enkel bepaalde bekommernissen meetbaar heeft gemaakt en verder geen concrete doelen vermeld. Men is vaak al tevreden met observatie op zich en een evaluatie van het type: het gaat beter (of het gaat slechter, of er is niets veranderd).

Het duiden van concrete doelen impliceert een 'politieke' verbintenis, is de uitdrukking van een concrete ambitie, waar men achteraf ook op getaxeerd kan worden. Misschien is dat de reden waarom men minder snel geneigd is om concrete normen af te spreken.

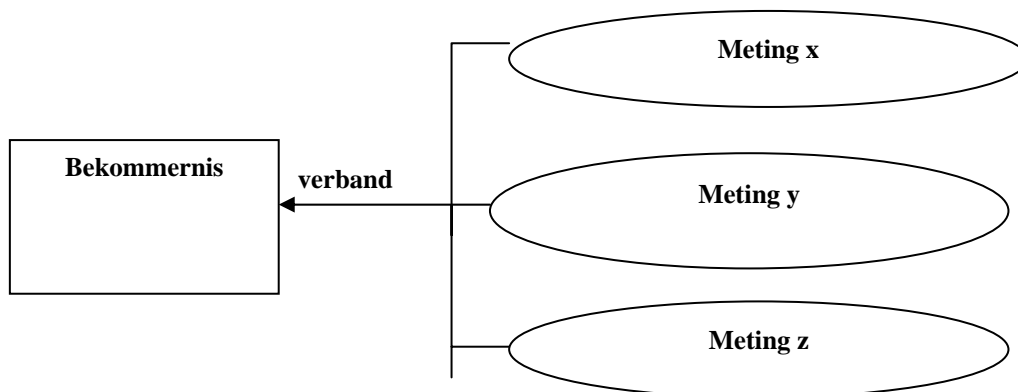
Compleetheid

Nemen we als becommernis welvaart dan kan het 'gemiddeld inkomen' een parameter zijn die iets zegt inzake welvaart, maar velen zullen daarbij benadrukken dat het gemiddeld inkomen onvoldoende informatie verschaft om een uitspraak te doen over de welvaart.

In vakjargon heet het dan dat inkomen niet 'compleet' is, met ander woorden: inkomen dekt onvoldoende de lading welvaart.

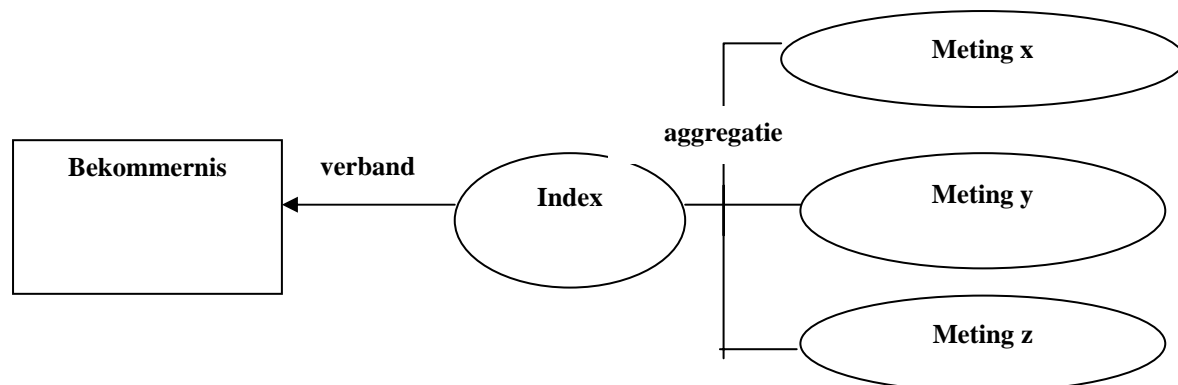
Mogelijks rijst het voorstel om nog andere zaken te meten, zodat er wel een completer (vollediger) evaluatie van de welvaart kan worden gemaakt.

Zo kan men er toe komen om niet alleen inkomen, maar ook de gemiddelde levensverwachting, het gemiddelde aantal jaren dat men school loopt, en andere relevante aspecten van welvaart te beschouwen.



In bovenstaande figuur worden drie metingen (x, y en z) in verband gebracht met één becommernis. Het oordeel met betrekking tot de becommernis zal dan ook berusten op de scores voor deze drie metingen.

In sommige gevallen worden de scores van meerdere metingen geaggregeerd in een zogenaamde 'index' of samengestelde indicator:



Merk op: de aggregatieprocedure die de berekening van één cijfer uit meerdere andere cijfers omvat, vervangt de discussie over het verband met de becommernis niet.

Het feit dat men een index maakt, wijst weliswaar op het feit dat men compleetheid nastreeft, maar daarmee is deze nog niet perfect. Omdat indexen vaak worden genoemd naar de becommernis die men wil observeren, kan dat wel eens tot misverstanden leiden, temeer daar indexen vaak ontwikkeld zijn op basis van praktische overwegingen: men werkt met metingen die beschikbaar zijn, liefst in meerdere landen tegelijk (zodat men kan vergelijken). De Human Development Index (HDI) bijvoorbeeld wordt in brede kringen gewaardeerd, terwijl velen het allicht eens zullen zijn met de bemerking dat de HDI onvolledig is om de sociale ontwikkeling van een land te evalueren.

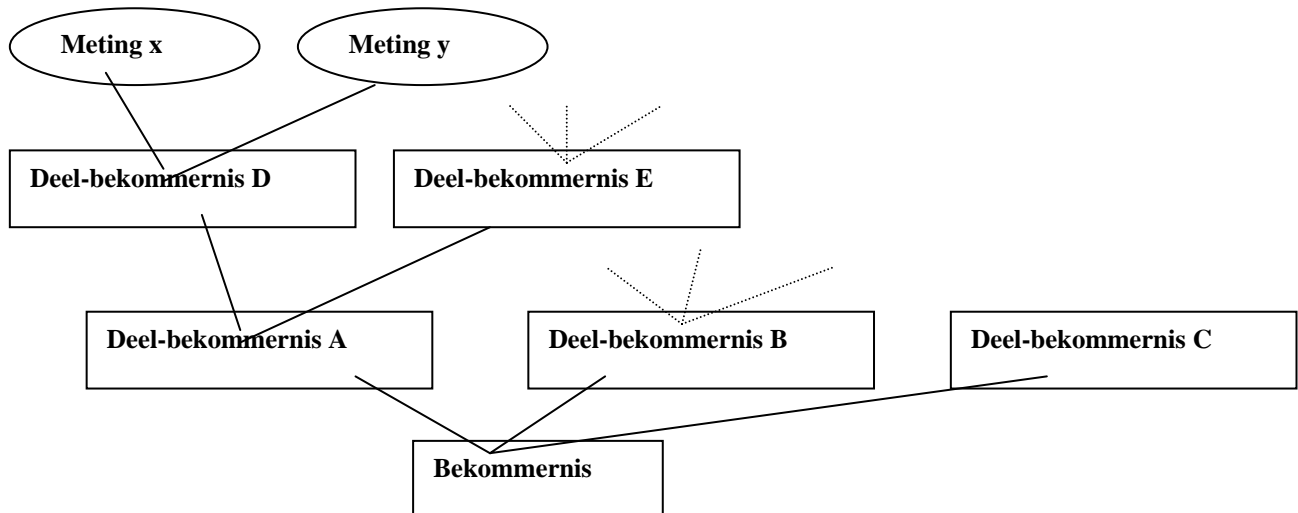
Wanneer nu het verband besproken wordt tussen een bepaalde meting (bijvoorbeeld x) en de bekommernis, dan zal deze bespreking ook de rol van x binnen het trio x, y en z aankaarten. De beschrijving van de ene indicator x is dus slechts volledig als men deze plaatst in de context waarbij x én y én z vermeld worden.

Nu blijkt al dat een beschrijving van een indicator vrij ingewikkeld kan worden als men stelt dat bij deze beschrijving ook het verband met een bekommernis hoort; als men stelt dat de beschrijving zowel die van de meting als de interpretatie van de meting moet omvatten.

Men kan dat vermijden door af te spreken dat met de term indicator bedoeld wordt: 'datgene wat gemeten wordt' en dus ook dat de beschrijving van die indicator enkel de beschrijving van de meting moet inhouden. Uit de wereldwijde praktijk betreffende IDO blijkt dat weinigen deze piste zullen volgen. Een indicator verliest zijn betekenis of functie als men die niet koppelt aan een bekommernis

Boomstructuur

De opsplitsing van de bekommernis 'duurzaamheid' in deelaspecten die vervolgens nogmaals in deelaspecten worden opgedeeld, om eventueel pas in een derde, vierde of vijfde stap tot meetbare parameters te komen, is een proces dat in het algemeen de constructie van een 'boom' oplevert, waar de hoofdbekommernis (duurzaamheid bijvoorbeeld) de stam vormt die via takken en zijtakken (deel-bekommernissen) eindigen bij de bladeren (metingen).



Zo'n boomstructuur is bij IDO praktisch vanzelfsprekend, zo blijkt uit de lezing van talrijke rapporten of websites waar IDO worden gepresenteerd.

Binnen dergelijke structuur kunnen overigens nog andere structuren zitten zoals bijvoorbeeld het 'DPSIR-frame'. Binnen het DPSIR frame wordt voor een bepaalde bekommernis een causaal verband gelegd tussen oorzaken (driving forces, D), gevolgen (pressure P, state S en impact I), en beleidsmatige antwoorden (response R).

Als men de ecologische bekommernis 'klimaatsverandering' bekijkt dan beschouwt men onder de Driving Forces alle activiteiten die bijdragen tot dit probleem (transport, verwarming,..) en men meet dan bijvoorbeeld het aantal kilometers dat met auto's gereden wordt.

Onder Pressure (P, druk) plaatst men gegevens over emissies van broeikasgassen (CO₂ en andere). State (S, toestand) kan gegevens bevatten betreffende de concentratie van broeikasgassen in de atmosfeer, terwijl Impact (I, effect) de aandacht vestigt op gevolgen voor mens en dier. Response (R, antwoord) tenslotte maakt plaats voor indicatoren die iets zeggen over de mate waarin men reageert op het probleem (investeringen in energiebesparende maatregelen bijvoorbeeld).

D, P, S, I en R zijn niet zozeer deel-bekommernissen maar verschillende metingen voor één bepaalde (deel)-bekommernis. Ze vormen samen een analytisch kader waarmee men een bekommernis zo grondig mogelijk en op overzichtelijke wijze behandelt.

Het gebruik van dergelijke schema's laat overigens nog steeds toe alle informatie inzake de IDO te plaatsen binnen een grote boomstructuur, waarbij D, P, D, S en R dan steeds weerkomende takjes zijn waaraan de metingen worden opgehangen.

II.2 Standaardisering van informatie

Eén van de eerste vragen die tijdens dit onderzoeksproject werd gesteld is: Kan men een standaard-document (fiche) ontwikkelen voor de beschrijving van Indicatoren voor Duurzame Ontwikkeling (IDO)? Met ander woorden: kan men een afspraak maken over de gewenste inhoud van een dergelijk document en de structuur waarin deze inhoud moet aangeboden worden?

Het antwoord was vrij snel duidelijk: nee.

Meer genuanceerd is het antwoord: ja, maar het is zeer twijfelachtig of velen deze standaard zullen gebruiken of blijven respecteren. En een standaard die niet gerespecteerd wordt is de facto geen standaard.

De praktijk binnen het eigen huis - het Centrum voor Duurzame Ontwikkeling (CDO) - biedt op zich al voldoende argumenten om deze antwoorden te onderbouwen. De afgelopen jaren heeft het CDO aan meerdere IDO-projecten meegewerkt en ander projecten lopen nog steeds. Nog nooit zijn in twee verschillende projecten dezelfde documentstructuren gebruikt. Zelfs al denk men binnen een bepaald project de perfecte samenstelling gevonden te hebben, bij een ander project wordt die weer gewijzigd.

Daarbij slaan de wijzigingen overigens niet alleen op de manier waarop de beschrijving van een indicator wordt opgedeeld in verschillende 'velden', maar ook over de benoeming van die velden wordt al eens van mening veranderd.

Om dat te demonstreren kunnen we verwijzen naar de 'barometer voor duurzame ontwikkeling' van de stad Gent en de verschillend tussen de editie van 1999 enerzijds en die van 2001 anderzijds.

<http://cdonet.UGent.be/onderzoek/bg/>

en

<http://cdonet.UGent.be/onderzoek/gent2001/>

Een summier overzicht van de respectieve fiches is in onderstaande tabel naast elkaar gezet. Op het gebied van het beheer van alle relevante data is er trouwens veel meer veranderd dan enkel de samenstelling van de technische fiches en de namen van de deelaspecten die in de fiches worden behandeld.

technische fiche in 1999	technische fiche in 2001
omschrijving_naam omschrijving_definitie omschrijving_eenheid omschrijving_betekenis omschrijving_evaluatie	id indicator_naam bronbestand meting_definitie meting_eenheid_kort
data_situatieschets data_gegevens data_databronnen data_gegevensgebrek	primaire_data_naam primaire_data_bron primaire_data_leverancier primaire_data_leverancier_contact primaire_data_kostprijs primaire_data_voorwaarden primaire_data_vorm primaire_data_uitwisselbaar primaire_data_betrouwbaarheid primaire_data_gegevensgebrek primaire_data_actualiteit primaire_data_updating primaire_data_vergelijkbaarheid primaire_data_geldigheid primaire_data_gegevens
datakwaliteit_beschikbaarheid datakwaliteit_verkrijgbaarheid datakwaliteit_betrouwbaarheid datakwaliteit_vergelijkbaarheid datakwaliteit_vergelijkbaarheid datakwaliteit_actualiteit datakwaliteit_kostprijs	omschrijving_berekening berekening uitvoering_berekening
validiteit gewicht	databeschikbaar feitelijke_data voorstelling gegevensgebrek vergelijkbaarheid actualiteit updating toelichting situatieschets
lokale_beleidsdoelstellingen bovenlokale_beleidsdoelstellingen_provinciaal bovenlokale_beleidsdoelstellingen_vlaams bovenlokale_beleidsdoelstellingen_belgisch bovenlokale_beleidsdoelstellingen_eu bovenlokale_beleidsdoelstellingen_internationaal	validiteit administratie_naam administratie_email administratie_contact
beleidsdomeinen participatie ngos verwijzingen	administrative_core_naam administrative_core_email administrative_core_contact administrative_core_datum
verwerkingsfase auteur contactpersonen	

Merk op dat de verschillen hier al groot zijn, ondanks het feit dat het hier tweemaal om dezelfde uitvoerders ging (CDO en stad Gent). Zowel in 1999 als in 2001 dachten betrokkenen een 'goede' fiche te hebben ontwikkeld, maar niemand durft nog te beweren dat die morgen ook 'goed' is.

Andere projecten die in dezelfde periode zijn uitgevoerd of nog lopende zijn leveren andere samenstellingen van fiches op. De klant is koning en hun behoeften of wensen zijn vaak sterk uiteenlopend.

Waar men in het ene project blijkbaar tevreden is met een zeer bondige beschrijving van de IDO, wordt elders een grondige analyse gepresenteerd, inclusief kritische beschouwingen omtrent de

kwaliteit van indicator, eventueel aangevuld met gedetailleerde informatie over de gegevensbronnen, het beheer van de gegevens en van de kwaliteit van dat beheer.

De voorbeelden die in dit hoofdstuk worden toegelicht spreken voor zich en leiden bij de auteurs van dit rapport alvast tot de conclusie dat het wellicht onmogelijk is om op brede schaal een consensus te bereiken omtrent de inhoud van een 'goede' beschrijving van een indicator. Anderzijds kan men uit een reeks voorbeelden wel een lijst opstellen met mogelijke opties, met ander woorden: men zou kunnen aangegeven wat er allemaal kan gezegd worden over een indicator en/of een collectie van indicatoren. Iedereen moet dan maar zelf beslissen welke elementen weerhouden worden of niet. De beslissingen terzake kunnen overigens beïnvloed worden door volgende aspecten:

- Het budget of het beschikbare personeel dat ingezet kan worden bij de ontwikkeling en het beheer van de IDO (evenals de beschikbare tijd). Beperkte middelen kunnen ertoe leiden dat men de inhoud (voorlopig) niet te omslachtig maakt.
- De doelgroepen of het publiek dat men wil aanspreken met de IDO. Men kan zich richten tot een breed publiek, of tot 'ingewijden' - t.t.z. mensen die nauw betrokken zijn bij het beleid of de besluitvorming. De doelgroep kan mee bepalen welke inhoud men relevant vindt of niet. Let wel, het is best mogelijk dat men dezelfde indicator op twee manieren presenteert. Enerzijds met een uitgebreide gedetailleerde beschrijving die alle informatie bevat die relevant is voor diegenen die actief betrokken zijn bij de ontwikkeling, het beheer en gebruik van de indicatoren. Anderzijds een verkorte versie die gericht is op een breder publiek. Het brede publiek kan ook toegang krijgen tot de uitgebreide beschrijving, zodat degenen die toch meer details willen, die gegevens kunnen raadplegen.
- De functie, rol of het gebruik van de IDO. In het eerste hoofdstuk is al aangekaart dat IDO dienen als monitoring- of evaluatie-instrument, als criteria die de besluitvorming kunnen ondersteunen en/of als instrument waarmee men het vage concept 'duurzame ontwikkeling' een meer eenduidige concrete invulling geeft (operationaliseren van het concept duurzame ontwikkeling). Bij dat laatste worden de IDO dus resultaat van visievorming. Over het algemeen overlappen deze functies - evaluatie, besluitvorming en visievorming - elkaar. Meer nog: ze zouden elkaar moeten overlappen. Maar dat neemt niet weg dat men in eerste instantie de IDO kan gebruiken voor ex-post evaluaties van het gevoerde beleid en daarom in de beschrijving van de indicator een dergelijke evaluatie van de (laatste) waarnemingen opneemt. Maar men kan ook deze evaluatie uitbreiden met beslissingen of opties voor het beleid dat verder kan of moet gevoerd worden.
- De context waarin de IDO worden gepresenteerd. Doelgroep en functie bepalen mee de context, maar in dit punt willen we de aandacht toespitsen op het feit dat de indicatoren niet noodzakelijk alleen in een strikt beleidsmatige context kunnen worden gebruikt, maar evenzeer een educatieve context of als 'propaganda'-middel dat bewustwording en gedragsverandering moet aanzwengelen. Ook dit kan er toe leiden dat men inhoudelijk andere klemtonen legt.
- De structurele samenhang met andere informatie. Informatie die relevant is met betrekking tot een bepaalde indicator, hoeft niet noodzakelijk opgenomen te worden in het document dat die indicator beschrijft. Stel bijvoorbeeld dat men in een stad de sociale aspecten van duurzame ontwikkeling evalueert op basis van vier indicatoren. Men kan nu in de beschrijving van elke indicator een oordeel vellen over de betekenis of aard van deze indicator in samenhang met de drie andere. Maar dat heeft dan tot gevolg dat men vier keer dezelfde informatie moet herhalen. De samenhang en/of complementariteit van de vier indicatoren kan echter in een apart document worden beschreven waarbij dan vier 'links' worden gelegd naar de vier documenten waar de meer specifieke informatie van de indicatoren wordt gegeven. In dit geval zit informatie die relevant is met betrekking tot één indicator verspreid over meerdere documenten (minstens twee in dit voorbeeld). Het moet opgemerkt dat er heel veel projecten zijn waar dit fenomeen zich voordoet, met ander woorden: waarbij men het gehele 'rapport' moet doornemen om alle relevante informatie betreffende een specifieke indicator gezien te hebben. Dit kan er soms toe leiden dat de documenten die één bepaalde IDO beschrijven zeer beknopt worden, omdat er al op een andere plaats is toegelicht waarom een bepaalde indicator werd weerhouden en hoe deze geïnterpreteerd moet worden in samenhang met andere indicatoren.

Nog enkele voorbeelden van indicatorbeschrijvingen, die vermelde vaststellingen bevestigen:

Een eerste voorbeeld is een fiche van UN-CSD (Commision on Sustainable Development of the United Nations):

<p>1. INDICATOR</p> <p>(a) Name: Unemployment Rate. (b) Brief Definition: Unemployment rate is the ratio of unemployed people to the labour force. (c) Unit of Measurement: %. (d) Placement in the CSD Indicator Set: Social/Equity/Poverty.</p>
<p>2. POLICY RELEVANCE</p> <p>(a) Purpose: The unemployment rate is a measure of the unutilized labour supply of a country. If employment is viewed as the desired portion of the economically active population (labour force), unemployment can be seen as, for the most part, the undesirable portion (although some short-term unemployment can be both desirable and necessary). Unemployment rates by specific groups- such as by age, sex, occupation or industry- are also useful statistics in identifying groups of workers and sectors most vulnerable to joblessness.</p> <p>(b) Relevance to Sustainable/Unsustainable Development (theme/sub-theme): Unemployment is useful and relevant to measuring sustainable development, especially if uniformly measured over time, and considered with other socioeconomic indicators. It is one of the main reasons for poverty in rich and medium income countries and among persons with high education in low income countries (no work, no income but compensation from insurance schemes or other welfare state systems whenever they exist). It should be noted, however, that it is common to find people working full-time but remaining poor due to the particular social conditions and type of industrial relations prevalent in their country, industry, or occupation.</p> <p>(c) International Conventions and Agreements: The measurements of unemployment and the labour force are defined in the International Labour Office (ILO): Resolution concerning statistics of the economically active population, employment, unemployment and underemployment, 13th International Conference of Labour Statisticians, Geneva, 1982.</p> <p>(d) International Targets/Recommended Standards: There are no international targets regarding the rate of unemployment.</p> <p>(e) Linkages to Other Indicators: The indicator is one among many that measure utilization or underutilization of the labour market. Other measurements focus on parts of the unemployment experience: youth unemployment, long-term unemployment, unemployment by educational attainment, time-related underemployment and the inactivity rate.</p>
<p>3. METHODOLOGICAL DESCRIPTION</p> <p>(a) Underlying Definitions and Concepts: The definitions for labour force, employed population, and unemployed population are well established by international agreements (see section 6 below).</p> <p>i) Labour Force: The current economically active population or labour force has two components: the employed and the unemployed population. The international standard definition of labour force established by the Thirteen International Conference of Labour Statisticians (International Labour Office (ILO), 1982) is based on the following elements:</p> <ul style="list-style-type: none"> - <i>The survey population:</i> All usual residents (de jure population) or all persons present in the country at the time of the survey (de facto population). Some particular groups, such as the armed forces or other populations living in institutions, nomadic people, etc., may be excluded. - <i>An age limit:</i> In countries where compulsory schooling and legislation on the minimum age for admission to employment have broad coverage and are widely respected, the age specified in these regulations may be used as a basis for determining an appropriate minimum age limit for measuring the economically active population. In other countries, the minimum age limit should be determined empirically on the basis of <ul style="list-style-type: none"> (i) the extent and intensity of participation in economic activities by young people, and (ii) the feasibility and cost of measuring such participation with acceptable accuracy. Some countries also determine a maximum age for inclusion in the labour force. - <i>The involvement in economic activities during the survey reference period:</i> The concept of economic activity adopted by the Thirteenth International Conference of Labour Statisticians

(1982) is defined in terms of production of goods and services as set forth by the United Nations System of National Accounts, (revised in 1993).

- *A short reference period:* For example, one week or a day.

ii) *Employed population:* According to the 1982 international definition of employment (ILO, 1983) the employed comprise all persons above the age specified for measuring the labour force, who were in the following categories:

- *Paid employment:* (i) at work: persons who, during the reference period, performed some work (at least one hour) for wage or salary, in cash or in kind; (ii) with a job but not at work: persons who, having already worked in their present job, were temporarily not at work during the reference period but had a formal attachment to their job;
- *Self-employment:* (i) at work: persons who, during the reference period, performed some work (at least one hour) for profit or family gain, in cash or in kind; (ii) with an enterprise but not at work: persons with an enterprise, which may be a business enterprise, a farm or a service undertaking, who were temporarily not at work during the reference period for some specific reason.

iii) *Unemployed population:* According to the 1982 international definition of employment (ILO, 1983) the unemployed comprise all persons above the age specified for measuring the labour force, who during the survey reference period were at the same time: (i) not in paid employment or self-employment, not even for an hour; (ii) available for work; and (iii) seeking work.

(b) **Measurement Methods:** Household or labour force surveys are generally the most comprehensive and comparable sources for unemployment statistics. Other sources include population censuses, "employment office records" and "official estimates". Data based on registration at employment offices tend to understate unemployment, in comparison with household or labour force surveys, because not all persons who are looking for work will register on account of eligibility requirements which may exclude those who have never worked or have not worked in a recent period. (In some countries, registration data can overstate unemployment, largely because of double-counting and failure to track persons registering, not all of whom, in any case, may be job-seekers). Official estimates are often based on a combination of sources.

Population censuses generally do not probe very deeply into labour force status, resulting in magnitudes of unemployment that differ substantially (either higher or lower) from those obtained from household surveys where more questions are asked.

(c) **Limitations of the Indicator:** As important as the unemployment rate is, it should not be interpreted as a measure of economic hardship. Doing so can produce some unfortunate results, giving unemployment a greater degree of significance than it deserves. The unemployment rate, if based on the internationally recommended standards, simply tells us the proportion of the labour force that does not have a job but is available and actively looking for work. It says nothing about the economic resources of the unemployed worker or the worker's family. The scope of unemployment should therefore be limited to its use as a measurement of the utilization of labour, and should not be extended to other spheres of the economy of a country. Broader measures, including income-related indicators, are needed to evaluate economic hardship.

Paradoxically, low unemployment rates may well disguise substantial poverty in a country, whereas high unemployment rates can occur in countries with significant economic development and low incidences of poverty. In countries without a safety net of unemployment insurance and welfare benefits, many individuals simply cannot afford to be unemployed. Instead they eke out a living in the informal sector. In countries with well-developed social protection schemes, workers can better afford to take the time to find desirable jobs.

(d) **Status of the Methodology:** Well developed and employed although discrepancies do occur.

In an effort to resolve the international comparability issue for its member-countries and building on work carried out by the United States Bureau of Labour Statistics in the 1960's, the Organization for Economic Co-operation and Development (OECD) initiated research on and has published "standardized unemployment rates" adjusted to the International Labour Office (ILO) concepts. The ILO extended the process even further, beginning in 1990. The ILO-comparable unemployment rates show historical data for 25 of the ILO member States based on the ILO-comparable series (produced in ILO: 1999 Key Indicators of the Labour Market, Geneva, 1999). This table represents unemployment rates from national labour force surveys that have been reconciled with and adjusted to

make the estimates conceptually consistent, with the strictest application of the ILO statistical standards. This implies that participating countries must be able to provide detailed information on the composite elements of their labour forces.

At the same time, the unemployment rates obtained are in conformity with the OECD's programme of standardized rates, which itself is based on the ILO standards. This avoids a proliferation of "international" estimates, which might not be the same. Further, all the data are expressed in terms of annual averages (or a period that is currently found to be the most representative over the year), thereby avoiding the variances that would occur if different reference periods were observed. These estimates, based on official national data, should provide the best basis currently available for making reasonable international comparisons and assumptions, although they may still contain very minor discrepancies.

(e) **Alternative Definitions/Indicators:** Underemployment rate; discouraged workers rate.

4. ASSESSMENT OF DATA

(a) **Data Needed to Compile the Indicator:** Labour force (total number of persons) and total number of unemployed persons, derived from the same survey.

(b) **National and International Data Availability and Sources:** Unemployment rate data are available for a total of 113 countries, all of which are broken down by gender, in the 1999 KILM with the majority of data resulting from household or labour force surveys with the remainder from employment office records, official estimates or population censuses.

(c) **Data References:** The data repositories used are International Labour Office (ILO) Yearbook of Labour Statistics, OECD Labour Force Statistics, and ILO Digest of Caribbean Labour. For seven countries, data were taken from national sources.

5. AGENCIES INVOLVED WITH THE DEVELOPMENT OF THE INDICATOR

(a) **Lead Agency:** The lead agency is the International Labour Office (ILO) of the United Nations, located in Geneva. The contact point is the Focal Point for Environment and Sustainable Development, ILO; fax no. (41-22) 798 8685.

(b) **Other Contributing Organizations:** None.

6. REFERENCES

(a) **Readings:**

Yearbook of Labour Statistics, ILO, Geneva;

Bulletin of Labour Statistics (quarterly) and its Supplement (January/February, April/May, July/August and October/November), ILO, Geneva;

Statistical yearbooks and other publications issued by the national statistical offices.

Surveys of Economically Active Population, Employment, Unemployment and Underemployment -An ILO Manual on Concepts and Methods, ILO, Geneva, 1992.

Sources and Methods: Labour Statistics, Volumes 3 and 5, ILO, Geneva, 1991 and 1990, currently updated.

System of National Accounts 1993, Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank, Brussels/Luxembourg, New York, Paris, Washington, D.C., 1993;

Current international recommendations on labour statistics, ILO, Geneva, 1988. See particularly the *Resolution Concerning Statistics of the Economically Active Population, Employment, Unemployment and Underemployment*, adopted by the Thirteenth International Conference of Labour Statisticians (October 1982).

(b) **Internet sites:**

For 1999 Key Indicators of the Labour Market, Geneva, 1999:
<http://www.ilo.org/public/english/employment/strat/polemp/kilm/>

For the text of the resolution concerning statistics of the economically active population, employment, unemployment and underemployment see:
<http://www.ilo.org/public/english/bureau/stat/res/ecacpop.htm>

For the ILO database on labour statistics, see <http://laborsta.ilo.org>

De algemene structuur die de CSD gebruikt en de inhoud is niet uitzonderlijk. Even belangrijk is echter de vaststelling dat er betreffende IDO ook fiches bestaan die sterk afwijken van deze van de deze van de CSD, in eerste instantie die van het CDO die we al eerder vermelden, maar ook:

De fiches uit Calgary in Canada, waarvan hieronder een voorbeeld:

<http://www.telusplanet.net/public/sustcalg/sooc2001.htm>

Crime Rate & Rate of Victimization

The facts: The estimated person crime rate per 100,000 people for 1999 in Calgary was 1,015. The estimated property crime rate was 6,724.

Definition: These statistics come from the Calgary Police Service Annual Statistical Report for 1995-99. Person crime includes attempted and committed homicide; street, financial, and commercial robbery; sex-related offences; and assault. Property crime includes fraud, break and enter, and theft and vehicle theft.

Trend: Since 1995 the rate of property crime has decreased consistently and is 9.3% lower in 1999 than the rate of 7,414 registered in 1995. The rate of person crime has been less consistent but has shown a 13% increase since 1995 when the rate was 899.

Incidence of domestic violence has been notoriously underreported, and in fact rigorous reporting only began in 1997. In 1998, the first full year of reporting, 4,700 offences were reported, rising to 4,900 incidents in 1999. Children were present in 49% of those incidents. A 1999 Angus Reid survey reported domestic violence as the top social concern of Calgarians.

Youth crime peaked in 1991 and has declined steadily since then, with 1999 having the lowest level in the past 10 years. The percentage of youth crime that is violent crime is increasing, but this may be due to a change in reporting procedures and lower tolerance levels. Since the tracking of school crime began in 1996, incidence has been decreasing.

Despite significant drops in national crime rates in the past 10 years, Statistics Canada points out that crime rates are about 10 to 13% higher than 20 years ago.

Importance: A community in which citizens do not feel safe, or are not in fact safe, is not sustainable – quality of life will deteriorate. Crime directly decreases the quality of life of victims through financial loss, physical injury, emotional trauma, and alienation. The repercussions of a crime spread beyond the immediate victim. Parents, children, friends, coworkers, witnesses, and the community also suffer after a crime has occurred. The fear of crime can lead people to secure themselves behind locked doors and resist stepping out into the community to build ties and contribute.

Linkages: Crime costs millions of dollars annually and therefore affects the economic development of a community. In areas hard hit by crime, housing prices drop and people who can afford it move to other neighbourhoods. Businesses bypass high crime areas.

Common community and individual measures to prevent crime aim to lessen the opportunities for crime to occur. These measures include organizing Neighbourhood Watch and Block Parent programs and modifying urban design (e.g., improving street lighting or encouraging increased pedestrian activity).

While these measures can be effective, they do not address the long-term root causes of crime. Building healthy, caring communities is one of the best ways to prevent crime. This approach is commonly referred to as “crime prevention through social development.”

Key factors in this approach include the provision of employment and educational opportunities, access to services, adequate housing, and accessible play and recreational facilities.

At the root of adult crime is exposure to violence in the home. According to The Canadian Fact Book on Poverty, a child's risk of poor outcomes dramatically diminishes as family income reaches \$30,000, and it diminishes further as incomes reach \$40,000.

Individual and Collective Actions:

- Support agencies and programs that help develop local crime prevention through social development strategies, such as the Action Committee Against Violence.

- Get involved in constructive activities for youth.

Terwijl men bij CSD veel plaats maakt voor 'methodologie' (kwaliteit van de observaties), is dat in Calgary minder het geval. Misschien heeft men zich daar ook wel om bekommerd, maar blijkbaar vond men het niet noodzakelijk om daar breedvoerig over uit te wijden in de steekkaart. We vermoeden dat veel verschillen kunnen verklaard worden in de verschillende doelgroepen die CSD enerzijds en Calgary anderzijds voor ogen hebben. De CSD richt zich in feite tot nationale administraties die aangespoord worden om in hun land een collectie van IDO op te stellen, geïnspireerd door de collectie van de CSD. De steekkaarten van CSD zijn dan ook opgesteld als handleidingen voor de potentiële gebruiker van deze indicator en vandaar allicht dat men alle relevante technische aspecten, met de nodige referenties, aankaart.

Calgary echter richt zich tot het brede lokale publiek en maakt duidelijke verbanden met het gevoerde of te voeren beleid. De vraag of de ene steekkaart nu beter is dan de anderen is in feite een vergelijking maken tussen appels en peren, omdat de twee steekkaarten verschillende functies hebben en in verschillende contexten staan.

Om de 'boomstructuur' van IDO-projecten te illustreren bekijken we een project van de Britse regering. Die situeert haar IDO-project als volgt:

"In December 1999 the UK Government published in 'Quality of life counts' a core set of about 150 indicators of sustainable development which will be used in future to monitor national progress. These indicators underpin the Strategy 'A better quality of life: a strategy for sustainable development in the UK' (May 1999) ...

The strategy has four main aims:

- *social progress which recognises the needs of everyone*
- *effective protection of the environment*
- *prudent use of natural resources*
- *maintenance of high and stable levels of economic growth and employment.*

The indicators give a broad overview of whether we are achieving a 'better quality of life for everyone, now and for future generations to come'. They cover the three pillars of sustainable development, namely social progress, economic growth and environmental protection, including people's everyday concerns - like health, jobs, crime, air quality, traffic, housing, educational achievement, wildlife and economic prosperity."

We belichten hier de zogenaamde Headline Indicators of Sustainable Development for the UK

<http://www.sustainable-development.gov.uk/indicators/headline/index.htm>

Voor een bredere context, zie:

http://www.sustainable-development.gov.uk/uk_strategy/factsheets/monitor/index.htm

Bij de Britten vinden we opnieuw een hiërarchische structuur die vertrekt bij de drie grote thema's economie-sociaal-milieu:

Framework and objectives

The headline indicators in the UK sustainable development strategy		
Themes, issues and objectives (Strategy reference)	Ref no	Headline Indicators
Maintaining high and stable levels of economic growth and employment		
<ul style="list-style-type: none"> • our economy must continue to grow (3.12) 	H1	GDP and GDP per head

<ul style="list-style-type: none"> investment (in modern plant and machinery as well as research and development) is vital to our future prosperity (3.14) 	H2	Total and social investment as a percentage of GDP
<ul style="list-style-type: none"> maintain high and stable levels of employment so everyone can share greater job opportunities (Box after 1.8) 	H3	Percentage of people of working age who are in work
Social progress which recognises the needs of everyone		
<ul style="list-style-type: none"> tackling poverty and social exclusion (7.17) 	H4	Indicators of success in tackling poverty and social exclusion (children in low income households, adults without qualifications and in workless households, elderly in fuel poverty)
<ul style="list-style-type: none"> equip people with the skills to fulfil their potential (3.16) 	H5	Qualifications at age 19
<ul style="list-style-type: none"> improve health of the population overall (3.17) 	H6	Expected years of healthy life
<ul style="list-style-type: none"> reduce the proportion of unfit (housing) stock (3.18) 	H7	Non-decent housing
<ul style="list-style-type: none"> reduce both crime and peoples fear of crime (3.19) 	H8	Level of crime
Effective protection of the environment		
<ul style="list-style-type: none"> continue to reduce our emissions of greenhouse gases now, and plan for greater reductions in longer term (3.20) 	H9	Emissions of greenhouse gases
<ul style="list-style-type: none"> reduce air pollution and ensure air quality continues to improve through the longer term (3.21) 	H10	Days when air pollution is moderate or higher
<ul style="list-style-type: none"> improve choice in transport; improve access to education, jobs leisure and services; and reduce the need to travel (7.29) 	H11	Road traffic
<ul style="list-style-type: none"> improving river quality (8.30) 	H12	Rivers of good or fair quality
<ul style="list-style-type: none"> reverse the long-term decline in populations of farmland and woodland birds (3.24) 	H13	Populations of wild birds
<ul style="list-style-type: none"> re-using previously developed land, in order to protect the countryside and encourage urban regeneration (3.25) 	H14	New homes built on previously developed land

Prudent use of natural resources		
<ul style="list-style-type: none"> • move away from disposal of waste towards waste minimisation, reuse, recycling and recovery (Box after 6.11) 	H15	Waste arisings and management

Op het World Wide Web kunnen nog vele andere voorbeelden geraadpleegd worden. Men kan zich daarbij zijn weg zoeken via de databases met indicatoren en indicatorprojecten die voordien in hoofdstuk één werden vermeld.

Men zou kunnen proberen een fiche of steekkaart te ontwerpen die zodanig is samengesteld dat alle data uit de concrete voorbeelden kunnen herschikt worden in het nieuwe ontwerp. Er is met andere woorden een standaard steekkaart die door iedereen gebruikt kan worden. Naarmate er steeds meer voorbeelden worden bekeken, stijgt het vermoeden dat men hier het onmogelijke probeert te realiseren.

Het voornaamste probleem is dan niet zozeer dat men in het ene voorbeeld een dataveld of data-element vindt dat het andere niet heeft, maar wel dat de ene bepaalde gegevens verdeeld over een aantal deelvelden – m.a.w. een bepaald aspect of kenmerk van een indicator wordt opgesplitst in deelaspecten-, die bij een andere worden samengevoegd in één veld.

Bijvoorbeeld:

In het ene project bevat de fiche verschillende velden waarin men informatie verschaft over de kwaliteit van de indicator, waarbij kwaliteit is opgesplitst in verschillende deelaspecten, zoals:

- begrijpbaarheid
- beschikbaarheid en toegankelijkheid van gegevens
- methodologische standvastigheid
- kostprijs
- gevoeligheid

In het ander project is echter slechts een veld voorzien:

- kwaliteit

Nu is het best mogelijk dat in het tweede geval eveneens informatie wordt gegeven over begrijpbaarheid, beschikbaarheid van gegevens, kostprijs, enzovoort. Maar men heeft blijkbaar beslist om de gegevens niet op te splitsen.

Indien men nu een 'standaard'-structuur wil ontwikkelen waarin beide partijen hun data moeten plaatsen, dan kan men moeilijk beide tegelijk een bevredigend aanbod doen. Ofwel kiest men voor de gedetailleerde aanpak zoals in voorbeeld één, maar dan zou de tweede partij de kwaliteitsgegevens waarover ze beschikken moeten opdelen over de verschillende velden.

Ofwel beslist men in de standaard ook slechts één veld op te nemen - 'kwaliteit' - waarbij de eerste partij dan al de data die met kwaliteit te maken hebben hierin moeten onderbrengen. In het laatste geval kan men overigens nog steeds de specifieke termen 'begrijpbaarheid', 'beschikbaarheid', e.d. handhaven. Met ander woorden, binnen het veld 'kwaliteit' kan een lijst van deelaspecten opgesomd worden.

Het probleem dat zich nu stelt is het volgende:

- Ofwel streeft men naar een standaard die flexibel is, met andere woorden een relatief grote vrijheid geeft aan de gebruikers van de standaard. Dan zal men eerder kiezen voor de optie met één veld – kwaliteit - waarbij de eindgebruiker moet beslissen hoe dit veld wordt ingevuld, al dan niet met een lijst van deelaspecten.

- Ofwel laat men zich leiden door de vakliteratuur over indicatoren waarbij vaak aangestuurd wordt op 'good practice' , wat ondermeer kan betekenen dat men specifieke of expliciete aandacht wenst voor de verschillende kwaliteitsaspecten.

Er is echter nog een tweede knelpunt. Wat verstaat men onder 'kwaliteit van de indicator'. Kan er geen onderscheid gemaakt worden tussen de kwaliteit van de observaties enerzijds en de kwaliteit van de indicator.

Wat betreft de observaties kunnen dan volgende deelspacten worden bekeken:

- beschikbaarheid en toegankelijkheid van data
- regelmaat van de observaties
- methodologische standvastigheid
- kostprijs

Wat betreft de kwaliteit van de indicator, focust men echter op het verband tussen datgene wat gemeten wordt en datgene wat via de meting geëvalueerd wordt (de becommernis). Stel dat men bijvoorbeeld een indicator voor 'gezondheid' maakt en daarbij overweegt om de 'gemiddelde levensverwachting' te meten, dan kan men onder de noemer 'kwaliteit' een antwoord geven op de vraag in hoeverre 'levensverwachting' een geschikte en toereikende parameter is om een oordeel te vellen over gezondheid. Sommigen zullen dan stellen dat voor dat laatste andere termen kan gebruiken, zoals 'validiteit' of 'evaluatie' in plaats van kwaliteit. Anderen zullen echter opmerken dat het begrip 'evaluatie' moet gereserveerd worden voor een eindbeschouwing die men koppelt aan de observaties.

De beschrijving van de kwaliteit (validiteit) van de indicator kan ook in deelaspecten worden opgesplitst: bijvoorbeeld in de deelaspecten compleetheid, complementariteit en alternatief. Onder 'compleetheid' wordt dan de vraag gesteld in welke mate bijvoorbeeld de 'levensverwachting' een voldoende of toereikend beeld (indicatie) geeft van de gezondheid. Onder 'complement' worden eventueel (verwijzingen naar) andere indicatoren vermeld die de eventuele onvolledigheid kunnen goedmaken, terwijl men in het veld 'alternatief' andere opties vermeld.

Daarmee komt dan een tweede probleem naar voor: begripsverwarring. Indien men verschillende IDO-projecten bekijkt dan kan men vaststellen dat niet overal dezelfde invulling wordt gegeven aan termen als 'kwaliteit' of 'evaluatie'. Lokaal kunnen die wel eenduidig gedefinieerd zijn, maar de afspraken die binnen de ene groep werden gemaakt, gelden daarom niet binnen een andere groep.

Binnen het Centrum voor Duurzame Ontwikkeling is geprobeerd om tot een akkoord te komen met betrekking tot de structuur van een fiche voor de beschrijving van een indicator en de namen (en hun betekenis) van de velden of elementen in zo'n fiche.

Zelfs binnen een beperkte groep is het al moeilijk om daarover een consensus te bereiken die iedereen honderd procent bijtreedt, of waarvan iedereen zegt: 'hieraan zal ik mij steeds houden'. Finaal blijkt dat men toch nog de vrijheid wil hebben om - afhankelijk van de context - wijzigingen aan te brengen, zowel in de structuur als in de namen waarmee men de velden of data-elementen benoemt. Het belangrijkste argument is dan dat men in elk geval ruimte wil laten voor inspraak van de 'klant', de stad of regio die beroep doet op het CDO om IDO te ontwikkelen. Uiteindelijk is het de eindgebruiker die zich goed moet voelen met de structuur en de aangewende terminologie.

Eén en ander leidt - althans binnen het CDO - tot de conclusie dat de ambitie om tot een standaard te komen die nationaal of internationaal aanvaard wordt, wellicht zeer hoog gegrepen is.

Op basis van bovenstaande beschouwingen heeft het CDO al meerdere structuurontwerpen gemaakt voor de beschrijving van indicatoren.

Het CDO was en is betrokken bij meerdere IDO-projecten, voornamelijk op stedelijk niveau.

Tot vandaag hebben al deze projecten hun eigen documentstructuur opgeleverd. Soms zijn de beschrijvingen van de indicatoren relatief beknopt, soms zijn ze vrij uitgebreid.

Vast staat dat de structuur voortdurend wijzigt, deels omdat de 'klant' wisselende eisen stelt, maar ook omdat men steeds lessen trekt uit de vorige projecten. Naarmate de jaren vorderen worden de opties steeds uitgebreider.

Een van de recente ontwerpen laat toe dat bepaalde velden (elementen) of series van velden meerdere keren voorkomen in het document. In onderstaande lijst van data-elementen wordt aangegeven of een bepaald veld nul keer (optioneel veld), één keer of eventueel meer dan één keer kan gebruikt worden. Onderstaande lijst is enkel bedoeld om te illustreren hoe lang het lijstje van optionele data-elementen kan worden als men er maar lang genoeg over nadenkt.

Een korte toelichting bij de terminologie: elke indicator wordt beschreven als een verband tussen iets dat gemeten wordt (meting) en iets dat men middels die meting wil evalueren (bekommernis). Een dergelijk bekommernis kan bijvoorbeeld 'gezondheid' zijn, terwijl een meting dan bijvoorbeeld 'levensverwachting' is.

Teken	Betekenis
?	Element komt nul of één keer voor
+	Minstens één, of meerdere keren
*	Nul, één of meerdere

- indicator_naam +
- compilatie?
 - compilatie_grootheid
 - compilatie_definitie
 - compilatie_eenheid_kort
 - compilatie_eenheid_voluit
 - compilatie_commentaar*
 - div*
- meting+
 - meting_grootheid +
 - meting_definitie +
 - meting_eenheid_kort +
 - meting_eenheid_voluit +
 - conversie +
 - meting_commentaar *
 - div *
- bekommernis+
 - naam_bekommernis
 - korte_beschrijving
 - lange_beschrijving
 - classificatie
 - belanghebbende*
 - beleidsdomein*
 - bekommernis_commentaar*
 - div *
- observatie
 - bron +
 - data leverancier
 - methodologie
 - data_beschikbaarheid
 - methode_standvastigheid
 - feitelijke_data
 - toelichting
 - maatstafvergelijking
 - bron_commentaar*
 - div *
- validiteit
 - compleetheid +
 - complementair*
 - alternatief*
 - gevoeligheid
 - begrijpbaarheid

- validiteit_commentaar*
- div *
- doelstellingen_lijst
 - doel +
 - geografische_schaal
 - tijds_horizon
 - kwalitatief_doel
 - kwantitatief_doel
 - formele_afspraak*
 - doel_commentaar*
 - div*
 - doelstellingen_commentaar*
 - div*
- participatie
 - actie +
 - actie_type
 - actie_beschrijving
 - actie_commentaar*
 - actor +
 - actor_naam
 - actor_type
 - bijdrage
 - actor_commentaar*
 - div *
 - div*
- div *
- referentie_lijst
 - referentie +
 - webadres
 - auteur
 - titel
 - referentie_bron
 - plaats
 - uitgever
 - datum
- document beheer
 - administrative_core +
 - name
 - email
 - contact
 - action
 - date
 - comment
 - vordering
 - version
 - daterange
 - status
 - comment

Merk op dat het ontwerp voorziet in de beschrijving van een 'samengestelde' indicator (index), met andere woorden een indicator die gebaseerd is op meerdere metingen die dan door een berekening (compilatie) één afgeleide grootheid opleveren.

Een dergelijke compilatie kan een eenvoudige verhouding zijn, bijvoorbeeld 'elektriciteitsgebruik per persoon', waarbij dan twee metingen worden verricht: het aantal personen én het totale

elektriciteitsverbruik. De compilatie kan echter ook ingewikkelder zijn, zoals dat bijvoorbeeld het geval is met de Human Development Index.

Regelmatig komt een zogenaamd 'div'-element voor, een element dat een opening creëert voor bijkomende informatie (diverse).

Hoewel het Centrum voor Duurzame Ontwikkeling probeert om tot een standaard document-structuur voor de beschrijving van indicatoren te komen, lukt het nauwelijks om het gebruik daarvan in de praktijk te handhaven.

Aangezien de 'klanten' (de organisaties waarvoor het CDO de ontwikkeling van IDO begeleid) ook hun zeg hebben in de data die worden geproduceerd en de manier waarop deze worden beheerd, wordt aanvaard dat de zogenaamde 'standaard' veeleer een basisontwerp is waarvan een concrete toepassing kan worden afgeleid. Het is dus veeleer een bron van inspiratie en geen 'regel'.

Bij wijze van voorbeeld kunnen we volgend project vermelden: 'Op weg naar duurzaamheidsindicatoren voor het kustgebied'¹. Hoewel bij dat project op termijn de ontwikkeling van vrij gedetailleerde steekkaarten was voorzien, is daar in de eerste verkennende fase een kleine steekkaart gebruikt met slechts zes data-elementen of velden.

Naam	Naam van de indicator
Definitie	Bondige en eenduidige omschrijving
Meeteenheid	Maat, hoeveelheid of grootheid
Betekenis	Relatie tussen indicator en 'duurzaamheid'
Evaluatie	Beoordeling van de evolutie van data
Opmerkingen	Andere informatie relevant in dit project

De samenstelling van de steekkaart is hier volledig gericht op de werkzaamheden, met ander woorden: voorlopig wordt slechts elementaire informatie ingezameld om verdere discussies betreffende de relevantie en het gebruik van de indicator te ondersteunen. Pas in latere fase -wanneer er zekerheid is dat de indicator wordt gebruikt - zouden de beschrijvingen verder gedetailleerd worden.

Op het gebied van databeheer wordt trouwens meer en meer overwogen om de gegevens 'modulair' te op te slaan. Dit zou betekenen dat er een specifiek document wordt gemaakt waarin een bekommernis kan worden beschreven, een ander document waarin een meting wordt beschreven, nog andere voor de beschrijving van observaties, de normen (doelstellingen), etc...

De beschrijving van een indicator wordt dan een 'collectie' van documenten die aan elkaar worden gelinkt.

II.3 Enkele conclusies

De opdrachtgever van dit project, het Federale Wetenschapsbeleid, had al eerder een project afgerond waarin een database werd aangelegd met gegevens inzake personen, organisaties, projecten, documenten en databanken die handelen over Duurzame Ontwikkeling. Dit Informatiesysteem voor Duurzame Ontwikkeling (ISDO) kan worden uitgebreid met een databank 'Indicatoren voor Duurzame Ontwikkeling' (IDO), zo luidde de ambitie bij aanvang van dit project.

De voor de handliggende vraag is dan wat de inhoud en structuur van de gegevens over de IDO moet zijn. Kan er een standaardfiche of steekkaart worden ontwikkeld die tegemoet komt aan de noden of wensen van de talloze eindgebruikers. Moet men een standaard steekkaart ontwikkelen voor de beschrijving van één indicator en/of een collectie van indicatoren?

Het is deze vraag die in dit hoofdstuk werd belicht.

¹ Onderzoek in opdracht van het Ministerie van de Vlaamse Gemeenschap - Administratie Waterwegen en Zeewezen (Afdeling Waterwegen Kust) - <http://cdonet.ugent.be/onderzoek/kustzonebeheer.htm>

De conclusie is dat men moet opletten met de verwachting dat er voor IDO een gestandaardiseerde inhoudelijke structuur kan worden opgedrongen aan al diegenen die met IDO omgaan. Men moet niet veel IDO-projecten bekijken om vast te stellen dat de informatie op zeer veel verschillende manieren wordt aangeboden en dat er inhoudelijk zeer veel varianten bestaan. Ook de structuur waarin men de gegevens plaatst en verdeelt over verschillende documenten is geen constante. Het Centrum voor Duurzame Ontwikkeling (CDO) waarvan de auteurs van dit rapport deel van uitmaken is betrokken bij meerdere IDO-projecten en stelt na enige jaren vast dat er geen twee projecten zijn waarin de gegevens over de IDO op dezelfde manier werden gestructureerd. Voor het CDO zelf zou het handig zijn dat er een standaard steekkaart voor de beschrijving van een IDO kon worden gebruikt, maar omdat elke collectie van IDO wordt ontwikkeld in samenspraak met de eindgebruikers van de indicatoren (stad, regio, ...) blijkt telkens weer dat de lokale informatiebehoeften en wensen een nieuwe variëteit voortbrengen.

Hebben de verschillende projecten qua inhoud en gegevensstructuur dan niets met elkaar gemeen, zo luidde een tweede vraag. Daar kan men wel een positief antwoord op geven, maar daarbij wordt tegelijk vastgesteld dat het gemeenschappelijke vrij beperkt is.

Een Indicator voor Duurzame Ontwikkeling is een meetbare parameter, die -eventueel in combinatie met andere indicatoren - informatie verschaft over (een deelaspect van) de maatschappelijke ontwikkelingen. Men meet iets - men observeert iets - en dat helpt een oordeel te vellen over iets anders, zoals een arts eventueel de bloeddruk meet om tot een oordeel te komen over de gezondheid van een patiënt. In het algemeen vindt men bij de beschrijving van indicatoren dan ook een beschrijving van datgene wat gemeten (bijvoorbeeld: aantal mensen met inkomen beneden bepaald bedrag) wordt en een beschrijving van de bekommernis die middels die meting geëvalueerd wordt (bijv. armoede).

Maar de manier waarop men meting en bekommernis en hun onderlinge verband beschrijft kan zeer uiteenlopende vormen aannemen.

Wanneer men IDO-projecten bekijkt dan komt nog een ander gemeenschappelijk 'praktijk' naar voor. Duurzame Ontwikkeling wordt in het algemeen voorwerp van een multicriteria-evaluatie, waarbij de bekommernis 'duurzaamheid' opgedeeld wordt in deelaspecten (bijvoorbeeld: sociale, economische en ecologische ontwikkeling); deelaspecten die dan nog eens verder opgedeeld kunnen worden in deelaspecten (het ecologische opdelen in bodem, water en lucht bijvoorbeeld). Zo wordt het complexe begrip duurzaamheid ontrafeld in een hiërarchie van deelaspecten waarvan de uiteinden gekoppeld worden aan meetbare parameters. Indicatoren zitten dus meestal als bladeren aan een boom, waarvan de stam 'duurzaamheid' heet. Elke indicator op zich kan pas goed gewaardeerd worden als men die kan plaatsen binnen dit groter geheel.

Juist het feit dat men in IDO-projecten in het algemeen een soort van multicriteria-evaluatie ontwikkelt, maakt dat de informatie die relevant is met betrekking tot een bepaalde indicator, verspreid kan zitten over meerdere documenten, waarin de hiërarchie van criteria en subcriteria (deelaspecten) wordt ontwikkeld en toegelicht.

Zo kan het zijn dat men een paar indicatoren gebruikt voor de meting van waterkwaliteit en voor elke indicator op zich een steekkaart heeft die beschrijft hoe en wat precies gemeten wordt. Echter, hun relatie met de bekommernis waterkwaliteit en de samenhang tussen het paar indicatoren kan beschreven zijn in een document dat over de bekommernis waterkwaliteit gaat. Juist omdat geen van beide indicatoren op zich volstaat, worden ze dan ook samen besproken en geëvalueerd in functie van de bekommernis waterkwaliteit.

Men kan tegen deze praktijk moeilijk bezwaren aantekenen, maar het leidt wel tot de vaststelling dat het moeilijk wordt om een 'standaard-formulier' te ontwikkelen voor de beschrijving van een indicator.

Verder in dit rapport is de aandacht dan ook verschoven naar 'metadata' over IDO. Metadata zijn data over data. Metadata kent iedereen die een bibliotheek bezoekt en gebruik maakt van steekkaarten of een database die de beschikbare boeken beschrijven (titel, naam van auteur, e.d), maar niet de boeken vervangen. Terwijl voor elk boek het metadatafiche dezelfde structuur heeft, kunnen de boeken zelf op zeer uiteenlopende manier gestructureerd zijn. Ook voor IDO kan men overwegen om het streven naar gelijkvormigheid te beperken tot een 'standaard' voor metagegevens die verder de weg wijzen naar de gegevens zelf.

III. Metadata en Dublin Core Metadata Initiative (DCMI)

III.1 Inleiding

In dit hoofdstuk wordt het Dublin Core Metadata Initiative (DCMI) belicht. Deels wordt dit metadata initiatief gebruikt als educatief voorbeeld waarmee verschillende aspecten in de ontwikkeling van metadata kunnen worden geïllustreerd. Overigens is de Dublin Core hier gekozen omwille van het feit dat dit initiatief de laatste jaren een snel stijgende populariteit kon noteren op het Internet.

Het is belangrijk om te vermelden dat de Dublin Core al eerder werd geëvalueerd in het kader van het ISDO-project van de DWTC (Informatiesysteem Duurzame Ontwikkeling, zie hoofdstuk 1). Het ISDO-project heeft in een eerste fase specifieke metadata ontwikkeld voor de beschrijving van:

- een project
- een organisatie
- een publicatie
- een persoon
- een databank

zie: <http://www.belspo.be/issd/>

Voor elk type werd een specifieke metadatasteekkaart ontwikkeld met specifieke metadata-elementen.

Een voorbeeld daarvan staat in hoofdstuk 1 van dit rapport.

De reden waarom men in deze eerste fase de Dublin Core niet heeft toegepast is eenvoudig: de metadata-elementen die de Dublin Core aanbiedt werden ontoereikend geacht voor de beschrijving van de vijf informatiebronnen die hierboven werden vermeld. Wanneer men de term 'ontoereikend' gebruikt moet men er echter rekening mee houden dat men een oordeel vormt binnen een bepaalde context. Over de Dublin Core of andere metadata-elementen kan men geen absoluut oordeel vellen of algemeen antwoord geven op de vraag of deze of andere metadata volstaan of niet. De Dublin Core is ontwikkeld vanuit een bepaalde behoefte, zoals de metadata in ISDO eveneens ontwikkeld zijn vanuit een bepaalde behoefte of ambitie die kan verschillen. ISDO heeft bijvoorbeeld de ambitie om een service te leveren aan al diegenen die bezig zijn met duurzame ontwikkeling. Voor de Dublin Core was dit geen uitgangspunt. De Dublin Core is niet ontwikkeld voor één specifiek domein, dus ook niet voor 'duurzame ontwikkeling'.

Destijds - in de eerste ISDO-fase- speelde wellicht nog een ander aspect in de evaluatie van de Dublin Core. Hoewel de betrokken onderzoekers die de eerste fase van ISDO mee ontwikkelden (dat zijn niet dezelfde als de auteurs van dit rapport) de kwaliteiten van het Dublin Core Initiatief erkenden en waardeerden, waren er onvoldoende doorslaggevende argumenten om het gebruik daarvan te forceren. De Dublin Core metadata waren veel minder 'populair' dan vandaag het geval is.

Ondertussen is er binnen het Dublin Core Metadata Initiative (DCMI) veel evolutie, en is er een snelle groei zichtbaar van het gebruik van de Dublin Core metadata. Dat is dan ook de reden waarom dit initiatief hier wordt herbekeken en mogelijke pistes worden ontwikkeld om eventueel de Dublin Core te integreren binnen een project zoals het ISDO. Deze oefening is ook belangrijk omdat ze de discussie betreffende metadata in de context van Duurzame Ontwikkeling kan inspireren.

In een eerste deel wordt het Dublin Core Metadata Initiatief toegelicht en aandacht besteed aan de populariteit van dit initiatief dat - pas sinds enkele jaren - fors is toegenomen.

Vervolgens wordt bekeken wat binnen het kader van ISDO voor of tegen het gebruik van de Dublin Core pleit, om daaruit een soort compromis af te leiden waarin de ene optie de andere complementeert.

III.2 Dublin Core Metadata en Internet

In deze sectie wordt de Dublin Core gedemonstreerd, een paar 'spelregels' toegelicht en in kort bestek beschreven hoe men dergelijke data 'online' brengt op het World Wide Web. Belangrijk is wel dat de leek duidelijk het onderscheid leert te maken tussen afspraken met betrekking tot de inhoud van metadata (welke metadata bieden we aan of wisselen we uit?), waarbij de afspraak kan zijn dat men de Dublin Core metadata gebruikt, en meer technische afspraken omtrent data-uitwisseling via het Internet die volledig los staan van de inhoud en geldig kunnen zijn voor om het even welke set van metadata, Dublin Core en/of andere. Wat dat laatste betreft focussen we op de standaard van het W3C, de organisatie die verantwoordelijk is voor de World Wide Web standaarden.

Dublin Core: een voorbeeld

Een voorbeeld zegt veel meer dan een beschrijving en we beginnen dan ook met een demonstratie van een data-beschrijving met behulp van Dublin Core metadata (DC).

De basis van de Dublin Core Metadata is een set van 15 elementen waarmee een informatiebron op het World Wide Web kan worden beschreven.

Volgende metadata beschrijven een publicatie van het Centrum voor duurzame ontwikkeling

DC.Title: The Co-design of Indicators on Urban Sustainable Development

DC.Identifier: http://cdonet.ugent.be/English/Urban_Sustainable_Development.pdf

DC.Creator: Block T.; Van Assche J.

DC.Contributor:

DC.Date: 2001

DC.Publisher: Centre for Sustainable Development – Ghent University, Belgium

DC.Description: Paper for First International Conference on Ecology and the City - Virtual Forum, Barcelona, January-March 2001 - "No one can specify if his or her city or district is developing satisfactorily in general. Everyone can however give a partial answer (i.e. as regards employment, housing, safety, environmental quality, etc.), but is unable to put all the pieces of the puzzle together. To promote rational decision-making on, participation in and communication about the city's future, the Centre for Sustainable Development (CSD - Ghent University) has developed a Barometer for Sustainable Development. This instrument has been tested for the first time in Ghent (Belgium). The barometer forms part of preparations for the Ghent Local Agenda 21, a local implementation of the global Agenda 21. The development of a Barometer for Sustainable Development is thus one of the first steps that the city of Ghent is taking towards implementing this sustainable development policy. The 1999 Edition Barometer comprises a basic list of 60 indicators from 25 policy areas as well as a selection list with 15 indicators from 10 policy areas".

DC.Subject: Co-design of Sustainable Development Indicators, Instrument for Decision Making, Urban Development, Monitoring Urban Sustainability, City, Belgium, Ghent, Participation

DC.Format: pdf

DC.Type: article

DC.Source:

DC.Language: English

DC.Coverage: Ghent, Belgium; 1995-1999

DC.Relation: Website about project: <http://cdonet.UGent.be/onderzoek/bg/index.html>, includes complete set of indicators used for the Ghent Barometer for Sustainable Development and the "Selection list Edition 1999", derived from this list

DC.Relation: Block T., De Jonge W., Van Assche J., Doom R. (1999); De Gentse Barometer voor Duurzame Ontwikkeling: Eindrapport Oktober 1999 (In opdracht van het Stadsbestuur Gent) - printed report

DC.Relation: Block T. and Van Assche J., (2000), Stad op koers? De Gentse Barometer voor Duurzame Ontwikkeling, In: "Overheid in Beweging", Diegem: Kluwer Editorial, Nr 4, december 2000, 28 p.

DC.Rights: Centre for Sustainable Development – Ghent University - <http://cdonet.UGent.be/>

Bovenstaand voorbeeld demonstreert alvast enkele 'spelregels':

- Niet alle 15 elementen hoeven ingevuld te zijn. Men mag dus bepaalde gegevens achterwege laten.
- Meer nog: men kan lokaal een DC-initiatief op poten zetten waarbij men bijvoorbeeld werk met slechts 3 of vijf DC-elementen.

Onderstaand voorbeeld is dus evengoed een Dublin Core als het eerste:

DC.Title: The Co-design of Indicators on Urban Sustainable Development

DC.Identifier: http://cdonet.UGent.be/English/Urban_Sustainable_Development.pdf

DC.Creator: Block T.; Van Assche J.

DC.Contributor:

DC.Date: 2001

DC.Publisher: Centre for Sustainable Development – Ghent University, Belgium

Indien men van mening is dat voorlopig met deze zes gegevens informatiebeheer kan worden ondersteund, dan kan men het voorlopig daarbij laten.

Weliswaar zijn de metadata dan wat 'mager', aangezien ze nauwelijks het opsporen van relevante informatie ondersteunen. Er is dan geen subject-element (trefwoorden) en geen samenvattende beschrijving (*description*), vaak belangrijke elementen die het opsporen van informatie met zoeksoftware vergemakkelijken.

Twee andere spelregels:

- Een DC-element mag meermaals gebruikt worden
- De volgorde waarin ze worden gepresenteerd of uitgewisseld is niet van belang

In het eerste voorbeeld is bijvoorbeeld drie maal het element DC.Relation gebruikt, omdat de auteur van de metadata het beschreven artikel wou relateren aan drie andere databronnen die er op een of andere manier verband mee houden.

Evengoed had men voor elke auteur van het artikel een apart DC.Creator-element kunnen gebruiken. Wanneer een artikel op meerdere websites kan gedownload worden kan men meerdere DC.identifier elementen gebruiken. Is het artikel beschikbaar in meerdere formaten (pdf en Word bijvoorbeeld), dan kan men meerdere DC.Format-elementen gebruiken.

Als men een Word-document heeft én een pdf-versie dan zijn er twee databronnen die mogelijk elk hun eigen metadata krijgen. Weliswaar verschillen ze dan enkel betreffende identifier en formaat, en moeten ze eventueel aan elkaar gekoppeld worden door het DC-element 'relation'.

De volgorde is dus niet vast bepaald. Men kan weliswaar een persoonlijke voorkeur hebben betreffende de volgorde waarin metadata aan een eindgebruiker worden aangeboden, maar - en ook dat is een cruciaal punt - het Dublin Core Metadata Initiative 'moeit' zich daar niet mee. Hoe men deze data lokaal verwerkt, presenteert, publiceert -zowel qua structuur als stijl - laat men over aan de smaak van de betrokkenen. Het Dublin Core Initiative maakt afspraken die men moet respecteren als men metadata wil uitwisselen met andere personen en/of organisaties. In feite zijn het richtlijnen die vooral van belang zijn voor diegenen die software ontwikkelen waarmee DC-metadata worden ingezameld en verwerkt. Wanneer de software-ontwikkelaar vindt dat hij de resultaten van zijn inzameling moet publiceren met een specifieke volgorde van de DC-elementen, dan is dat zijn/haar probleem. Bij de inzameling moet hij/zij aanvaarden dat de aanbieders van DC-metadata een andere volgorde gebruiken en hij/zij zal in de software dan maar een sorteer- of rangschikkingoperatie moeten inbouwen, of een filteroperatie als men niet alle elementen wenst. Softwarematig zijn dat overigens 'elementaire' operaties die elke ingewijde moet aankunnen.

Metadata op het World Wide Web

Hoewel men voor de namen van de DC-elementen steeds Engelse termen gebruikt - DC.Title en niet DC.Titre - is dat voor de waarde of inhoud van die elementen helemaal niet het geval. Een Dublin Core mag inhoudelijk zelfs een mengeling van talen bevatten, zoals dat overigens gedemonstreerd is in het eerste voorbeeld.

Weliswaar zal men dan bij het uitwisselen van informatie enkele begeleidende codes moeten toevoegen - wat in het eerste voorbeeld niet is gebeurd.

Daarmee belanden we evenwel op een terrein dat niet DC-specifiek is, maar in eerste instantie te maken heeft met World Wide Web-standaarden. Wanneer men op het WWW metadata aanbiedt, met de bedoeling dat ze geconsulteerd zullen worden door zoekrobots dan kan men de spelregels volgen die door het W3C (World Wide Web Consortium) zijn vastgelegd. Dit zijn regels die gelden voor alle metadata- en niet uitsluitend voor de Dublin Core.

Concreet worden metadata in een webpagina dan als volgt gecodeerd:

```
<meta name="DC.Title" lang="en" content="Co-Design of Indicators" />
```

Sommige lezers hebben dit wellicht nog nooit gezien. Ze verschijnen overigens ook niet op het scherm wanneer een browser een pagina met die code ontvangt. Wat men meestal ziet zijn de data en niet de metadata.

De metadata die aldus zijn gecodeerd zijn dan ook bedoeld om 'gelezen' te worden door een zoekrobot, die de metadata dan eventueel hercodeert in een andere webpagina, zodanig dat ze wel zichtbaar worden op het scherm.

Mag de producent van metadata, zijn gegevens dan niet aanbieden in een 'zichtbare' vorm? Vanzelfsprekend, maar als men verwacht dat zoekrobots zoals Google e.a. de metadata raadplegen, dan moet men ze ook aanbieden volgens de strikte regels van het W3C.

Volgend voorbeeld moet de metadatapraktijk op het Internet illustreren. Organisatie CDO maakt een website en deze begint bij een homepage. Deze homepage bevat data en deze kan men zien als men ze met de browser download: zie <http://cdonet.UGent.be>

De webmaster van het CDO moet er evenwel voor zorgen dat zoveel mogelijk mensen die informatie zoeken inzake duurzame ontwikkeling, ook de website van het CDO raadplegen.

In eerste instantie is het dan van belang om zoveel mogelijk in het vizier te komen van webrobots zoals Google e.a. die het web afschuimen en metadata-banken aanleggen, die dan door bezoekers van Google kunnen geraadpleegd worden. Als de webmaster voor deze zoekrobots geen gestructureerde metadata aanbiedt, volgens de W3C-codes dan is het risico groot dat de robot een aantal woorden selecteert uit de data in de homepage, op basis van het vermoeden dat woorden die bovenaan voorkomen of vaak voorkomen mogelijks de beste trefwoorden zijn. Iedereen die Google e.a. raadpleegt heeft al kunnen vaststellen dat de geooogste trefwoorden vaak naast de kwestie zitten.

Dat is niet de schuld van de webrobots, maar veeleer van de webmasters die te weinig metadata aanbieden op een gestructureerde manier, volgens de WWW-standaarden.

De basisstandaard, waarbij men aan zijn webdocument gestructureerde metadata toevoegt, zoals in bovenstaand voorbeeld gaat ervan uit dat men de metadata kan opnemen in hetzelfde webdocument dat ook de data bevat en zodanig dat de metadata niet op het scherm komen bij diegene die de data wil lezen. Een webmaster kan ze in beeld brengen, maar het hoeft dus niet.

De metadata die opgeslagen liggen in de homepage van het CDO worden zichtbaar als men de 'broncode' of 'source-code' van de pagina bekijkt. Elke browser beschikt over een 'view'-knop met een menu waarin men 'source' kan selecteren. Als men deze functie aanklikt opent *Notepad* of andere text- of HTML-software waarin de volledige inhoud van de webpagina zichtbaar wordt. Als er gestructureerde metadata aangeboden worden zitten die bovenaan in het zogenaamde head-element van de webpagina. Dit head-element kan naast metadata nog allerlei codes bevatten die niet op het canvas van de browser moet verschijnen. Wat wel op het scherm verschijnt, is de inhoud van het body-element (zie volgend hoofdstuk over HTML, XML en XHTML).

De vaste minimale structuur van een webpagina is als volgt gecodeerd:

```
<html>
<head>
Hier zit code die browser bepaalde opdrachten geeft, bijvoorbeeld stijl-
opdrachten (kleur, lettertype), én metadata voor webrobots
</head>
<body>
Hier staat inhoud die op scherm moet verschijnen
</body>
</html>
```

De zogenaamde html-tags - <html> en </html> - melden de browser waar de pagina begint en eindigt. De head- en body-tags delen de inhoud op in een deel dat niet op het scherm moet verschijnen en een deel dat wel moet verschijnen. Webrobots 'lezen' beide en zullen metadata in het head-element oogsten, mits ze daar aanwezig zijn.

Bijvoorbeeld:

```
<html>
<head>
<meta name="DC.Title" content="Centrum voor Duurzame Ontwikkeling" />
<meta name="DC.Creator" content="Stefan Deconinck" />
<meta name="DC.Subject" content="duurzaamheid; duurzaamheidsbarometer;
leefbaarheid; Agenda 21; verkeer; mobiliteit; verkeersveiligheid;
snelheidsbegrenzer; ISA; leefbaarheid; indicatoren; duurzame ontwikkeling;
voorraadbeheer; factor 10; factor 4; onderwijs; onderzoek; publications;
publicaties; milieugebruiksruimte; environmental utilization space; eau;
water; ontwikkelingssamenwerking; ketenanalyse; ecologiecriterium;
milieuzorg; woordenboek; universiteit; gent; wetenschap; België; Belgique;
développement durable; durabilité; recherches; centre; gand; ecologie;
écologie; behoeften; kustzonebeheer; integraal" />
</head>
<body>
....
</body>
</html>
```

Het belangrijkste dat men met de W3C-standaard kan concluderen is dat er voor het Internet - het World Wide Web - een internationaal aanvaarde standaard praktijk bestaat voor het aanbieden en uitwisselen van metadata op het web en dat:

- deze standaard geen beperkingen oplegt inzake de hoeveelheid, aard en inhoud van de metagegevens
- geen beperkingen oplegt inzake taalgebruik
- eenvoudig en eenduidig is

Volgende voorbeeld is een perfecte toepassing van die standaard:

```
<meta name="DC.Title" lang="en" content="Co-Design of Indicators" />
<meta name="DC.Title" lang="nl" content="Cooperatief ontwerp van
indicatoren" />
<meta name="DC.Creator" content="Block T.; Van Assche J." />
<meta name="AC.Name" content="Johnny" />
<meta name="DC.Date" scheme="W3C-DTF" content="2000-06-11" />
```

Het vierde meta-element behoort niet tot de Dublin Core, maar is een element uit de A-Core (Administrative Core), die we verder nog toelichten. In dit geval wordt met AC.name aangegeven dat Johnny de auteur is van de metadata, niet de auteur van het document "Co-Design of Indicators", want daarvoor zijn dus Block en Van Assche verantwoordelijk.

Een combinatie van DC-elementen met andere metadata-elementen is dus mogelijk. Een interessant gegeven dat we verder nog zullen exploiteren.

Tevens kan men dus dezelfde metadata in verschillende talen tegelijk opgeven, en zoekrobots sturen dor middel van het gebruik van het lang-attribuut (language) waarmee aangegeven wordt in welke taal het content-attribuut ingevuld is. De codes "en" (Engels) en "nl" (Nederlands) zijn vastgelegd in een internationale Internet-standaard, met name¹ RFC1766.

In de laatste regel van het laatste voorbeeld duikt ook een attribuut 'scheme' op. Hier wordt gestipuleerd dat de datum geschreven is volgens het Date-Time-Format van het W3C, wat wil zeggen: eerst komt het jaar, vervolgens de maand en dan de dag (in het voorbeeld is de datum dus 11 juni 2000 en niet 6 november).

Het attribuut scheme wordt ook ingeschakeld als men een gecontroleerde vocabulair of thesaurus gebruikt. Stel bijvoorbeeld dat men binnen het kader van een informatiesysteem inzake duurzame ontwikkeling afspreekt te werken met de GEMET-thesaurus (zie hoofdstuk 1), dan kan men dit als volgt combineren met het DC-element voor trefwoorden DC.Subject:

```
<meta name="DC.Subject" scheme="GEMET" content="trefwoorden uit Gemet-  
vocabulary" />
```

Zijn alle webrobots dan vertrouwd met GEMET? Neen. Vandaag zullen de meeste robots niets anders doen dan de trefwoorden opnemen zoals alle andere trefwoorden die met een DC.Subject worden opgevoerd. Maar als men binnen een bepaalde groep van organisaties expliciet gebruik wil maken van GEMET, om metadata thematisch te classificeren bijvoorbeeld, dan kan men de scheme-boodschap wel exploiteren. De robotsoftware of software voor de archivering van metadata moet dan wel aangepast worden natuurlijk.

Voor de volledigheid moet hierbij nog worden vermeld dat met betrekking tot metadata en het World Wide Web nog andere belangrijke ontwikkelingen gaande zijn, die we hier niet aan bod laten komen omdat het te maken heeft met geavanceerde internet-standaarden op basis van XML (eXtended Markup Language), RDF (Resource Description Framework), en dergelijke. Tijdens ons onderzoek stelden we vast dat terzake nog geen standaard praktijk bestaat die op brede schaal wordt toegepast. Er circuleren voorstellen en wellicht zal dit op termijn wel een consensus opleveren.

Voor de meeste lezers zou het een geruststelling moeten zijn te weten dat de productie van gegevens zoals die hierboven staan gecodeerd, geen problemen levert wanneer men later naar meer geavanceerde WWW-technieken zoals het Resource Description Framework zou willen overschakelen.

Lezers die zich niet kunnen of willen verdiepen in de technische materie die hierboven werd aangekaart, kunnen deze zaken wel degelijk overlaten aan hun lokale webmaster. Het belangrijkste is wel dat men weet dat de Internet- en WWW-standaarden volgende zaken gelden:

¹ RFC1766 definieert en verklaart de taal-codes die men gebruikt in een lang-attribuut in webpaginas. Samengevat komt het hierop neer: een taal-code bestaat uit een primaire code en eventueel een serie van subcodes:

```
language-code = primary-code ( "-" subcode )*
```

Enkele voorbeelden:

- "en": Engels
- "en-US": de U.S. versie van Engels.
- "en-cockney": de Cockney versie van Engels.
- "i-navajo": de Navajo taal
- "x-klinton": een experimentele taal, zijnde klinton

De twee-letter primaire code (vaak de enigedie men gebruikt) baseert zich op de ISO639-standaard voor taal-afkortingen. Bij deze twee-letter codes heeft men bijvoorbeeld: fr (Frans), de (Duits), it (Italiaans), nl (Nederlands), es (Spaans), ar (Arabisch), he (Hebreeuws, ru (Russisch), zh (Chinees), ja (Japans). Elke twee-letter subcode wordt geïnterpreteerd als een ISO3166 landen-code.

- er zijn gestandaardiseerde codes voor de uitwisseling van metadata op het Internet
- de standaarden zijn in eerste instantie bedoeld voor de webmasters die metadata aanbieden en de ontwikkelaars van webrobots, zoek-software en metadata-archieven die metadata verzamelen en eventueel met andere geëigende instrumenten zichtbaar en consulteerbaar maken voor eindgebruikers (online-database bijvoorbeeld)
- de webstandaarden laten toe om verschillende metadata te mixen (Dublin Core met andere metadata)
- het is mogelijk om met een mengeling van verschillende talen te werken
- men kan gebruik maken van een gecontroleerde woordenschat of thesaurus

Na deze technische aspecten keren we terug naar het inhoudelijke.

III.3 Dublin Core: verfijning en uitbreiding

Als men metadata zoals Dublin Core beoordeeld, dan is het niet onbelangrijk om niet alleen de samenstelling van de metadata te bekijken, maar ook een evaluatie te maken van de slagkracht van diegenen die deze metadata propageren. Het ontwikkelen van een metadata-set is één zaak. Het gebruik daarvan gangbaar maken is evenwel andere koek. Een evaluatie van het Dublin Core Metadata Initiatief in zijn geheel - en niet alleen van de Dublin Core metadata- kan er mogelijk toe leiden dat men er anders tegenaan kijkt. Immers wanneer men de Dublin Core niet gebruikt, verliest men tegelijk alle diensten die binnen het DCMI worden aangeboden.

Alle informatie met betrekking tot de Dublin Core kan gevonden worden op de website van het Dublin Core Metadata Initiative.

<http://dublincore.org>

Hier kunnen we geen correct overzicht garanderen, aangezien de ontwikkelingen elk overzicht snel gedateerd maken. Bedoeling is vooral aan te geven wat een metadata-initiatief kan inhouden.

"The Dublin Core Metadata Initiative (DCMI) is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models. Its primary offering, the Dublin Core Metadata Element Set, is the de facto worldwide standard for the description of information resources across disciplines and languages. DCMI's activities include consensus-driven working groups, global workshops, conferences, standards liaison, and educational efforts to promote widespread acceptance of metadata standards and practices".

"Dublin Core metadata is being used as the basis for descriptive systems by several community interest groups such as:

- *educational organizations*
- *libraries*
- *government institutions*
- *scientific research sector*
- *Web page authors*
- *businesses requiring more searchable sites*
- *corporations with vast knowledge management systems"*

Deze citaten die we van de Dublin Core - site hebben geplukt vatten de enorme ambities van het initiatief bondig samen. Niet alleen organiseert het DCMI een open forum waarin iedereen kan participeren, het staat ook in voor het sturen van het zoeken naar consensus, middels werkgroepen, congressen, educatie, enz.

Het ontwikkelen van metadata vraagt management. Niet alleen moet men het met een groep eens worden over de metadata-elementen, maar ze moeten ook eenduidig gedefinieerd worden en bij de Dublin Core die internationale ambities heeft moeten definities bij voorkeur worden vertaald, moeten

handleidingen worden gemaakt die de gebruiker leren omgaan met metadata, enzovoort. Op de website van het DCMI worden al deze diensten aangeboden.

We belichten eerste de definiëring van de 15 basis-elementen. Hier worden de Engelse definities deels gekopieerd, hoewel er ook Nederlandse en Franse vertalingen bestaan. Geen onbelangrijk detail: vandaag moet men versie 1.1 van de Dublin Core gebruiken. Versie 1.0 is achterhaald.

De 15 elementen focussen op drie informatie-types: inhoud, intellectueel eigendom en medium (de manier waarop de data worden aangeboden).

Inhoud	Intellectueel Eigendom	Medium
Coverage	Creator	Date
Description	Contributor	Format
Type	Publisher	Identifier
Relation	Rights	Language
Source		
Subject		
Title		

Beschrijving van metadata elementen:

element	element description
dc.title	<p>The name given to the resource. Typically, a Title will be a name by which the resource is formally known.</p> <p>If in doubt about what constitutes the title, repeat the Title element and include the variants in second and subsequent Title iterations... .</p>
dc.subject	<p>The topic of the content of the resource. Typically, a Subject will be expressed as keywords or key phrases or classification codes that describe the topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.</p> <p>Select subject keywords from either the Title or Description information. If the subject of the item is a person or an organization, use the same form of the name as you would if the person or organization were a Creator, but do not repeat the name in the Creator element.</p> <p>In general, choose the most significant and unique words for keywords, avoiding those too general to describe a particular item. This element might well include classification data (for example, Library of Congress Classification Numbers or Dewey Decimal numbers) or controlled vocabularies (such as Medical Subject Headings or Art and Architecture Thesaurus descriptors) as well.</p>
dc.description	<p>An account of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.</p> <p>Since the description field is a potentially rich source of indexable vocabulary, care should be taken to provide this element when possible. Some metadata collections could include content description (spectral analysis of a visual resource, for example) that may not be embeddable in current network systems. In such a case this field might contain a link to such a description rather than the description itself.</p> <p>Descriptive information can be taken from the item itself, if there is no abstract or other structured description available. Normally, if a Description cannot be found either in the introductory or front matter, or the first few paragraphs, it should be set up "on the fly" by the metadata provider. Normally, Description should be limited to a few brief sentences.</p>
dc.type	<p>The nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the working draft list of Dublin Core Types</p>

	<p>List of Resource Types. Dublin Core Draft Working Group Report, http://dublincore.org/documents/dcmi-type-vocabulary/</p> <p>To describe the physical or digital manifestation of the resource, use the FORMAT element.</p> <p>If the resource has content of multiple mixed types then multiple or repeated Type elements should be used to describe the main components.</p>
dc.source	<p>A Reference to a resource from which the present resource is derived. The present resource may be derived from the Source resource in whole or part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.</p> <p>In general, include in this area information which does not fit easily into Relation.</p>
dc.relation	<p>A reference to a related resource. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.</p> <p>The recommended method for expressing a relationship in qualified DC is: Title="the present resource"</p> <p>Relation="relationship-type [space] unique identifier for the related resource" where "relationship-type" is a token drawn from the approved list of qualifiers.</p> <p>Note: In the case where the DC metadata is embedded in the present resource, the value for Identifier is implied (i.e. the present resource). In qualified DC the two components given in Relation here will be structured using sub-elements for easier automated processing.</p>
dc.coverage	<p>The extent or scope of the content of the resource. Coverage will typically include spatial location (a place name or geographic co-ordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [Getty Thesaurus of Geographic Names, http://shiva.pub.getty.edu/tgn_browser/]) and that, where appropriate, named places or time periods be used in preference to numeric identifiers such as sets of co-ordinates or date ranges.</p> <p>Whether this element is used for spatial or temporal information, care should be taken to provide consistent information that can be interpreted by users. For most simple applications, where place names or coverage dates might be useful, whether the information is numeric or alphabetical may be enough to differentiate. For more complex applications, consideration should be given to additional qualification.</p>
dc.creator	<p>An entity primarily responsible for making the content of the resource. Examples of a Creator include a person, an organisation, or a service.</p> <p>Guidelines for creation of content: Creators should be listed separately in the same order that they appear in the publication. Personal names should be listed surname or family name first, followed by forename or given name. When in doubt, give the name as it appears, and do not invert.</p> <p>In the case of organizations where there is clearly a hierarchy present, list the parts of the hierarchy from largest to smallest, separated by full stops. If it is not clear whether there is a hierarchy present, or unclear which is the larger or smaller portion of the body, give the name as it appears in the item.</p> <p>If the Creator and Publisher are the same, do not repeat the name in the Publisher area. If the nature of the responsibility is ambiguous, the recommended practice is to use Publisher for organizations and Creator for individuals. In cases of lesser responsibility, other than creation, use Contributor.</p>
dc.publisher	<p>The entity responsible for making the resource available. Examples of a</p>

	<p>Publisher include a person, an organisation, or a service. Typically, the name of a Publisher should be used to indicate the entity.</p> <p>The intent of specifying this field is to identify the entity that provides access to the resource. If the Creator and Publisher are the same, do not repeat the name in the Publisher area. If the nature of the responsibility is ambiguous, the recommended practice is to use Publisher for organizations and Creator for individuals. In cases of lesser responsibility, other than creation, use Contributor.</p>
dc.contributor	<p>An entity responsible for making contributions to the content of the resource. Examples of a Contributor include a person, an organization or a service. Typically, the name of a Contributor should be used to indicate the entity.</p> <p>The same general guidelines for using names of persons or organizations as Creators apply here.</p>
dc.rights	<p>Information about rights held in and over the resource. Typically a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.</p> <p>At present, used for either a textual statement or a URL.</p>
dc.date	<p>A date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [Date and Time Formats, W3C Note, http://www.w3.org/TR/NOTE-datetime] and follows the YYYY-MM-DD format.</p> <p>If the full date is unknown, month and year (YYYY-MM) or just year (YYYY) may be used. Many other schemas are possible, but if used, they may not be easily interpreted by users or software.</p>
dc.format	<p>The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [http://www.isi.edu/in-notes/iana/assignments/media-types/media-types] defining computer media formats).</p> <p>In principle, formats can include physical media such as books, serials, or other non-electronic media.</p> <p>Information concerning the size of a resource may be included in the content of the Format element if available. In resource discovery this might be used as a criterion to select resources of interest, since a user may need to evaluate whether they can make use of the resource within the infrastructure available to them.</p>
dc.identifier	<p>An unambiguous reference to the resource within a given context. Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. An example of formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).</p> <p>This element can also be used for local identifiers (e.g. ID numbers or call numbers) assigned by the Creator of the resource to apply to a particular item.</p>
dc.language	<p>A language of the intellectual content of the resource. Recommended best practice for the values of the Language element is defined by RFC 1766 [Tags for the Identification of Languages, Internet RFC 1766, http://info.internet.isi.edu/in-notes/rfc/files/rfc1766.txt] which includes a two letter</p>

	<p>Language Code (taken from the ISO 639 standard [Codes for the representation of names of languages, ISO 639, http://www.oasis-open.org/cover/iso639a.html]), followed optionally, by a two letter Country Code (taken from the ISO 3166 standard [Codes for the representation of names of countries, ISO 3166, http://www.oasis-open.org/cover/country3166.html]). For example, "en" for English, "fr" French or "en-uk" for English used in the United Kingdom.</p> <p>Coded or textual information can be represented here. If the content is in more than one language, the element may be repeated.</p>
--	---

Bovenstaande lijst toont aan dat de definities van meta-elementen technisch zwaarwichtig worden als men precisie of eenduidigheid nastreeft. Eigenlijk leert men de Dublin Core best kennen door voorbeelden. Maar als men twijfelt, kan men nog steeds de definities raadplegen in het officiële document: <http://dublincore.org/documents/dces/>

De vertalingen die men op de site van het DCMI kan raadplegen zijn wellicht toegankelijker (zie: <http://dublincore.org/resources/translations/>), maar nogmaals dient beklemtoond dat bij uitwisseling van DC-metadata via het Word Wide Web de Engelse namen van de DC-elementen worden gebruikt. Indien men metadata aan een Franstalig eindgebruiker wil tonen, dan kan men een presentatie gebruiken in de volgende vorm:

DC.Creator (créateur): Emile Zola

De Franstalige lezer is bediend én het internationale publiek is bediend.

Nogmaals: de inhoud zelf van de metadata-elementen mag in om het even welke taal geschreven zijn.

Refinements

Het DCMI is niet gestopt bij deze vijftien basiselementen. De 'vijftien' kunnen in een andere meer gedetailleerde vorm worden gebruikt, waardoor men in feite de beschikking krijgt over 34 metadata-elementen.

Zo kan men niet alleen DC.Title gebruiken, maar ook DC.Title.Alternative, wat als volgt is gedefinieerd:

"Any form of the title used as a substitute or alternative to the formal title of the resource. This qualifier can include Title abbreviations as well as translations".

Vaak hebben projecten een lange naam en wordt een accroniem gebruikt in de omgangstaal. Dit kan men dus kwijt in DC.Title.Alternative.

Volgens DCMI kan men ook vertalingen van de titel kwijt in DC.Title.Alternative. We vinden het een slecht advies en verwijzen terzake naar de eerder belichte Internet-standaard die een voorziet in het gebruik van het attribuut lang (language). Alleen zo vangt men op structurele wijze 'meertaligheid' op, en ook het DCMI meldt dat via sommige handleidingen. Anders had het DCMI een 'alternative' moeten voorzien voor elk DC-element, wat dus niet het geval is.

De taal van de 'originele' of 'oorspronkelijke' titel, kan men achterhalen door de inhoud van het DC.Language element.

In bovenstaande definitie heeft men het over een 'qualifier'. De 'verfijning' van Title met Alternative is één type qualifier, naast het gebruik van een 'codeer-schema', waarover we het reeds eerder hadden in het luik dat WWW-code beschreef (toepassing van standaard-notatie voor de datum bijvoorbeeld). Men kan dus verschillende DC.Subject-elementen (trefwoorden) hebben die elk een specifieke gecontroleerde woordenschat (thesaurus) hanteren.

Voorlopig blijven we even halt houden bij de DC-qualifiers die bekend staand als 'refinement', verfijning.

Men kan DC.Description verfijnen door:

DC.Description.Abstract

of

DC.Description.TableOfContents

De termen spreken allicht voor zich en wie toch mocht twijfelen kan de definities raadplegen:

<http://dublincore.org/documents/dcmes-qualifiers/>

Hieronder tonen we enkel een lijst van alle beschikbare refinements.

DCMES Element	Element Refinement(s)	Element Encoding Scheme(s)
Title	Alternative	-
Creator	-	-
Subject	-	LCSH MeSH DDC LCC UDC
Description	Table Of Contents Abstract	-
Publisher	-	-
Contributor	-	-
Date	Created Valid Available Issued Modified	DCMI Period W3C-DTF
Type	-	DCMI Type Vocabulary
Format	Extent	-
	Medium	IMT
Identifier	-	URI
Source	-	URI
Language	-	ISO 639-2 RFC 1766
Relation	Is Version Of Has Version Is Replaced By Replaces Is Required By Requires Is Part Of Has Part Is Referenced By References Is Format Of Has Format	URI
Coverage	Spatial	DCMI Point ISO 3166 DCMI Box TGN
	Temporal	DCMI Period

		W3C-DTF
Rights	-	-

Deze verfijningen kunnen, maar moeten niet gebruikt worden. Merk op dat er nog steeds 15 elementen zijn. Een organisatie die zonder verfijningen werkt zal dus eventueel DC.Title.Alternative verwerken als een tweede DC.Title element.

De mogelijkheden qua databeheer nemen met deze verfijning sterk toe.

Dublin Core Structured Values

Er is nog een belangrijke optie die een antwoord biedt op specifieke lokale noden die op het eerste zicht niet kunnen bevredigd worden binnen de spelregels van de Dublin Core.

Stel bijvoorbeeld dat men lokaal een ondubbelzinnig onderscheid wil maken tussen voornaam en familienaam. Ondubbelzinnig wil zeggen dat men later heel doelgericht door gebruik van software alle familienamen uit de metadata kan plukken.

Daarvoor propageert het DCMI het gebruik van zogenaamde Dublin Core Structured Values (DCSV).

Een voorbeeld van DCSV maakt veel duidelijk:

```
<meta name="DC.Creator" scheme="DCMIDCSV" content="name.given=Thomas ; name.family=Block; email=thomas.block@UGent.be" />
```

De gestructureerde waarden (van de meta-elementen) zijn dus tekst met volgende patronen:

- "u1; u2; u3"
- "cA=v1"
- "cA=v1; cB.part1=v2; cB.part2=v3"
- "cA=v1; u2; u3"

waarbij:

- u1, u2 en u3 componenten zijn zonder label,
- cA en cB labels zijn van de gestructureerde waardecomponenten,
- v1, v2 en v3 waarden van de componenten zijn
- part1 en part2 verfijningen zijn van het label.

Drie specifieke tekens worden gebruikt:

- gelijkheidstekens scheiden label en waarde binnen een component
- punt-komma scheidt componenten
- punt is indicatie voor hiërarchische structuur in labels

Deze gestructureerde waarden moeten vooral gezien worden als een middel waarmee men vaak een oplossing kan vinden voor specifieke lokale behoeften die men wil bevredigen binnen de spelregels van de Dublin Core.

Meer hierover: <http://dublincore.org/documents/2000/07/28/dcmi-dcsv/>

A-Core: Administrative Core

Indien de Dublin Core nog steeds ontoereikend is, dan kan men de aanvulling met andere metadata overwegen. Een eerste - vaak gebruikte - complementaire metadata-set is de A-core of 'administrative core' die informatie verschaft over o.a. de auteur en status van de metadata. De A-Core (AC) bevat

dus meta-metadata omdat het informatie is over de metadata, en speelt vaak een belangrijke rol in het beheer van metadata.

Belangrijkste referentie op het gebied van de A-Core was (begin 2002) nog steeds het ontwerp van R. Iannella en D. Campbell, meer bepaald versie 2.0.

Zie daarvoor: <http://metadata.net/admin/draft-iannella-admin-01.txt>

Terwijl we met de Dublin Core beschrijven wie een bepaald artikel heeft gemaakt, wordt dus met de A-Core beschreven wie de metadata heeft geproduceerd. Dat kan dezelfde persoon zijn, maar dat is niet noodzakelijk het geval. De auteur van de data kan reeds lang dood zijn als de metadata worden gepubliceerd.

Bijvoorbeeld:

```
<meta name = "AC.name" content = "Walter De Jonge" />
<meta name = "AC.email" content = "walter.dejonge@Ugent.be" / >
<meta name = "AC.contact" content = "http://cdonet.Ugent.be " />
<meta name = "AC.activity" content = "created" />
<meta name = "AC.date" content = "2001-04-11" />
<meta name="DC.Creator" content="Karl Marx" />
<meta name="DC.Title" content="Das Kapital; Kritik der Politischen
Oekonomie" />
<meta name="DC.Date" content="1867" />
```

In dit geval heeft Walter De Jonge de metadata gecreëerd op 11 april 2001. Maar Marx zelf heeft dus wel de data gecreëerd.

Het element AC.activity heeft hier de waarde "created", maar kan bijvoorbeeld ook "modified" zijn. Soms ziet men een opeenvolging van meerdere A-Cores, waarbij bovenaan een A-core staat die werden gemaakt op het ogenblik van de creatie van de metadata, vervolgens een A-core die een wijziging meldt.

Men kan ook AC.dateRange gebruiken waarmee men aangeeft gedurende welke periode de metadata (dublin Core en andere) geldig zijn. Hiermee kan men dus een vervaldatum aankondigen. Overigens is er ook een AC.rights waarin men eventueel auteursrechten betreffende de metadata vermeld.

Verder is er het element AC.location dat gebruikt wordt wanneer de A-Core gescheiden opgeslagen is van de metadata zelf (location is dan meestal het webadres of de URL van de metadata). Een praktijk die zelden wordt toegepast. AC en DC zitten meestal in dezelfde metadata-bron.

Binnen het DCMI is er wel een discussie aan de gang betreffende verdere ontwikkeling van de A-core. Zie: <http://dublincore.org/groups/admin/>

Het is nu wellicht voldoende gedemonstreerd dat men verschillende metadata kan onderscheiden middels een prefix, zoals DC en AC, waardoor er een duidelijk verschil is tussen DC.Date en AC.Date.

EdNA: Education Network Australia

Deze piste wordt verder bewandeld door het Australische EdNA (Education Network Australia): <http://www.edna.edu.au/index.html>

EdNA focust op educatieve databronnen en gebruikt de Dublin Core aangevuld met volgende elementen:

EDNA.Audience: A category of user for whom the resource is intended

EDNA.Approver: Email of person or organisation approving the item for inclusion in EdNA Online

EDNA.CategoryCode: A numerical code derived from the database tables which support the EdNA Online Browse Categories.

EDNA.Entered: Date item was entered as an entry in the EdNA Online item database

EDNA.Indexing: A parameter which identifies to what extent the EdNA Online indexing ("spidering") software should follow hyperlinks from the described page.

EDNA.Review: A third-party commentary or formal review of the resource. This element is defined such that two forms of review are accommodated.

EDNA.Reviewer: Name of person and/or organisation or authority affiliated with the review.

EDNA.Version: Version of the EdNA Metadata Standard applied

Zie voor eventuele update: <http://www.edna.edu.au/edna/go/pid/385>

III.4 Evaluatie van Dublin Core

Het DCMI werkt voortdurend aan de promotie van de Dublin Core, blijkbaar niet zonder succes, want men heeft effectief de status bereikt van 'de facto'-standaard op het Internet, wat dus wil zeggen dat er nergens formeel beslist werd dat iedereen de Dublin Core moet gebruiken, maar wel dat het de meest gebruikte standaard is geworden. Overigens is dat een uitspraak die men pas sinds enkele jaren mag doen. De populariteit van de Dublin Core is eigenlijk voorwerp van een exponentiële groei. Elk nieuwe gebruiker is automatisch promotor van de Dublin Core en kan 'twijfelaars' mee overtuigen.

Een lijst van op Dublin Core gebaseerde projecten is beschikbaar: <http://dublincore.org/projects/>

Wel dient beklemtoond dat de toenemende populariteit van de Dublin Core niet uitsluitend gebaseerd is op de volgzzaamheid van twijfelaars. Al een aantal malen is de Dublin Core ten gronde doorgelicht en vergeleken met andere opties, om finaal met een goede score uit de evaluatie te komen. Dergelijke kritische gebruikers komen tot een positief oordeel wanneer ze als voornaamste criteria voor de keuze van metadata volgende vooropstellen:

- eenvoud
- flexibiliteit
- brede toepasbaarheid

De Dublin Core is ontworpen voor de beschrijving van data-bronnen op het World Wide Web. Die databronnen kunnen om het even wel onderwerp behandelen. De Dublin Core is dus niet ontwikkeld voor een specifiek domein of onderwerp. Evenmin had men een specifieke doelgroep voor ogen, tenzij de gebruiker van het World Wide Web, maar dat kunnen dus evengoed kinderen als bejaarden zijn, experts en leken.

Een onderzoeksproject waar men een dergelijk grondige evaluatie heeft gemaakt op basis van vermelde criteria is het Nordic Metadata Project van Scandinavische metadata experts. Zij vergeleken verschillende opties en concludeerden in hun finaal rapport (July 1998):

"During the lifetime of the Nordic metadata project the Dublin Core Metadata Element Set has gained a strong status as the format for the description of document-like Internet objects. At the moment it does not have any serious rivals – if we do not take domain-specific formats such as MARC as competitors with the more generic Dublin Core. We do not see the relationship of Dublin Core and domain specific formats as a competitive one. On the contrary, due to its flexibility the Dublin Core is an ideal exchange format, which can enable us to pass bibliographic information from one community to another in the same way that we now can exchange data between library systems thanks to ISO 2709 Exchange format."

Het rapport kan worden gedownload: <http://www.lib.helsinki.fi/meta/nmfinal.htm>

De homepage van het Nordic Metadata Project: <http://www.lib.helsinki.fi/meta/index.html>

Merk op dat men in Scandinavië de Dublin Core niet alleen hoge scores geven op basis van genoemde criteria, maar eveneens in de Dublin Core een goed glijmiddel zien om verschillende metadata-standaarden te koppelen, t.t.z. om metadata uit te wisselen tussen verschillende standaarden.

Stel bijvoorbeeld dat organisatie of netwerk X werkt met metadata-standaard A die dertig metadata-elementen bevat, weliswaar toegespitst op het domein 'sociologie'. Een andere organisatie of netwerk Y werkt met metadata-standaard B die deels andere elementen gebruikt toegespitst op economische informatie.

In elk geval zal men inhoudelijke gelijkenissen vinden tussen A en B. Beiden zullen wellicht wel een element hebben dat ruimte geeft voor de titel van de databron, de auteurs en de publicatiedatum. Mogelijks heet het bij A 'author' en bij B 'editor', maar beide kunnen dus hetzelfde bedoelen, voornamelijk datgene wat in de Dublin Core gedefinieerd is als 'creator'.

Wanneer de twee netwerken metadata willen uitwisselen zullen ze moeten uitzoeken welke metadata in A overeenkomen met metadata-elementen in B. Stel nu dat er een derde actor wil aansluiten met metadata-standaard C, dan zou men opnieuw een oefening moeten maken, waarbij A met C én B met C wordt gekoppeld.

Enkele lezers zullen wellicht opmerken dat een koppeling tussen A en C, automatisch een koppeling tussen B en C voortbrengt aangezien A al met B was gerelateerd. Maar zo eenvoudig liggen de zaken niet.

Elke overdracht van het ene systeem naar het andere houdt vaak enkele 'schoonheidsfoutjes' in. Sommige metadata-elementen zijn bijna hetzelfde, maar niet helemaal. Waar de overdracht van het ene naar het andere systeem nog net aanvaardbaar is, kan bij een tweede conversie -opnieuw met een kleine fout- een té grote fout ontstaan.

Wanneer men afspreekt om de Dublin Core als exchange-format te gebruiken, zal men voor elke nieuwkomer slechts de relatie moeten leggen tussen de elementen van de nieuwkomer en de Dublin Core. Daarmee zijn niet alle standaarden perfect twee-aan-twee met elkaar gekoppeld, maar men kan toch een goede basis leggen voor de uitwisseling van metadata. Niets verbiedt de eindgebruiker om naast de gedeelde Dublin Core metadata, ook nog steeds de metadata volgens standaarden A, B en C te consulteren. Het feit dat men de Dublin Core gebruikt als standaard voor uitwisseling, betekent dus niet dat men zijn domeinspecifieke standaard niet langer gebruikt, integendeel.

Zo kan men bijvoorbeeld vaststellen dat de experts inzake GIS (grafische informatie systemen) weliswaar hun domeinspecifieke metadata-standaard ontwikkelden, maar tegelijk voorzien in de productie van Dublin Core, waarin dan slechts een deel van de 'professionele' metadata worden opgenomen, maar tegelijk aansluiting wordt gezocht met andere metadata-bronnen. In dit GIS-kader wordt de Dublin Core deels geprofileerd als een metadata-standaard waarmee men een breder publiek bedient, naast de specifieke 'professionele' metadata-standaard.

Zie: <http://www.diffuse.org/gis.html>

Voorlopig stellen we vast dat de Dublin Core in drie toepassingsgebieden veld windt:

- daar waar men zoek naar een eenvoudige, flexibele en algemeen toepasbare metadata-set
- daar waar men zoekt naar een formule voor uitwisseling van metadata tussen verschillende metadata-systemen of een consensus zoekt tussen verschillende domeinspecifieke metadata-initiatieven (waarbij elk zijn domein-specifiek aanbod kan blijven handhaven)
- daar waar men onderscheidt maakt tussen 'professionele' gedetailleerde metadata en een vereenvoudigde verkorte versie voor het breder publiek

III. 5 Conclusies met betrekking tot Dublin Core

In het eerste hoofdstuk werden verschillende metadata-initiatieven in het domein duurzame ontwikkeling aangekaart. Hoewel sommige van deze initiatieven zich niet expliciet aankondigen als metadata-initiatief gebruiken ze wel metadata om het opsporen of bekendmaken van informatie te ondersteunen.

Tot eind 2002 hadden we nog geen initiatieven gevonden die focussen op het uitwisselen van metadata tussen verschillende kennissystemen die duurzame ontwikkeling viseren.

Ze werken dus naast elkaar en delen hun kennis niet. Men kan dat betreuren en samenwerking nastreven. Daarbij zijn er echter twee opties:

- ofwel probeert men een akkoord te bereiken over de metadata-elementen die elke betrokken partij voortaan zou moeten gebruiken, ook binnen zijn lokaal project, waarbij men eventueel de Dublin Core als consensus weerhoudt.
- ofwel aanvaardt men dat elke betrokken partij verder werkt met zijn specifieke metadata-elementen, maar de diensten die men lokaal aanbiedt uitbreidt met een zogenaamde crosswalk naar een gemeenschappelijk metadata-systeem, eventueel de Dublin Core metadata.

De laatste optie is de interessantste omdat ze meest flexibel is en van bestaande projecten een minimale aanpassing vraagt.

Een illustratie:

ISDO heeft o.a. steekkaarten met beschrijvingen van organisaties. De metadata-elementen werden reeds in hoofdstuk 1 vermeld.

We sommen niet alle elementen opnieuw op, maar maken een (alsnog onvolledige) crosswalk van enkele elementen naar de Dublin Core.

ISDO	Dublin Core
Organisatie	DC.Title
Algemene beschrijving	DC.Description
Vrije sleutelwoorden	DC.Subject
Sleutelwoorden (GEMET)	DC.Subject (scheme="gemet")

De metadata die bij ISDO in het element of dataveld 'organisatie' werden opgeslagen, worden in de Dublin Core in het element 'Title' geplaatst. 'Title' is in het Dublin Core Initiatief immers gedefinieerd als "de naam waarmee men de informatiebron aanduidt".

Tijdens de *crosswalk* kan specifieke informatie verloren gaan die in ISDO wordt geëxpliciteerd. Zo wordt het mogelijks voor de eindgebruiker van de Dublin Core minder snel duidelijk dat het hier om metadata gaat die een organisatie beschrijven.

Echter:

- men zou in eerste instantie al kunnen overwegen om het DC.Type element in te vullen met "organisatie".
- men kan ook in DC.Title schrijven: "organisation_name = Center for Sustainable Development ". In dit geval worden dus structured values geproduceerd.

Men moet echter beseffen dat een *crosswalk* wordt gemaakt om de uitwisseling van metadata tussen verschillende kennissystemen te ondersteunen, niet om die respectieve kennissystemen te vervangen. ISDO heeft bepaalde metadata-elementen gekozen op basis van lokale behoeften en ambities die niet noodzakelijk dezelfde zijn als die in andere kennissystemen en evenmin dezelfde zijn als de 'doorsnee'-behoeften waarvoor de Dublin Core werd ontwikkeld.

De Dublin Core is effectief ontwikkeld voor niet-domein specifieke toepassingen en evenmin voor een bepaald type van kennissysteem. De Dublin Core kan en zal dus nooit kunnen tegemoet komen aan domeinspecifieke noden of specifieke eisen die worden gesteld binnen een bepaald kennissysteem.

Informatiesysteem A maakt dus een *crosswalk* naar de Dublin Core en biedt deze aan bij diegenen die een informatiesysteem hebben dat met de Dublin Core werkt, of diegenen die het Dublin Core metadata willen opnemen in hun kennissysteem (dus mogelijk een tweede *crosswalk* uitvoeren: van Dublin Core naar hun metadata-elementen).

Punt is dus dat men niet eindeloos moet overleggen om verschillende systemen nauwer bij elkaar te laten aansluiten, maar dat men afsprekt dat elk een *crosswalk* naar Dublin Core aanbiedt, en ieder doet daar dan mee wat hij/zij wil.

Vermeldenswaardig hierbij is wellicht een ontwikkeling die we bij internationale metadata-projecten inzake GIS (Graphical Information Systems) noteerden. De metadata-elementen die men binnen dit domein nodig acht zijn voor een leek wellicht nogal omslachtig en voor sommige gebruikers misschien te breedvoerig of gedetailleerd. Men opteert daar dan ook voor een *crosswalk* naar de Dublin Core metadata om diensten te leveren naar een breder - niet gespecialiseerd - publiek.

Onzes inziens kan men door middel van *crosswalks* naar de Dublin Core ook een antwoord geven op een confrontatie van standpunten met betrekking tot het beheer van metadata.

Er zijn ingewijden die een uitgesproken voorkeur hebben voor een gedecentraliseerde productie van metadata en anderen die pleiten voor de productie op een centraal punt.

ISDO demonstreert het laatste: als men metadata in het systeem wil invoeren moet men inloggen op de database van ISDO en daar zijn gegevens invullen.

Daartegenover staat de optie dat organisaties of personen hun metadata aanbieden op hun website, op zodanige manier dat deze via het Internet kunnen geogst worden met webrobots die ze dan eventueel opslaat in databases waaraan zoekfaciliteiten zijn gekoppeld. In dit geval kunnen metadata dus ook worden gecentraliseerd, maar ze worden niet op een centraal punt geproduceerd en blijven ook decentraal beschikbaar.

Beide hebben hun voor- en nadelen. In het laatste geval moet de aanbieder van metadata wel weten hoe ze op het WWW moeten worden aangeboden en dat vereist dat minstens één persoon binnen de organisatie bekwaam is in webtechnieken. Ingewikkeld is het niet. Een doorsnee webmaster kan op een half uur leren hoe op een perfecte manier Dublin Core metadata op het WWW moeten worden gepresenteerd. Zelfs een volslagen leek kan snel in de materie worden ingewijd.

Mogelijks is het omslachtig voor een individu of organisatie die snel de metadata van één of een paar informatiebronnen wil aanbieden. Op dat moment zal de voorkeur allicht uitgaan naar een centrale database die niets meer vraagt dan enkele veldjes in de vullen.

Anderzijds is het dan wel zo dat men - met het oog op een zo breed mogelijk publiek - bij verschillende metadatabanken zijn gegevens moet deponeren en steeds opnieuw 'met knippen en plakken' gelijkaardige veldjes moet invullen, tenminste zolang de verschillende informatiesystemen geen metadata uitwisselen.

Wellicht moet men hier de complementariteit van beide methoden zien, een complementariteit die wellicht in de praktijk slechts goed kan werken wanneer men de Dublin Core gebruikt in *crosswalks* (uitwisseling tussen verschillende informatiesystemen) en bij het gedecentraliseerde aanbod. Let wel: in theorie hoeft men niet noodzakelijk de Dublin Core te gebruiken, maar het is wellicht de beste of meest ontwikkelde optie.

In het volgende hoofdstuk ontwikkelen we een procedure die toelaat om zowel domein-specifieke metadata, als DC-metadata aan te geven in één document. Met ander woorden: we ontwikkelen een standaard formaat waarin zowel lokale metadata-behoeften als internationale metadata-standaarden kunnen gecombineerd worden.

Nog enkele belangrijke bronnen:

<http://www.diffuse.org/meta.html> - overzicht van metadata-initiatieven

<http://www.openarchives.org/> - OAI - het Open Archives Initiatief

<http://www.ulb.ac.be/ceese/meta/meta.html>

IV. HTML, XML en XHTML

IV.1 Open Standaarden

Met de opkomst van het World Wide Web zijn internationale standaarden ontwikkeld met betrekking tot de dataformaten die men kan/moet gebruiken om informatie over het Internet uit te wisselen. Vandaag is omzeggens alles mogelijk en vele websites tonen de meest flitsende toepassingen van multimedia (video, geluid,...). Indien men zich echter concentreert op een gecoördineerde uitwisseling van data en metadata tussen verschillende personen en organisaties, dan gaat de focus naar die formaten waarmee de uitwisseling van tekst wordt gestandaardiseerd.

Voor tekst bestaan talloze formaten, zoals Word bijvoorbeeld, maar vele van deze formaten zijn eigenlijk 'eigendom' van een commercieel bedrijf, zoals Word aan Microsoft toebehoort. Men kan weliswaar gratis een 'viewer' downloaden, maar als men zelf zo'n document wil produceren dan moet men meestal software kopen bij het bedrijf dat ene licentie heeft op het formaat.

Daarnaast bestaan er ook formaten die aan niemand toebehoren, of beter: aan iedereen. Die zijn voornamelijk ontstaan met de introductie van het Internet en het World Wide Web waarbij er behoefte was aan standaarden die niet gebonden waren aan een specifiek platform (Windows, Mac, Unix, Linux,...) of een specifiek Office-pakket zoals dat van Microsoft.

Deze standaarden zijn vertrouwde instrumenten voor alle webmasters die de zekerheid willen dat alle browsers (Explorer, Netscape, Opera,..) deze perfect zullen interpreteren.

De ontwikkeling van deze WWW-standaarden is in handen van het W3-consortium, een instelling die weliswaar samenwerk met deskundigen van de grote softwarehuizen als Microsoft en Sun, maar er tegelijk over waakt dat iedereen deze standaarden kan gebruiken zonder de minste vergoeding aan diegenen die betrokken waren bij de ontwikkeling daarvan. Het W3C levert schitterend werk en de resultaten komen uit een intensieve samenwerking van onafhankelijke academici en deskundigen van commerciële softwarehuizen.

Vandaag catalogeert men deze WWW-standaarden onder de verzamelnaam 'open standaarden'

Dat standaarden nuttig zijn, ervaren we dagelijks in het maatschappelijke verkeer. In Europa bijvoorbeeld betalen we in standaard munteenheden (de Euro) en maken we afspraken in standaard tijdseenheden (seconden, minuten, uren etc.).

Standaarden kunnen 'open' zijn of 'gesloten'. Onder een 'open standaard' verstaan we een standaard die voldoet aan de volgende vijf eisen:

1. De standaarden worden op basis van een open beslissingsprocedure (consensus of meerderheidsbeslissing, etc.) vastgesteld;
2. Het beheer van de standaard ligt bij een not-for-profit organisatie die een volledig vrij toetredingsbeleid kent;
3. De standaarden zijn gepubliceerd en men kan er dus alles over te weten komen;
4. De kosten voor het gebruik van de standaard zijn laag en vormen geen drempel voor toegang tot de standaard. Eventueel aanwezig intellectueel eigendom dat aan een open standaard ten grondslag ligt, wordt royalty-free ter beschikking gesteld;
5. Er zijn geen beperkende voorwaarden over het (her)gebruik van een standaard.

Het gebruik van open standaarden wordt steeds meer gepropageerd in het onderwijs en bij overheidsinstellingen. Het gebruik van deze standaarden kan immers voordelen opleveren op het gebied van:

- **Toegankelijkheid:** de mogelijkheid om de informatie aan en zo breed mogelijk publiek beschikbaar te stellen;
- **Aanpasbaarheid:** de mogelijkheid om de informatie snel aan te passen aan nieuwe noden;
- **Duurzaamheid:** de zekerheid dat de informatie bruikbaar blijft ongeacht de verdere ontwikkeling op het vlak van informatie en communicatietechnologie;
- **Interoperabiliteit** en

- **Herbruikbaarheid:** de mogelijkheid om informatie uit te wisselen tussen verschillende platformen (Window, Linux,...) en softwaretools (tekstverwerkers, web servers,...)

Voorbeelden van open standaarden zijn de technische standaarden van het Internet (de zgn. RFC's die o.a. standaard protocollen beschrijven voor de transmissie van gegevens, bijvoorbeeld het http- en ftp-protocol – zie: <http://www.faqs.org/rfcs/index.html>) en de web-standaarden van het World Wide Web Consortium (W3C). Dit zijn echte open standaarden in die zin dat niet één organisatie eigenaar is van de specificatie. Er zijn ook standaarden die veel gebruikt worden maar daarom nog niet open zijn omdat ze eigendom zijn van één organisatie.

Een open standaard is bijvoorbeeld de eXtensible Markup Language (XML) die we verder in dit hoofdstuk toelichten. XML wordt beheerd door W3C. Deze non-profit organisatie beheert samen met IETF (Internet Engineering Task Force) de meeste internetstandaarden. Op de website <http://www.w3c.org> zijn de specificaties van XML gratis verkrijgbaar. Er zijn fora waar over een eventuele nieuwe versie van XML gediscussieerd kan worden. Aan het gebruik van XML worden geen beperkingen gesteld.

Het gebruik van standaarden bevordert de duurzaamheid van de geproduceerde data, omdat men dezelfde gegevens kan gebruiken met verschillende versies van software en besturingssystemen. Het is namelijk erg gemakkelijk als iets wat op een Apple Macintosh is gemaakt, ook op een Linux computer te bekijken is. Interoperabiliteit en herbruikbaarheid tussen verschillende tools en platforms kan mogelijk gemaakt worden door bijvoorbeeld gebruik te maken van een standaard zoals XML (Extensible Markup Language)

Het is immers zo dat alle ontwikkelaars van software of programmeertalen zich vandaag aanpassen aan het bestaan van open standaarden. Sommige lezers zijn misschien nog niet vertrouwd met deze tendens. Men mag overigens stellen dat pas rond het jaar 2000 echt een exponentiële groei van het gebruik van open standaarden plaats vond. De vijf jaren die daaraan vooraf gingen waren de open standaarden enkel nog maar een zaak voor professionals in ICT en databeheer en het is nog maar zeer recent dat deze standaarden brede toepassingsgebieden veroveren waarbij vooral het toenemende gebruik bij de overheid en in het onderwijs mag beklemtoond worden.

Open standaarden en metadata

Onderstaande tabel geeft een overzicht van toonaangevende organisaties die zich bezig houden met standaarden in het onderwijs:

Bron: <http://www.cetis.ac.uk/static/who-does-what.html>

	IMS	ADL/ SCORM	CEN/ ISSS	Dublin Core	Prome- teus	IEEE Learning Tech.	BSI/ ISO	ARIAD- NE	AICC
Metadata	Ja	-	Ja	Ja	-	Ja	Ja	Ja	-
Content	Ja	Ja	-	-	-	Ja	-	-	Ja
Enterprise	Ja	-	-	-	-	Ja	-	-	-
Learner Information	Ja	-	-	-	-	Ja	Ja	-	-
Question & Test	Ja	-	-	-	-	-	Ja	-	-
Accessibility	Ja	-	-	-	-	-	-	-	-
Learning Design	Ja	-	-	-	Ja	-	-	-	-
Collaboration	-	-	-	-	-	-	Ja	-	-
User requirements	Ja	-	-	-	Ja	-	-	-	-

Merk op dat organisaties die open standaarden gebruiken ook veel aandacht schenken aan de standaardisering van metadata. Beide gaan immers samen bij het bevorderen van de toegankelijkheid van informatie.

De belangrijkste organisaties die zich bezighouden met specificaties voor e-learning zijn:

- AICC – Aviation Industry CBT Committee – <http://www.aicc.org>

- ADL (SCORM) (Verenigde Staten, nu internationaal) – Advanced Distributed Learning – <http://www.adlnet.org>
- IMS – voluit IMS Global Learning Consortium – <http://www.imsproject.org>
- IEEE – voluit IEEE Learning Technology Standards Committee – <http://ltsc.ieee.org>
- ARIADNE (Europa) - <http://www.ariadne-eu.org/>
- CANCORE (Canada) - <http://www.cancore.ca/>

Beide laatste hebben banden met IMS of de IEEE (en er is een overlap tussen IMS en IEEE)

Tijdens dit onderzoek is geen grondige analyse gemaakt van de metadata-initiatieven in de onderwijssector. Er waren teveel vernieuwende initiatieven gaande en er ontstonden voortdurend nieuwe samenwerkingsverbanden. Een 'stabiel' plaatje kon moeilijk gemaakt worden. In elk geval moet vermeld worden dat er een samenwerking bestaat tussen het Dublin Core Metadata Initiatief (DCMI) enerzijds en IMS en IEEE anderzijds (het akkoord staat bekend als het Ottawa Communiqué). Men kan overigens vrij gemakkelijk van het ene metadatasysteem naar het andere overschakelen.

Zie: http://www.chin.gc.ca/English/Standards/metadata_educational.html

Open source

Open standaarden worden meestal gepropageerd samen met open source ('vrije software'). Open source houdt in dat de broncode (source code) van een programma beschikbaar is voor iedereen (als men de broncode van een programma kan lezen, dan weet men precies hoe het programma werkt. Van Microsoft krijg je meestal alleen de gecompileerde programma-code, praktisch 'onleesbare' bits en bytes of machinetaal).

Met open source code kunnen heel veel mensen hun eigen inbreng hebben in projecten waar software gedeeld wordt, waardoor de ontwikkeling van het project erg snel vooruit kan gaan. Meer specifiek staat open source eigenlijk voor zowel open software als open data formaten (open standaarden). Dit laatste is erg belangrijk aangezien software eigenlijk maar een hulpmiddel is om met gegevens om te kunnen gaan. Door open standaarden wordt interoperabiliteit met andere programma's mogelijk. In hoofdstuk VI wordt 'open source' verder toegelicht.

Officiële bronnen

De open standaarden die in dit rapport gehanteerd worden zijn: XML, CSS, WAI en HTTP. In volgende secties worden leken ingewijd in de materie, maar hier melden we voor de volledigheid de officiële bronnen die men moet raadplegen als men de standaard feilloos wil gebruiken.

acroniem: HTML naam: HyperText Markup Language versie: 4.01 organisatie: W3C organisatie website: http://www.w3.org/ doel: Representatie van webpagina's documentatie http://www.w3.org/TR/html401/
acroniem: XML naam: Extensible Markup Language versie: 1.0 organisatie: W3C organisatie website: http://www.w3.org/ doel: standaard format voor het uitwisselen van data over het internet documentatie: http://www.w3.org/TR/1998/REC-xml-19980210
acroniem: XHTML naam: Extensible HyperText Markup Language versie: 1.0 organisatie: W3C organisatie website: http://www.w3.org doel: Representatie van webpaginas documentatie http://www.w3.org/TR/xhtml1/

<p>acroniem: CCS naam: Cascading Style Sheets organisatie: W3C organisatie website http://www.w3.org doel: stijl van webpagina's coderen documentatie http://www.w3.org/Style/CSS/</p>
<p>acroniem: HTTP naam: Hypertext Transfer Protocol nummer: RFC 1945 en RFC 2616 versie: 1.0/1.1 organisatie: W3C organisatie website: http://www.w3.org doel: transfer van data over het internet documentatie: http://www.w3.org/Protocols/</p>
<p>acroniem: MIME naam: Multipurpose Internet Mail Extensions nummer: RFC 2045 to RFC 2049 organisatie: IETF organisatie website: http://www.ietf.org doel: e-mail client of webbrowser helpen bij het identificeren van bestandstype documentatie http://www.ietf.org/rfc/rfc2045.txt</p>
<p>acroniem: ECMA-262 versie: 3e editie organisatie: ECMA organisatie website: http://www.ecma-international.org/publications/standards/ECMA-262.HTM doel: scripting language, voornamelijk client side gebruik en correct geïnterpreteerd door meeste browsers omschrijving: De ECMA-standaard voorziet in een goedgekeurde subset van JavaScript</p>
<p>acroniem: UTF-8 naam: Universal Character Set (UCS) Transformation Format - 8 versie 3.0 organisatie: unicode organisatie website: http://www.unicode.org doel karakterset documentatie: ftp://ftp.rfc-editor.org/in-notes/rfc2279.txt omschrijving: Komt overeen met ISO/IEC 10646 part three. Karakterset die vooral op het web gebruikt worden.</p>
<p>acroniem: UTF-16 naam: Universal Character Set (UCS) Transformation Format - 16 versie: 3.1 organisatie: unicode organisatie website: http://www.unicode.org doel: karakterset documentatie ftp://ftp.rfc-editor.org/in-notes/rfc2781.txt</p>
<p>acroniem : WCAG naam: Web Content Accessibility Guidelines versie 1.0 organisatie: W3C organisatie website: http://www.w3.org documentatie: http://www.w3.org/TR/WCAG10</p>

Voor verder ontwikkelingen zie onze website: <http://cdonet.UGent.be/issd/> of <http://purl.oclc.org/NET/cdo>

We laten ons voorlopig inspireren door de nationale overheden die terzake toonaangevend zijn:

Nederland: <http://www.ososs.nl/matrix/matrix.jsp?id=545&f> =aanbevolen

Canada: http://www.cio-dpi.gc.ca/its-nit/index_e.asp

Duitsland: <http://www.kbst.bund.de/Themen-und-Projekte/eGovernment-182/SAGA.htm>

Belangrijke standaardisatie organisaties zijn in elk geval:

IETF - <http://www.ietf.org/>

OASIS - <http://www.oasis-open.org/home/index.php>
W3C - <http://www.w3.org/>

Enkele andere organisaties die het gebruik van open standaarden aanmoedigen

<http://www.iso.org/iso/en/commcentre/wsd/2003wsdindex.html> - Website van de Worlds Standards Day 2003

<http://www.openstandaarden.be/>

<http://www.ososs.nl/index.jsp> - programma Open Standaarden en Open Source Software voor de Nederlandse overheid

<http://openstandards.net/index.jsp> - nieuwsbrief

<http://www.opensource.org/> - Open Source Initiative

<http://www.osdn.com/about/> - Open Source Development Network

IV.2 HTML: HyperText Markup Language

De meest bekende WWW-standaard is HTML: HyperText Markup Language, een code om webpagina's te structureren en met elkaar te verbinden via links.

Een webpagina heeft steeds een inhoud en een structuur. De inhoud is de tekst zelf, datgene wat je leest. De structuur, dat is: de opdeling van de tekst in paragrafen, of een onderscheid tussen titels en gewone tekst, of een tabel, of een lijstje van genummerde items, ...

Als men een tekst beschikbaar wil maken op het World Wide Web en je wenst dat een browser die tekst op een vlot leesbare manier op het scherm brengt van diegene die uw website bezoekt, dan moet je de inhoud -de tekst- zelf vervolledigen met codes die aan de browser duidelijk maken: hier begint een nieuwe paragraaf, hier begint een tabel, hier begint de eerste rij van de tabel...hier eindigt de tabel, hier eindigt het document. Een eindgebruiker ziet die codes niet, maar ziet dus wel een gewenste opdeling in paragrafen, een tabel, een genummerd lijstje, enzovoort.

Stel bijvoorbeeld dat je een tekst op het scherm van een browser wil laten verschijnen zoals in onderstaand kader:

Dit is een titel

Dit is een eerste korte paragraaf.

Dit is nog een paragraaf die nu gevolgd wordt door een lijstje met drie items:

- eerste item
- tweede item
- derde item

De tekst met structurele HTML-code ziet er dan als volgt uit:

```
<html>
<head></head>
<body>
<h1>Dit is een titel</h1>
<p>Dit is een eerste korte paragraaf</p> <p>Dit is nog een paragraaf die nu
gevolgd wordt door een lijstje met drie items:</p>

<ul>
<li>eerste item</li>
<li>tweede item</li><li>tweede item</li>
</ul></body></html>
```

Dit is dus een HTML-document. Met de zogenaamde 'tags' <html> en </html> signaleren we aan de browser waar het document begint en eindigt. Het verschil tussen een begin- en eindtag zit dus in het voorwaartse 'slash'. Zo wordt met <h1> en </h1> aangegeven waar een titel (header) begint en eindigt en met <p> en </p> aangegeven waar een paragraaf begint en eindigt.

 en geven begin en einde van het ongenummerde lijstje aan (unordered list) en en doen dat voor de items in het lijstje.

<body> en </body> bakenen datgene af wat op het scherm van de browser moet verschijnen. In het 'head'-element wordt meestal andere code geplaatst zoals de code waarmee wordt aangegeven welke lettertypes of kleuren moeten worden gebruikt voor achtergrond, titel, paragraaf, etcetera. Deze code is dan weer een andere standaard, met name CSS (Cascading Style Sheets, ook van het W3-Consortium).

De aandachtige lezer zal gemerkt hebben dat de tweede paragraaf in de broncode niet op een nieuwe lijn begint en het derde item in het lijstje evenmin. De browser zal de tekst echter wel op een nieuwe lijn laten starten omdat dit 'gecommandeerd' wordt door de tags.

Een HTML-**element** is een stukje tekst dat afgebakend is met een begin- en eind-tag. Alle witruimte tussen twee opeenvolgende elementen wordt door een browser genegeerd. We hadden dus tussen de twee p-elementen (de twee paragrafen) evengoed een hele pagina witruimte kunnen plaatsen, maar voor de browser blijven het twee paragrafen waarbij de tweede op een volgende regel, direct na de eerste, moet worden geplaatst.

We proberen hier enkel het algemeen concept aan een leek uit te leggen. De finesses -bijvoorbeeld het maken van een link, of het invoegen van een afbeelding - worden in talloze handboeken voor beginners uiteengezet, en ook op het WWW vindt je snel gratis inleidingen als je bijvoorbeeld zoekt met het trefwoord 'HTML-tutorial' of 'inleiding HTML' e.d.

Een snelle intro wordt ondermeer gegeven bij het w3-consortium:

zie: <http://www.w3.org/MarkUp/Guide/>

Er is ook een korte introductie inzake Cascading Style Sheets, de beste methode om aan een webpagina stijl-codes mee te geven.

<http://www.w3.org/MarkUp/Guide/Style.html>

De css-stijlcode wordt in de head van een html-document geplaatst. Bijvoorbeeld:

```
h1, h2 {color: red; font-family:verdana}
```

Deze code 'meldt' aan de browser dat de inhoud van alle h1- en h2-elementen in het rood en met het lettertype 'verdana' moet weergegeven worden.

Met CSS kan je alles onder controle krijgen: de lettergrootte, vet en cursief, de marges aan linker- en rechterzijde van het scherm, achtergrondkleur, afstand tussen de regels,...enzovoort

Je hoeft meestal slechts éénmaal css-code te schrijven die dan voor een eindeloos aantal webpagina's kan worden gebruikt. Men slaat dan de css-code op in een document, bijvoorbeeld met de bestandsnaam: style.css

Vervolgens wordt in alle HTML-documenten die deze stijl moeten gebruiken een link naar die css-code gelegd. (Die link staat dan ook in het head-element)

```
<link type="text/css" rel="stylesheet" href="style.css" />
```

Als men in een document bepaalde elementen een afwijkende stijl wil geven, dan kan men in de html-elementen het attribuut class gebruiken om de elementen te onderscheiden.

Bijvoorbeeld:

```
<h1 class="chapter">Titel van een hoofdstuk</h1>
```

```
<h1 class="section">Titel van een sectie</h1>
```

Met css kunnen we verschillende stijlen toekennen aan beide h1-elementen

```
H1.chapter {color:red;}
```

```
H1.section {color: silver;}
```

Hadden we verder nog volgende code opgegeven

```
H1 {color: black;}
```

...dan zullen alle h1-elementen in het zwart verschijnen, tenzij het een element is met class-waarde chapter of section.

Een ruim aanbod van handleidingen (tutorials) vindt men bij: <http://www.w3schools.com/>

Men heeft er wel baat bij om deze materie in het Engels te bestuderen, omdat beginners vaak wel in meerdere talen worden bediend, maar gevorderden niet. Het is dan jammer als men het Engelse vakjargon niet kent als men zich verder wil verdiepen in de materie.

Een html-document maak je bij aanvang best met een eenvoudige text-editor, zoals kladblok in de bureau-accessoires van MS-office. Bewaar het html-bestand met de extensie .htm of .html (bijvoorbeeld: test.htm) en als je daar op klikt zou (in normale omstandigheden), automatisch uw browser moeten openen en de tekst tonen. Ofwel start je eerst de browser en gebruik je de procedure 'open' in het File menu van uw browser.

Belangrijk is de term 'element' die we verder nog vaker zullen gebruiken: een stuk informatie met begin- en eindtag.

Merk op dat HTML-documenten een boomstructuur hebben. In ons voorbeeld hebben we het html-element als basis (stam van de boom) die vertakt in twee grote delen, head en body. De head is alsnog niet verder vertakt, maar de body heeft drie takken: twee p-elementen (paragrafen) en een ul-element (lijst). De lijst wordt verder vertakt in drie li-elementen.

Die boomstructuur is typisch voor deze en andere standaarden die we verder toelichten.

Men spreekt dan ook over element x dat dochters is van element y, of element y dat 'ouder' is van x (li is in ons voorbeeld een dochter (daughter) van ul, body een ouder (parent) van ul en p).

IV.3 XML: eXtensible Markup Language

Het hele World Wide Web draait voornamelijk op uitwisseling van HTML-documenten en toch is het een standaard die nog niet iedereen tevreden kon stellen. De standaard heeft zijn beperkingen als je hoge eisen gaat stellen aan flexibiliteit en mogelijkheden inzake databeheer en uitwisseling van data over het net.

Daarom is er een tweede standaard, XML of eXtensible Markup Language, die de basis werd van meer geavanceerde toepassingen op het Internet. Met HTML toont het veel gelijkenissen: een XML-document is ook opgebouwd uit elementen (stukken tekst afgebakend met begin- en eind-tag, of openende en sluitende tag), en men heeft eveneens steeds een boomstructuur waarin de elementen zich wederzijds verhouden als ouders en dochters.

Wat is het verschil? In HTML liggen de namen van de elementen vast: met ander woorden je moet kiezen uit een beperkte lijst met daarin p, li, ul, body, head, table, Weliswaar een vrij lange lijst, maar wel eindig.

Bij XML kan men de elementen zelf benoemen, met ander woorden men kan zelf een lijstje van element-namen verzinnen die eventueel beter aangepast zijn aan een specifieke behoefte.

Stel dat je binnen organisatie een standaardfiche met adressen wil invoeren. Dan zou je kunnen afspreken dat deze informatie met de XML-standaard wordt bewaard en uitgewisseld via het Internet. Een fiche kan er dan zo uitzien:

```
<?xml version="1.0" ?>
<persoon>
<familienaam>Marx</familienaam>
```

```
<voornaam>Piet</voornaam>
<adres>
<straat>Tinkstraat</straat>
<nummer>60</nummer>
<gemeente>Gent</gemeente>
</adres>
<contact>
<email>marx@hotmail.com</email>
<telefoon></telefoon>
<fax>09/222222</fax>
<fax>09/1213123</fax>
</contact>
</persoon>
```

De namen van de elementen hebben we zelf gekozen, en we hebben zelf beslist dat bijvoorbeeld het element contact drie dochters kan hebben: e-mail, fax en telefoon.

Zo kan men binnen een groep van mensen een afspraak maken over de elementen die men kan gebruiken en de manier waarop deze als dochters en ouders in een documentstructuur moeten worden opgenomen.

Al deze afspraken kunnen worden vastgelegd in een DTD - Definition Type Document - waarmee men de kwaliteit van de aangeboden data kan verifiëren. Met gebruik van een dergelijk DTD kan een software-toepassing bijvoorbeeld beslissen dat alle persoonsbeschrijvingen zonder voornaam-element geweigerd worden ('not valid', zo heet dat in vakjargon). Maar men kan in de DTD ook vastleggen dat een bepaald element optioneel is, dat het eventueel mag ontbreken, zoals men ook kan vastleggen dat een bepaald element meerdere keren gebruikt wordt (*fax* in ons voorbeeld). Merk op: in ons voorbeeld is het element telefoon weliswaar niet ingevuld, maar het is wel gebruikt (het is leeg, maar aanwezig). Kortom alle finesses die men kan bedenken zijn mogelijk en men kan dan ook zeer scherpe afspraken maken over de structuur van documenten en de doorstroming van data daarop controleren. Voor een ontwikkelaar van software is dit een droom van een instrument. Als je data wil verwerken, dan is het essentieel dat de data aangeboden worden in een structuur waarmee de software kan omgaan. Alle software-programma's beginnen vaak met dergelijke stappen: zijn de aangeboden data wel correct gestructureerd – zoniet: verwerp ze want anders volgt verder in het programma allicht een crash?

Als men XML-data gebruikt dan kan een software-ontwikkelaar standaard procedures gebruiken voor de validering van XML-documenten. En alle data die men kan bedenken, zelfs de meest complexe structuren, kunnen in XML-formaat gegoten worden (zelfs grafische afbeeldingen, maar daar gaan we verder niet op in).

XML is nu eenmaal ontwikkeld als flexibele standaard voor data-transfert tussen computers en computerprogramma's en de ontwerpers van deze standaard hebben dan ook op alle mogelijke toepassingen geanticipeerd. Fantastisch, maar

Er is een maar. De vele kwaliteiten van XML kunnen pas gevaloriseerd worden als men kan omgaan met de XML-technologie. En dat heeft ons tijdens het onderzoek meerdere keren parten gespeeld. Zo kan je afspreken dat alle metadata in XML-bestanden worden opgeslagen. Als men die echter op een prettig leesbare manier op het scherm wil brengen dan moet men XSL inschakelen (eXtensible Style Language).

XSL is één van de standaarden in de XML-familie waarmee je het XML-data kan 'verwerken' of 'bewerken' (de familie wordt steeds uitgebreider: Xpath, XSLT, DOM, RDF, ...) . Dergelijke standaarden vragen echter veel grotere inspanningen om ze onder de knie krijgen. Iemand bijbrengen hoe een correct XML-document wordt gemaakt, duurt hooguit een half uurtje. De andere standaarden vragen dagen studeerwerk en oefening als je wat geavanceerde toepassingen beoogt.

Een aantal van die standaarden hebben we de afgelopen jaren geprobeerd. Enerzijds stel je met toenemend enthousiasme vast hoe groot de slagkracht van de XML-technologie wel is, maar anderzijds moet je erkennen dat een doorsnee organisatie die met 'duurzame ontwikkeling' bezig is, moeilijk de expertise in huis kan halen.

XML-fanaten zullen allicht zeggen dat het allemaal nog meevalt en dat er voldoende software ontwikkeld is, waarmee ook een leek zich snel uit de slag kan trekken. Daarbij is dan de vraag, wat hun perceptie van 'leek' is. Tijdens ons onderzoek hebben we herhaaldelijk 'probleempjes' kunnen noteren, die intern snel konden verholpen worden. Vraag was evenwel: kunnen we XML nu aanbevelen aan collega's of een organisatie waar niemand beschikbaar is om 'probleempjes' op te lossen. Immers: wij kunnen terzake geen service garanderen.

De eindgebruikers die wij viseren zijn zelden bezig met informatica. MS-Word, e-mail en de browser zijn hun voornaamste instrumenten. Bovendien hebben ze 'noden' die niet direct met een simpele XML-editor op te vangen zijn (zie verder: XHTML maken met MS-Word). Meer nog: de meerwaarde die XML biedt hebben ze (voorlopig) slechts zelden nodig. Men kan moeilijk voor één bepaald project, of voor het produceren van enkele metadata, een volledige omschakeling naar XML-technologie verantwoorden.

Volgende vraag kwam naar voor: bestaat er geen optie die tegelijk de voordelen van XML beschikbaar maakt en tegelijk beheersbaar is voor een doorsnee organisatie die actief is in het domein 'duurzame ontwikkeling'. Of nog anders gesteld: kan men geen eerste stappen zetten die meerwaarde creëren zonder daarbij hoge eisen moeten te stellen aan de kennis op het gebied van deze technologie. Het eerste antwoord lag in de standaard XHTML.

IV.4 XHTML: Extensible HyperText Markup Language

Omdat XML niet makkelijk implementeerbaar is in een organisatie die geen expert terzake kan consulteren en bovendien slechts zelden de meerwaarde van deze technologie kan valoriseren, hebben we gezocht naar een middenweg die enerzijds voor deze organisatie toch direct nut oplevert en anderzijds een voorbereiding is op toekomstig gebruik van de XML-technologie.

Daarbij is een belangrijke optie om in eerste instantie XHTML-documenten te maken. Dit zijn documenten die voldoen aan de XML-standaard en tegelijk ook HTML-documenten zijn die door iedere browser -ook oudere versies- feilloos op het computerscherm worden gezet en kunnen uitgeprint worden met een gepaste letterstijl en lay-out die het lezen vergemakkelijkt. Men hoeft daarvoor geen XSL te gebruiken. Een Cascading Style Sheet (CSS) volstaat daarvoor. Dit wil zeggen: technieken die een doorsnee webmaster vrij goed onder de knie heeft. CSS is veel gemakkelijker aan te leren dan XSL. Ook oudere browsers verwerken CSS perfect.

Men kan een XML-document altijd lezen, maar met de 'tags' heeft men qua stijl niet de meest gewenste presentatie. In sommige boeken wordt weliswaar gesteld dat men XML ook kan begeleiden met een CSS, maar dat is een gebrekkige optie. Of beter gesteld: het is een optie die pas goed werkt als je de XML-documenten zodanig ontwerpt dat men dan maar beter omschakelt naar XHTML.

Sommige browsers laten het gebruik van css met xml toe, maar men vergist zich als men denkt hiermee een praktische oplossing te hebben.

Stel dat men bijvoorbeeld volgend xml-element heeft:

```
<naam>W. De Jonge </naam>
```

Men wil dit in het rood op het canvas zetten, dan zal dat in sommige browsers lukken met de css code

```
naam {color: rood;}
```

Het punt is echter dat men daarmee niet het woordje 'naam' niet naar wens in beeld krijgt, en dat is vaak wel de bedoeling. Men wil allicht volgende output op het scherm:

Naam: W. De Jonge

XML-presentatie moet men in feite ondersteunen met XSL(T) en niet met CSS, maar zoals we al eerder hebben gesteld: XSLT heb je niet snel onder de knie. Terwijl veel webmasters CSS beheersen, is dat met XSLT veel minder het geval.

Om maar meteen met de deur in huis te vallen, tonen we een voorbeeld van een XHTML-document met metadata om vervolgens de voordelen ten opzichte van andere XML te bespreken. De lezer focust best eerst op een div-element, waarvan in onderstaand voorbeeld een exemplaar vetjes aangeduid is.

De eerste brok code die het vetgedrukte div-element voorafgaat, kan de leek afschrikken, maar eigenlijk hoeft die zich daar niet veel om te bekommeren. Het is een reeks codes die steeds letterlijk gekopieerd kunnen worden (op de inhoud van het element title na). Die codes geven allerlei hints aan browsers en andere software om aan te geven welk type document het gaat en welke standaarden gehanteerd werden.

```

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>Metatada eindrapport voorraadbeheer</title>
<meta http-equiv="Content-type" content="text/html; charset=iso-8859-1" />
<link rel="stylesheet" type="text/css" href="metadata_body1.css" />
</head>

<body>

<div class="element">
    <div class="dc">dc.title</div>
    <div class="local_name">Titel</div>
    <div class="lang">nl</div>
    <div class="scheme">- </div>
    <div class="content">Voorraadbeheer binnen de
milieugebruiksruimte</div>
</div>

<div class="element">
    <div class="dc">dc.creator</div>
    <div class="local_name">Auteurs</div>
    <div class="lang">nl</div>
    <div class="scheme">-</div>
    <div class="content">Walter De Jonge</div>
</div>

<div class="element">
    <div class="dc">dc.description</div>
    <div class="local_name">Beschrijving</div>
    <div class="lang">nl</div>
    <div class="scheme">-</div>
    <div class="content">Eindrapport van beleidsondersteunend onderzoek
naar de operationalisering van voorraadbeheer binnen het gangbare
milieubeleid. Centrale vraag was: hoe kan men werken aan een efficiënte,
effectieve en rechtvaardige allocatie van schaarse natuurlijk hulpbronnen.
Het concept 'milieugebruiksruimte' werd daarbij geëvalueerd als indicator
voor deze verdeel-vraagstukken.</div>
</div>

<div class="element">
    <div class="dc">dc.identifier</div>
    <div class="local_name">WWW-adres</div>
    <div class="lang">-</div>
    <div class="scheme">URI</div>

```

```
<div class="content"><a href="http://cdonet.ugent.be/">http://cdonet.ugent.be/</a></div>
</div>

<div class="element">
  <div class="ac">ac.name</div>
  <div class="local_name">Metadata-auteur</div>
  <div class="lang">nl</div>
  <div class="scheme">-</div>
  <div class="content">wdj</div>
</div>

</body>
</html>
```

Het element `div` is een HTML element dat voor alles en nog wat gebruikt wordt. Het enige wat `div` implementeert is dat een browser de weergave van de inhoud op een nieuwe regel start en ook weergave van de inhoud van een volgend element op een volgend regel start. Men heeft dus hetzelfde effect als het element `p`.

Op het scherm van een browser krijgen we dus volgend beeld:

```
dc.title
Titel
nl
-
Voorraadbeheer binnen de milieugebruiksruimte

dc.creator
Auteurs
Nl
-
Walter De Jonge
```

Een gelijkaardig element is `span`, ook bedoeld voor 'onbepaald' gebruik, met dit verschil dat een browser niet naar een nieuwe regel springt als een `span`-element begint of wordt afgesloten.

```
<div class="element">
  <span class="dc">dc.identifier</span>
  <span class="local_name">WWW-adres</span>
  <span class="lang">-</span>
  <span class="scheme">URI</span>
  <span class="content"><a href="http://cdonet.ugent.be/">http://cdonet.ugent.be/</a></span>
</div>

<div class="element">
  <span class="ac">ac.name</span>
  <span class="local_name">Metadata-auteur</span>
  <span class="lang">nl</span>
  <span class="scheme">-</span>
  <span class="content">wdj</span>
</div>
```

Dit levert volgende output op een scherm:

```
dc.identifier WWW-adres - URI - http://cdonet.UGent.be
ac.name Metadata-auteur nl - wdj
```


We hebben overigens een vermoeden dat de combinatie van div en span beter is voor weergave zonder css, maar in elk geval kan men met css de weergave altijd wijzigen en aan een div-element het gedrag van een span-element geven en omgekeerd.

Met deze structuur vangen we vele vliegen in één klap:

We hebben een structuur waarin we de verschillende gegevensvelden in een steekkaart een eigen naam kunnen geven:

Titel: Voorraadbeheer binnen de milieugebruiksruimte

Auteurs: Walter De Jonge

Beschrijving: Eindrapport van beleidsondersteunend onderzoek naar de operationalisering van voorraadbeheer binnen het gangbare milieubeleid. Centrale vraag was: hoe kan men werken aan een efficiënte, effectieve en rechtvaardige allocatie van schaarse natuurlijk hulpbronnen. Het concept 'milieugebruiksruimte' werd daarbij geëvalueerd als indicator voor deze verdeel-vraagstukken.

WWW_adress: `http://cdonet.UGent.be/`

Metadata-auteur: wdj

Met een css kunnen we ervoor zorgen dan enkel deze informatie in beeld verschijnt.

Merk op dat we van het www-adres een link gemaakt hebben naar het document met URL <http://cdonet.UGent.be/>.

Binnen de div-elementen kunnen we alle kanten uit en paragrafen (p), lijstjes (ul, li), tabellen en andere structurele html code gebruiken. Ook links naar grafische afbeeldingen zijn mogelijk, of naar externe primaire bronnen zoals een spreadsheet-bestand. Inhoudelijk zijn er dus geen beperkingen.

Bovendien kunnen we ook mappen met de Dublin Core en de A-Core. We hebben in ons voorbeeld naast onze eigen benaming van de velden ook de overeenkomstige DC-namen aangegeven. We kunnen overigens mappen naar een eindeloos aantal Metadata-initiatieven.

Voorts is er ruimte voor het duiden van de taal en eventueel voor het duiden van een scheme (die verwijst naar een standaard code of gecontroleerd vocabularium of thesaurus).

Belangrijk is echter dat je dit document direct kan publiceren op het Word Wide Web en dat elke browser de inhoud perfect in beeld zal brengen.

Even belangrijk is echter ook dat we een XML-document hebben dat met velerlei XML-software (XSLT, DOM) kan verwerkt worden.

Enkele eisen bij XHTML

Eerder hebben we gesteld dat XHTML zowel HTML als XML is. Maar het is vooral XML en daarmee wordt komaf gemaakt met een aantal vrijheden die bekend zijn bij HTML.

Een lijstje van spelregels die het verschil maken met HTML:

Goed	Fout	Spelregel
<code><div>...</div><div>...</div></code>	<code><div>...<div></code>	alle elementen moeten een openende en sluitende tag hebben.
<code><div></code>	<code><DIV></code>	de namen van elementen moeten met kleine letters gespeld worden. ('XML is case-sensitive')
<code><title></code>	<code>< title ></code>	geen spaties in tags, behalve in lege elementen (zie verder)
<code>
</code> <code><hr /></code>	<code>
</code> of <code>
</code> <code><hr></code>	lege elementen moeten een forward slash hebben en er moet een spatie zitten tussen de element-naam en die slash. Denk eraan bij img, meta e.d.
<code><div lang="nl"></code>	<code><div lang=nl></code>	attributen moeten met quotes en ze moeten

		een waarde hebben.
DOCTYPE	Doctype	doctype declaration niet vergeten en de term doctype moet in grote letter
<html>	<html xmlns="http://www.w3.org/1999/xhtml">	namespace vermelden

Men kan xhtml-documenten valideren (controleren op correcte syntax) op volgende site:

als je document op uw PC wil testen:

<http://validator.w3.org/file-upload.html>

als je document op website wil testen:

<http://validator.w3.org/detailed.html>

Als je zelf geen encoding in document opgeeft, dan selecteer je best: iso-8859-1 (west-europees)

De belangrijkste referentie is steeds:

<http://www.w3.org/TR/2002/REC-xhtml1-20020801/> - XHTML 1.0 The Extensible HyperText Markup Language (Second Edition); A Reformulation of HTML 4 in XML 1.0; W3C Recommendation 26 January 2000, revised 1 August 2002

Schemes:

Op het gebied van het gebruik van schemes is onze voornaamste referentiebron nu de 'metadata terms' van het Dublin Core Metadata Initiative (DCMI)

zie: <http://dublincore.org/documents/dcmi-terms/>

en meer specifiek inzake 'encoding-schemes':
<http://dublincore.org/documents/dcmi-terms/#H4>

De taalcodes kunnen geraadpleegd worden op:

<http://www.loc.gov/standards/iso639-2/langcodes.html>

Gebruikt men de twee-lettercode dan wordt scheme="RFC3066"

Gebruikt men de drie-lettercode dan wordt eveneens scheme="RFC3066", men gebruikt bijvoorkeur een twee letter code

Enkele voorbeelden:

Taal	RFC3066
Nederlands	nl
Frans	fr
Engels	en
Duits	de
Spaans	es
Italiaans	it
Arabisch	ar
Turks	tr

Tip voor programmeurs: Bij het programmeren (scripts die metadata verwerken): accepteer voor de waarde van scheme ook kleine letters en negeer spaties of koppeltekens, bv. RFC - 3066), aanvaard ook RFC1766, ISO639, ISO 639-1 of ISO-639-2 en beschouw deze alle zoals RFC3066.

Voorbeeld:

```
<div class="element">
  <div class="dc">dc.language</div>
  <div class="local_name">Taal</div>
  <div class="lang">-</div>
  <div class="scheme">RFC3066</div>
  <div class="content">nl</div>
</div>
```

Het meest gebruikelijk schema voor een datum is: W3CDTF, het 'date en time format' van het W3-consortium

Zie: <http://www.w3.org/TR/NOTE-datetime>

Voorbeeld:

1994-11-05 staat voor 'vijf november 1994'

Men schrijft dus eerst jaar (vier cijfers), dan maand (twee cijfers), tenslotte dag (twee cijfers), gescheiden door een tussenschreefje

Bij het programmeren (scripts die metadata verwerken): accepteer voor de waarde van scheme ook kleine letters en negeer spaties of koppeltekens, bv. w3c -dtf)

Voorbeeld:

```
<div class="element">
  <div class="dc">dc.date</div>
  <div class="local_name">Datum</div>
  <div class="lang">-</div>
  <div class="scheme">W3CDTF</div>
  <div class="content">2003-06-30</div>
</div>
```

Format: men kan putten uit een lijst van Internet Media Types (IMT, ook gekend als MIME):

Zie: <http://www.isi.edu/in-notes/iana/assignments/media-types/media-types>

Format	IMT
platte tekst	text/plain
Hypertext Markup Language	text/html
Cascading Style Sheets	text/css
eXtended Markup Language	text/xml
XHTML	text/html; text/xml
zip-bestand	application/zip
Portable document File (Adobe)	application/pdf
Rich Text Format	application/rtf
Microsoft Word	application/msword

Bij het programmeren (scripts die metadata verwerken): accepteer voor de waarde van scheme ook kleine letters en negeer spaties of koppeltekens, bv. imt), aanvaard ook mime of mime-type i.p.v. imt.

Voorbeeld:

```
<div class="element">
<div class="dc">dc.format</div>
<div class="local_name">format</div>
<div class="lang">-</div>
<div class="scheme">IMT</div>
<div class="content">application/msword</div>
</div>
```

Elke update vindt U op de website: <http://cdonet.UGent.be/issd/>
Of <http://purl.oclc.org/NET/cdo>

Ter afronding tonen we de toepassing van onze XHTML-structuur op een steekkaart van ISDO (zie hoofdstuk 1, pagina 19 en 20). We tonen slechts enkele fragmenten die wellicht volstaan om de aanpak te demonstreren. De attribuutwaarde 'issd_organisation' geeft dus aan dat het een label is van het issd-project, meer bepaald de metadata-fiches voor de beschrijving van organisaties. Een mapping naar Dublin Core is toegevoegd.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title></title>
<meta http-equiv="Content-type" content="text/html; charset=iso-8859-1" />
<link rel="stylesheet" type="text/css" href="metadata_body1.css" />
</head>

<body>

<div class="element">
  <span class="dc">dc.title</span>
  <span class="issd_organisation">organisatie</span>
  <span class="lang">nl</span>
  <span class="scheme"> - </span>
  <span class="content" lang="nl">Universiteit Gent - Centrum voor
Duurzame Ontwikkeling</span>
</div>

<div class="element">
  <span class="dc">dc.title</span>
  <span class="issd_organisation">naam_fr</span>
  <span class="lang">fr</span>
  <span class="scheme">-</span>
  <span class="content"> Université de Gand - Centre de Développement
Durable</span>
</div>
...

<div class="element">
  <span class="dc">dc.title.alternative</span>
  <span class="issd_organisation">acroniem</span>
  <span class="lang">nl</span>
  <span class="scheme">-</span>
  <span class="content">CDO</span>
</div>
...
```

```
<div class="element">
  <span class="dc">dc.identifier</span>
  <span class="issd_organisation">URL</span>
  <span class="lang">nl</span>
  <span class="scheme">URI</span>
  <span
    class="content"><a
href="http://cdonet.UGent.be/">http://cdonet.UGent.be/</a></span>
</div>

<div class="element">
  <span class="dc">dc.subject</span>
  <span class="issd_organisation">sleutelwoorden</span>
  <span class="lang">en</span>
  <span class="scheme">GEMET</span>
  <span class="content">energy, sustainability, prevention, labelling,
ecolabel, documentation centre, environmental data, environmental
legislation, Agenda 21, building material, energy resources, social
participation, pedagogy, wastes, association, social movement</span>
</div>
</body>
</html>
```

V. Web Accessibility Initiative (WAI)

V. 1 Toegankelijkheid van informatie op het WWW

De stad Gent heeft op haar website een zijdsprong naar 'brailenet', waar informatie wordt aangeboden aan mensen met een visuele handicap: <http://www.gent.be/brailenet/>

"In de eerste plaats wordt informatie aangeboden in de vorm van persmededelingen. Hiervan is de lay-out individueel aanpasbaar. Het is ook mogelijk om de belangrijkste persmededelingen in audioformaat te beluisteren. Hiervoor is wel een MP3 - speler noodzakelijk, die vanaf deze website kan gedownload worden. Verder is ook een pagina voorzien met richtlijnen voor webbouwers. Brailenet wil op deze manier een aanzet geven tot het doordacht ontwerpen van websites, die voor iedereen optimaal toegankelijk zijn."

Het is zeker niet de eerste stad of overheidsdienst die er aan denkt om de informatie die ze via het Internet aanbiedt ook toegankelijk te maken voor blinden en slechtzienden. Gent propageert tegelijk ook deze aanpak en verwijst naar bronnen die terzake informatie verschaffen.

Toegankelijkheid betekent 'leesbaarheid' en bij blinden vereist dit dat de inhoud van een webpagina automatisch kan omgezet worden in braille of geluid (de meeste blinden laten een pagina voorlezen door de 'speech'-software op de computer). Zit de webpagina ingewikkeld in elkaar, dan maakt de computer er natuurlijk een zootje van. Ook andere groepen mensen kunnen trouwens moeilijkheden ondervinden wanneer zij geconfronteerd worden met minder goed ontworpen webpagina's:

- mensen die zich niet goed kunnen bewegen (moeilijk een muis kunnen besturen),
- mensen die problemen hebben met het lezen of begrijpen van tekst,
- mensen die niet over de nieuwste internet-technologieën beschikken,
- mensen die werken met oudere besturingssystemen of browsers,
- mensen die een trage internetverbinding hebben

Hoe kan men er nu voor zorgen dat de webpagina's wel toegankelijk zijn voor blinden en slechtzienden (overigens een steeds grotere groep door de vergrijzing)?

Door verschillende organisaties zijn er sets met toegankelijkheidsrichtlijnen ontwikkeld. Met name twee organisaties spelen hierbij een belangrijke rol: het Access-Board die de officiële richtlijnen voor de Verenigde Staten heeft opgesteld en het W3C (WWW- consortium) dat een hoofdrol speelt bij de ontwikkeling van Internet en WWW-standaarden (zie vorig hoofdstuk). De richtlijnen van het Access-Board zijn overigens voor 90 % gebaseerd op de richtlijnen van het W3C. Bij de bespreking van de richtlijnen van het Access-Board wordt stilgestaan bij de minimale verschillen.

Die richtlijnen zijn weliswaar technisch van aard maar vrij vlot verteerbaar voor webmasters die regelmatig webpagina's ontwerpen.

Een volledig overzicht van de W3C richtlijnen is te vinden op de website van het W3C:

<http://www.w3.org/TR/WCAG10/>

Er is tevens een Nederlandse vertaling van de W3C richtlijnen beschikbaar op de website van het Benelux kantoor van het W3C:

<http://www.w3c.nl/Vertalingen/2000/WAI-WEBCONTENT/WAI-WEBCONTENT-NL.html>

"Deze richtlijnen geven aan hoe je Webcontent toegankelijk kunt maken voor mensen met een handicap. Zij zijn bedoeld voor alle Webcontentontwikkelaars (auteurs van webpagina's en ontwerpers van websites) en voor ontwikkelaars van authoring tools. Het belangrijkste doel van deze richtlijnen is de bevordering van de toegankelijkheid. Een neveneffect zal evenwel zijn dat Webcontent meer binnen handbereik komt van alle gebruikers, ongeacht welke user agent zij gebruiken (bijvoorbeeld een desktopbrowser, een voicebrowser, een mobiele telefoon, een PC in de auto, etc.) en ongeacht de beperkingen waaronder zij werken (bijvoorbeeld lawaaierige omgeving, slecht verlichte of felverlichte ruimten, hands-free bellen, etc.). Verder zal het toepassen van deze richtlijnen gebruikers ook helpen om informatie over het Web sneller te vinden. Deze richtlijnen raden

Webcontentontwikkelaars niet af om gebruik te maken van beelden, video etc., maar leggen uit hoe zij multimedia voor een breed publiek toegankelijk kunnen maken."

De officiële richtlijnen van het Access-Board voor Amerikaanse overheidssites:
<http://www.accessibility.nl/checkit/acces-board-web.html>

Naast de Amerikaans en internationaal georiënteerde initiatieven heeft men ook in ons land lokale initiatieven die 'web toegankelijkheid' (web accessibility) promoten.

De organisatie "Blindenzorg Licht en Liefde" onderzoekt websites op hun toegankelijkheid voor blinden en slechtzienden.

Toegankelijke websites krijgen het "Blindsurferlabel". Criteria zijn onder andere:

- het geven van tekstalternatieven bij grafische elementen,
- goede navigatiemogelijkheden,
- eenvoudige en overzichtelijke tabellen,
- vlot invulbare formulieren,
- enzovoort.

Als aan deze voorwaarden voldaan is, kunnen blinden en slechtzienden zonder problemen de informatie op een website raadplegen.

Zie: <http://www.blindsupport.be/> - Het project werd opgericht door zijn geestelijke vader Rudi Canters (inspirator van dit hoofdstuk, maar jammer genoeg overleden op 14 april 2003) en is momenteel een samenwerkingsverband tussen Blindenzorg Licht en Liefde vzw en Oeuvre Nationale des Aveugles asbl (ONA).

In het najaar van 2003 meldde Staatssecretaris Vanvelthoven, verantwoordelijk voor het federaal ICT-beleid:

"Bij het uittekenen van het e-governmentbeleid beschouw ik het als mijn taak ervoor te zorgen dat niemand uitgesloten wordt. Ik heb jammer genoeg gemerkt dat zelfs de federale portaalsite het BlindSurfer-label nog niet heeft. De overheidssites die het label al hebben gekregen, zijn die van de Directie-Generaal Internationale Samenwerking, het ministerie van Sociale Zaken en de Rijksdienst voor Sociale Zekerheid van de provinciale en plaatselijke diensten. Ook de site van voormalig regeringscommissaris Greet Van Gool draagt het label.

In de begroting voor 2004 heeft Fedict 30.000 euro uitgetrokken voor een coördinerende rol. Volgend jaar zullen alle sites worden gescreend en aangepast om het BlindSurfer-label te kunnen krijgen. De federale portaalsite moet het label dit jaar al kunnen krijgen".

De toon is daarmee duidelijk gezet. Het is in feite een tendens die de laatste jaren internationaal een enorme vaart heeft gekregen, mede door 'het Europees jaar van de personen met een handicap' (2003)

Zie: <http://www.eypd2003.org/eypd/index.jsp> (European Year of People with Disabilities of EYPD 2003)

"EYPD2003.org should be accessible for everyone. Therefore, the EYPD2003 team works with a combination of the Web Content Accessibility Guidelines (WCAG) published by the Web Accessibility Initiative (WAI), and the opinion of accessibility experts," zo luidt het op de website van de Europese campagne.

De Belgische officiële websites die dit Europees jaar in eigen land organiseerden waren niet minder begaan met web-toegankelijkheid. Zie o.a:

http://www.vlafo.be/nederlands/ejph/alt_frameset.html

EU-instellingen hebben overigens in het verleden al gehamerd op de toegankelijkheid van informatie voor alle mensen die kunnen/moeten participeren in de kennismaatschappij.

Het Europees Economisch en Sociaal Comité, het Comité van de Regio's en het Europees Parlement hebben respectievelijk op 20 februari 2002, 15 mei 2002 en 13 juni 2002 een gunstig advies uitgebracht over de strategie inzake de toegankelijkheid van websites van de overheid en de inhoud daarvan, die in een mededeling van de Commissie van 2001 was uiteengezet.

Het Europees Parlement heeft de Commissie en de lidstaten verzocht de dialoog met vertegenwoordigers van gehandicapten en ouderen te verbeteren teneinde deze in staat te stellen gebruik te maken van de nieuwe technologieën zodat elke vorm van discriminatie kan worden voorkomen. Ook de Raad heeft gewezen op de noodzaak om een grotere inspanning te doen. In het bijzonder worden de lidstaten aangemoedigd de toegankelijkheid van deze websites en de inhoud ervan door middel van adequate nationale maatregelen te verbeteren. In een resolutie over de toegankelijkheid van Internet (e-accessibility) heeft de Raad een beroep gedaan op de lidstaten en de Commissie om de deelname van gehandicapten aan de kennismaatschappij te verbeteren door zich ertoe te verbinden de technische, juridische en andere obstakels daarvoor uit de weg te ruimen.

Zie o.a: http://europa.eu.int/eur-lex/nl/com/cnc/2001/com2001_0529nl01.pdf en

http://europa.eu.int/information_society/topics/citizens/accessibility/web/wai_2002/ces_opinion_web_wai_2002/a_documents/ces_opinion_web_wai_2002_nl.pdf

Men stelt dus vast dat op allerlei bestuurlijke niveaus – van stad tot EU – en dit wereldwijd, aandacht wordt besteed aan de toegankelijkheid van informatie op het WWW, en dat men daarbij de lat heel hoog legt. Zelfs blinden moeten kunnen meegenieten. Maar nogmaals dient beklemtoond dat naast blinden en slechtzienden, nog veel meer groepen van mensen worden geholpen met een 'overzichtlijke' aanbod van informatie op het Internet.

V.2 Enkele tips voor webmasters¹

Webpagina's ontwerpen die toegankelijk zijn voor personen met visuele of andere handicaps, is helemaal niet zo moeilijk. Het gaat niet om speciale nieuwe technieken of programmeertalen, maar om oog hebben voor enkele details.

Men baseert zich best op de richtlijnen die het World Wide Web consortium (W3C - <http://www.w3.org>) ontwikkelt, want dat wordt internationaal een de facto standaard. Zij hebben daarbij binnen het zogenoemde "Web Accessibility Initiative" (WAI) oog voor personen met visuele of andere handicaps die hen de toegang tot het internet bemoeilijken. Eerder in dit hoofdstuk zijn de links naar de Engelstalige en Nederlandstalige richtlijnen vermeld.

Ook op de website <http://www.accessibility.nl/> vindt u allerlei nuttige informatie over het maken van toegankelijke sites.

Als men zelf websites bouwt, dan bestaan er een aantal instrumenten die kunnen helpen bij het bouwen van toegankelijke webpagina's. De meeste HTML-editors houden tegenwoordig bij het schrijven van webpagina's rekening met toegankelijkheidsaspecten. Bij sommige editors is dit rechtstreeks ingebouwd, bijvoorbeeld:

- Hotmetal - <http://www.hotmetal.com>, en
- Dreamweaver MX - <http://www.macromedia.com/exchange/>

Bij andere kunt u de functionaliteit uitbreiden door een aantal (gratis) extensies te installeren, bijvoorbeeld: FrontPage - <http://www.hisoftware.com/msacc/>.

Daarnaast bestaan diverse instrumenten waarmee u webpagina's op toegankelijkheid kunt testen. Het bekendste daarvan is Bobby Worldwide (<http://www.cast.org/bobby>). Met Bobby kunt u webpagina's gratis online laten testen op toegankelijkheid. Vaak ziet men op toegankelijke websites het label 'Bobby-approved'.

¹ Bron: http://www.vlafo.be/nederlands/ejph/alt_frameset.html (zie toegankelijkheid, onderaan op de webpagina)

Om webpagina's offline te testen, of om een hele website ineens te testen, dient u te beschikken over een eigen versie van Bobby. Terwijl de vorige versie (versie 3.2) van Bobby gratis te downloaden was, dient voor de nieuwe versie (Bobby WorldWide) betaald te worden. Een zeer gebruiksvriendelijk en gratis alternatief voor Bobby is A-prompt:

<http://www.aprompt.ca/> . Waar Bobby enkel aangeeft waar de toegankelijkheidsproblemen zich bevinden en welke dat zijn, biedt A-prompt de mogelijkheid om de gevonden problemen onmiddellijk te corrigeren.

Interessant voor personen met een visuele handicap is de website <http://access.adobe.com> (van, onder andere, de Acrobat Reader). Op deze site vindt u een aantal instrumenten waarmee personen met een visuele handicap pdf-documenten kunnen omzetten naar tekst- of html-documenten. Op internet kan men online een pdf-document laten omzetten naar html.

http://www.adobe.com/products/acrobat/access_simple_form.html . Per e-mail kunt u een pdf-document opsturen, dat dan automatisch teruggestuurd wordt als html- of tekstdocument. De adressen zijn: pdf2html@adobe.com voor html-documenten, pdf2txt@adobe.com voor tekstdocumenten.

Nog enkele bronnen:

<http://www.esat.kuleuven.ac.be/teo/docarch/toegankelijkheid/accweb.htm> - Jan Engelen and Filip Evenepoel (ed.) Producing Web Pages that Everyone can Access - A COST219bis Guidebook

<http://www.accessibility.nl/> - Stichting Accessibility in Nederland
Aanmaak tabellen : <http://www.accessibility.nl/informatie/tabellen.html>

<http://www.htmlhelp.com/nl/design/accessibility/> - Web Design Group over toegankelijkheid

<http://www.w3.org/WAI/> - Homepage The Web Accessibility Initiative
<http://www.w3.org/TR/WAI-WEBCONTENT/> - Web Content Accessibility Guidelines 1.0
<http://www.w3.org/WAI/Policy/> - Policies Relating to Web Accessibility
<http://www.w3.org/WAI/WCAG1-Conformance.html.en> - WAI Conformance Logos

<http://www-3.ibm.com/able/> - Accessibility Center van IBM

<http://www.modemadvies.be> - MODEM Communicatie- en Computercentrum (zie: Werkgroep Ondersteunde Communicatie Vlaanderen (WOCV))

<http://www.e-accessibility.org> - EDeAN - European Design for All e-Accessibility Network

<http://www.d4allnet.gr/> - Design for All Network of Excellence

<http://www.idcnet.info/idcnet/home.xhtml> - Inclusive Design Curriculum Network

V.3 Toegankelijkheid van informatie over (indicatoren voor) duurzame ontwikkeling

... At the national level, each individual shall have appropriate access to information concerning the environment that is held by public authorities..., and the opportunity to participate in decision-making processes. States shall facilitate and encourage public awareness and participation by making information widely available. Effective access to judicial and administrative proceedings, including redress and remedy, shall be provided.
-- Principle 10 of the Rio Declaration on Environment and Development (United Nations)

Als men in het kader van duurzame ontwikkeling 'participatie' van zoveel mogelijk stakeholders ernstig neemt dan is ook de toegankelijkheid van informatie voor personen met een handicap een aspect dat men niet mag negeren, en dus ook de toegankelijkheid van informatie over indicatoren. In vorige hoofdstukken hebben we een basisstructuur voor XHTML-document ontwikkeld voor de beschrijving van indicatoren en/of voor de overdracht van metadata via het internet.

Voldoet deze basisstructuur nu aan de toegankelijkheidsrichtlijnen van WAI?

We testten volgend document:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title></title>
<meta http-equiv="Content-type" content="text/html; charset=iso-8859-1" />
<link rel="stylesheet" type="text/css" href="metadata_body1.css" />
</head>

<body>

<div class="element">
<span class="dc">dc.title</span>
<span class="issd_organisation">organisatie</span>
<span class="lang">nl</span>
<span class="scheme"> - </span>
<span class="content">Universiteit Gent - Centrum voor Duurzame
Ontwikkeling</span>
</div>

<div class="element">
<span class="dc">dc.title</span>
<span class="issd_organisation">naam_fr</span>
<span class="lang">fr</span>
<span class="scheme">-</span>
<span class="content"> Universit  Gand - Centre de D veloppement
Durable</span>
</div>

<div class="element">
<span class="dc">dc.title.alternative</span>
<span class="issd_organisation">acroniem</span>
<span class="lang">nl</span>
<span class="scheme">-</span>
<span class="content">CDO</span>
</div>

<div class="element">
    <span class="dc">dc.identifier</span>
    <span class="issd_organisation">URL</span>
    <span class="lang">nl</span>
    <span class="scheme">URI</span>
    <span class="content"><a
href="http://cdonet.ugent.be/">http://cdonet.ugent.be/</a></span>
</div>

<div class="element">
<span class="dc">dc.subject</span>
<span class="issd_organisation">sleutelwoorden</span>
<span class="lang">en</span>
<span class="scheme">GEMET</span>
<span class="content">energy, sustainability, prevention, labelling,
ecolabel, documentation centre, environmental data, environmental
legislation, Agenda 21, building material, energy resources, social
participation, pedagogy, wastes, association, social movement</span>
</div>
```

```
</div>
</body>
</html>
```

Als we die pagina eerst controleren op XHTML-kwaliteiten dan lijkt ze perfect te zijn. Deze test kan men online verrichten middels een zogenaamde XHTML-validator:

Zie: <http://validator.w3.org/>

Ons XHTML-document leverde volgend commentaar:

"This Page Is Valid XHTML 1.0 Strict!"

"The uploaded file was checked and found to be valid XHTML 1.0 Strict. This means that the resource in question identified itself as "XHTML 1.0 Strict" and that we successfully performed a formal validation using an SGML or XML Parser (depending on the markup language used)."

We mogen deze pagina dan ook een het officiële XHTML-logo geven, zie:
<http://www.w3.org/lcons/valid-xhtml10>

Als we het document nu testen met A-prompt (software die men kan downloaden is te vinden op <http://www.aprompt.ca/>) dan worden volgende bemerkingen gemaakt:

Er is geen formele aanduiding van de taal waarin de inhoud van het document, of delen daarvan zijn opgesteld.

In feite verwacht A-Prompt dat we in de start-tag html aangeven wat de belangrijkste taal is binnen het document. Als er bepaalde stukken tekst van afwijken, dan moeten we dat daar ook aan geven.

De eerste regel wordt dus in ons geval:

```
<html xmlns="http://www.w3.org/1999/xhtml" lang="nl">
```

...waarmee we melden dat de taal voornamelijk Nederlands is.

Het laatste div-element is evenwel deels in het Engels, en dit duiden we als volgt aan:

```
<div class="element">
<span class="dc">dc.subject</span>
<span class="issd_organisation">sleutelwoorden</span>
<span class="lang">en</span>
<span class="scheme">GEMET</span>
<span class="content" lang="en">energy, sustainability, prevention,
labelling, ecolabel, documentation centre, environmental data,
environmental legislation, Agenda 21, building material, energy resources,
social participation, pedagogy, wastes, association, social movement</span>
</div>
```

Het kan op het eerste zicht overbodig lijken, omdat we al in de metadata hadden aangegeven dat de taal hier Engels was, middels de regel:

```
<span class="lang">en</span>
```

Het verschil is evenwel dat het hier inhoud van de metadata betreft, die we ook op het scherm willen zien, terwijl het lang-attribuut bedoelt is voor voice-software voor blinden die moet 'ingelicht' worden inzake de taal die hij/zij moet verwerken. Als men dat niet doet, dan kan die Nederlandse zinnen met een zwaar Amerikaans accent lezen.

Verder heeft A-Prompt ons gewaarschuwd over de link naar cdonet.UGent.be. “Is de link duidelijk?”, zo luidde de vraag. Sommigen beschrijven immers een link met ‘Klik hier’ of ‘Ga door’. A-Prompt beveelt aan om duidelijk te maken waar men terecht komt als men iets aanklikt.

Structureel had A-Prompt evenwel geen problemen met het document. Als taal aanduidingen en links helder genoeg zijn verdient ons document het A-Prompt-logo en/of WAI-lolo.

Het kan omslachtig lijken dat men herhaaldelijk twee keer een taal-aanduiding moet maken in het document. Echter zullen we verder in dit rapport een methode belichten voor de aanmaak van dergelijke documenten, door het gebruik van Word-documenten met verborgen code. In normale omstandigheden kan via die truuk heel veel dubbel werk vermeden worden.

Met Bobby kunt u webpagina's gratis online laten testen op toegankelijkheid. Vaak ziet men op toegankelijke websites het label ‘Bobby-approved’. Het zal allicht niet verbazen dat Bobby vrijwel dezelfde bemerkingen maakt als A-prompt.

De beste test is evenwel een lezing door een blinde persoon. En die kan men steeds vinden bij een blindenorganisatie waar meestal wel mensen zitten die online uw website kunnen ‘bekijken’.

VI . Archivering van metadata

VI.1 Beheer met relationele database

In het standaard XHTML-formulier dat we in hoofdstuk IV ontwikkelden, kan een onbeperkt aantal metadata zitten. Formeel kunnen er nul, één of meer metadata-elementen in één document voorkomen.

Dit lijkt op het eerste zicht een beheer met een relationeel database management systeem (RDBMS) uit te sluiten aangezien men in een dergelijk database in principe records maakt met een gelimiteerd aantal velden. Men stelt effectief vast dat men in metadata-projecten, waarin de metadata worden ingezameld in een RDBMS, records worden ontwikkeld waarin elke record alle metadata bevatten die één specifieke databron beschrijven.

Het feit dat het aantal metadata-elementen kan variëren, alsook de aard van de metadata - enkel Dublin Core, of Dublin Core gecombineerd met andere - is een gegeven dat kan aansporen tot databeheer met behulp van XML-software. In dat geval worden de XHTML-formulieren zelf de records in de database en men zal dan alle zoek-, sorteer- en andere operaties implementeren met gebruik van XML-technieken zoals XSLT en de DOM.

Dat kan relatief gemakkelijk georganiseerd worden als men mag veronderstellen dat de inkomende XHTML-formulieren perfect opgesteld zijn. Met ander woorden als deze documenten well-formed en valid zijn. Het is heel gebruikelijk om bij XML-systemen alle inkomende data te weigeren die niet voldoen aan de regels van een goed samengesteld XML-document. Er wordt dus eerst gekeken of het document voldoet aan de afspraken en indien dat niet het geval is, dan worden geen verdere operaties meer uitgevoerd op die data. Dit is overigens een interessant systeem, aangezien men daarmee vele problemen in het verdere verloop van de verwerking kan vermijden.

Wanneer we echter met een grote groep organisaties afspreken dat er metadata over IDO (Indicatoren voor Duurzame Ontwikkeling) en duurzame ontwikkeling over het algemeen worden aangeboden met behulp van een XHTML-document zoals beschreven in hoofdstuk IV, dan zullen we er moeten vanuit gaan dat er regelmatig wel eens een fout wordt gemaakt. Een optie is dan dat men diegene die informatie aanbieden hiervan op de hoogte stelt en hen de gelegenheid geeft om de fouten te corrigeren, zodat men vervolgens de correcte XHTML-documenten kan aanbieden.

We vrezen echter dat dit een vrij frustrerende methode kan worden.

Een organisatie heeft er op termijn alle baat bij om de XHTML-documenten correct op te stellen (omdat ze die dan zelf met XML-software kunnen bewerken), en er bestaan gratis beschikbare tools om fouten op te sporen in XHTML-documenten (HTML-tidy), evenals websites met validators waar men zijn XHTML-document kan uploaden en een verslag krijgt met de fouten in het document (<http://validator.w3.org>, zie ook hoofdstuk V, pagina 92). Deze argumenten zouden moeten volstaan om het streven naar perfectie te verantwoorden. Maar men moet er ook rekening mee houden dat een krap budget en een gebrek aan tijd eveneens argumenten zijn die bepalen wat men doet of laat. Stel dat men bijvoorbeeld honderd XHTML-documenten heeft waarin regelmatig meerdere kleine foutjes zijn gesloten. Indien het altijd dezelfde fouten zijn, dan is correctie misschien nog snel realiseerbaar. Indien er in de verschillende XHTML-documenten zeer uiteenlopende fouten voorkomen, dan kan het een hele klus worden om die te corrigeren. En de betrokken organisatie kan voldoende redenen hebben om dit uit te stellen. Met het risico dat het er nooit van komt.

Overigens moet men er rekening mee houden dat de foutmeldingen van XML-software vrij cryptisch kunnen zijn voor iemand die niet echt vertrouwd is met HTML- en XML. Een archief dat 'onverstaanbare' boodschappen terugstuurt, zal zeker niet populair worden bij een breder publiek.

Dit alles leidde tot de conclusie dat we een systeem moeten ontwikkelen dat aanbieders van metadata weliswaar waarschuwt voor de fouten die ze maken, maar anderzijds voorlopig de input met gebreken aanvaardt. De aanbieder van metadata krijgt dan volgende boodschap:

"Uw data zijn opgeslagen in archief. Voorlopig kunnen die geconsulteerd worden door derden. Evenwel stellen we vast dat er in de documenten fouten zitten die ze mogelijks ongeschikt maken voor verwerking met XML-software. Voor sommige gebruikers van uw data kan dat een hinderpaal zijn. U kan die fouten corrigeren en de verbeterde documenten later opnieuw aanbieden. Dit archief zal echter voorlopig de aangeboden metadata bewaren en deze verwerken met software die minder

eisen stelt."

Iemand die perfecte XHTML-documenten aanbiedt, krijgt dan volgende boodschap.

"Uw data zijn opgeslagen in archief en kunnen geconsulteerd worden door derden. Overigens stellen we vast dat uw documenten perfect samengesteld zijn en voldoen aan de richtlijnen voor XHTML. Proficiat."

Dit kan natuurlijk slechts meegedeeld worden als het archief alle binnenkomende documenten valideert. Maar het is zeker geen slechte optie om dat te doen. Het is immers een belangrijke service die vele organisaties kan helpen om - weliswaar stap voor stap- om te schakelen naar het gebruik van XML bij data-uitwisseling.

Wanneer nu XHTML-documenten worden aangeboden aan het archief, dan zal dit eerst volgende stappen ondernemen:

Elk document wordt volledig opgeslagen in de database. We ontwikkelen daarvoor een record met volgende velden:

id
url
file
datetime
validation
harvest
management

id is een uniek identificatienummer van de record (key), zodat elke record in de database-tabel kan onderscheiden worden van andere.

url is het webadres van het XHTML-document. We gaan ervan uit dat dit niet kan gebruikt worden als id (key) binnen de database, omdat een organisatie altijd veranderingen kan aanbrengen op zijn website en dat op een foute manier meldt bij het archief.

file bevat dan het hele XHTML-document

valid kan vier waarden bevatten: unknown, valid, well-formed en no:

- Unknow als er nog geen validering is gebeurd;
- valid als na validering blijkt dat het om een perfect (well-formed en valid) XHTML-document gaat
- wellformed als uit de validering blijkt dat well-formed maar niet valid is (volgens XHTML-DTD)
- no als het document niet well-formed is

harvest verwijst naar een ander document dat we verder in dit hoofdstuk toelichten

management is een veld waarin we gegevens opslaan betreffende de verdere verwerking van de binnenkomende gegevens in het archief. Het is immers de bedoeling dat we het hier niet bij laten en dat het binnenkomende document verder verwerkt wordt waarbij de verschillende metadata worden opgeslagen in aparte records.

Regular expressions

In elk geval beschikken we nu over de data die voorwerp kunnen worden van zoekoperaties. Als het een valid document is kunnen de zoekoperaties geïmplementeerd worden met XML-software, indien niet dan zullen we ons moeten redden met regular expressions (pattern matching). Regular expressions zijn een instrument dat beschikbaar is in vrijwel alle programmeer- of scripttalen die ontwikkeld zijn voor tekstverwerking.

Stel we hebben een document met volgend metadata-element:

```
<div class="element">
  <span class="dc">dc.title</span>
  <span class="local_label">Naam van de indicator</span>
  <span class="lang">nl</span>
  <span class="scheme">-</span>
  <span class="content">Gemiddeld energiegebruik van een inwoner van de
stad</span>
</div>
```

Als men bijvoorbeeld zoekt naar een document met een DC.Title die in zijn content het woord 'energie' bevat dan kan men volgende opdracht formuleren:
zoek in de tekst naar een tekst- string die

begint met class="dc"
vervolgens één of meerdere willekeurige tekens bevat,
vervolgens het woord title bevat,
vervolgens één of meerdere willekeurige tekens bevat,
vervolgens de tekens class="content" bevat,
vervolgens één of meerdere willekeurige tekens bevat, met uitzondering van de string '',
vervolgens het woord energie bevat,
vervolgens één of meerdere willekeurige tekens bevat,
en eindigt met

Volgend ECMA-script (javascript) demonstreert deze match-test:

```
<title>testing regular expression</title>
<script language="JavaScript" type="text/javascript">
<!--

// xhtml-document

var original_string = 'x???? ... <class="DC" ...?++++/>...
title...class="content"...?llB....energie...</span>...';

// first test
var reg_exp1 = /class="dc".+title.+class="content".+energie.+<\span>/i;
var result1 = original_string.search(reg_exp1);

if (result1 != -1) {
document.write ("the first pattern was found" + "<br>")
}
else {
document.write ("the first pattern was NOT found" + "<br>")
}
document.write ("index first match = " + result1);
document.write ("<br>");

//second test
var reg_exp2
=/class="dc".+title.+class="content".+<span.+energie.+<\span>/i;
var result2 = original_string.search(reg_exp2);

if (result2 != -1) {
document.write ("the second pattern was found" + "<br>")
}
else {
document.write ("the second pattern was NOT found" + "<br>")
}
document.write ("index second match = " + result2);
document.write ("<br>");
```

```
if (result1 != -1) {  
  if (result1 != result2) {  
    document.write ("You have a hit")  
  }  
  else {  
    document.write ("No hit")  
  }  
} else {  
  document.write ("No hit")  
}  
-->  
</script>  
</head>  
<body></body>  
</html>
```

De variabele `original_string` speelt in dit script de rol van het XHTML-document dat we willen doorzoeken.

Met de huidige waarde van die variabele levert het script volgend resultaat:

```
the first pattern was found  
index first match = 11  
the second pattern was NOT found  
index second match = -1  
You have a hit
```

Eerst wordt gekeken naar volgende match:

```
class="dc"...title...class="content"....energie...</span>
```

Result1 wordt gelijk gesteld aan -1 als er geen match is, of de index van het eerste teken in de string die een match oplevert

Vervolgens wordt gekeken naar volgende match:

```
class="dc"...title...class="content" ...<span>...energie...</span>
```

Result2 wordt gelijk gesteld aan -1 als er geen match is, of de index van het eerste teken in de string die een match oplevert.

Indien `result1` gelijk is aan `result 2` dan hebben we een `` tussen `class="content"` en `energie`, waardoor we de match niet kunnen weerhouden. 'Energie' komt wel voor in XHTML-document maar niet in content van `dc.title`

Het script is misschien nog niet perfect, maar het toont allicht de manier waarop men de zoekoperatie kan aanpakken. Zo zullen we ondermeer moeten voorkomen dat er in de variabele `original_string` newlines aanwezig zijn. Immers: het meta-karakter dot (.) matcht met elk karakter, behalve newline. Eveneens moeten in de variabele `original_string` alle single quotes vervangen door een double quotes, vooraleer er met bovenstaande regular expressions kan gewerkt worden.

Overigens zou men bij een betere applicatie het XHTML-document eerst kunnen splitten in div-elementen en die in een array stoppen om ze vervolgens met regular expression te doorzoeken.

Indien men echter te maken heeft met een well-formed of valid document, dan zal men regular expressions wellicht gebruiken in combinatie met het Document Object Model.

Maar met regular expressions kunnen we de eisen betreffende de kwaliteit van de XHTML-documenten relatief beperkt houden.

Management

Nadat het volledige XHTML-document is opgeslagen in een tabel met records zoals hierboven beschreven, kunnen we overwegen om de gegevens te verdelen over verschillende tabellen die elk een specifiek metadatum bevatten. Zo kunnen we bijvoorbeeld een tabel maken die records bevat met alle DC.title gegevens uit de binnenkomen documenten.

De tabel heet dan DC.title en de velden in een record zijn:

id
content
scheme
lang
other_labels
source

Daarin kunnen dus alle gegevens opgeslagen worden uit een div-element dat een element `dc.title` bevat, dus zowel de content, het scheme dat eventueel gebruikt werd, de taal (lang) en de namen van andere labels die men eventueel, naast DC.Title aan de content gaf (locale labels en eventueel ISSD-labels).

Other_labels bevat steeds koppels: als men lokaal in het XHTML_document schreef:

```
<span class="local_label">naam van de indicator</span>
```

Dan wordt er in het veld other_labels de waarde (local_label, naam van de indicator) opgeslagen.

id is hier de key van het DC.Title record

Source is de id (key) van de record waarin het volledige XHTML-document ligt opgeslagen.

Indien nu iemand zoekt op een trefwoord dat in de titel moet voorkomen dan moet enkel de tabel DC.Title doorzocht worden. Wordt er een titel met dit trefwoord gevonden dan kan de archief-software meerdere dingen doen:

- De volledige titel terugzenden met een link naar het volledige XHTML-document dat in de database opgeslagen ligt
- Daarbij ook de URL meegeven naar de oorspronkelijke locatie (het origineel op het WWW)
- Eerst een beperkte hoeveelheid metadata terugsturen, bijvoorbeeld DC.Title, DC.Subject en DC.Description. We gaan er hier immers vanuit dat er ook voor DC.Subject en DC.Description tabellen worden gemaakt en dat men dus snel in deze tabellen alle records kan opsporen die dezelfde waarde hebben voor source.

De bezoeker van het archief kan dan zelf beslissen of hij/zij meer gegevens wil.

De mogelijkheden zijn eindeloos. In elke geval durven we hier vaststellen dat het niet onverstandig is om het gehele XHTML-document in een aparte tabel te bewaren. Men kan het dan immers gemakkelijk in zijn geheel naar de bezoeker van het archief sturen. Vroeger was het 'not done' om in een database twee maal dezelfde gegevens op te slaan, omdat elke byte geheugen kostbaar was. Vandaag ligt de kostprijs van geheugenruimte een stuk lager en men doet er zelfs beter aan vooral de loonkosten van databeheer zo laag mogelijk te houden.

Indien we zouden overwegen om de data die opgeslagen zijn in specifieke metadata-tabellen te verwijderen uit ons XHTML-document, dan moeten we ook zeker zijn dat we het document kunnen reconstrueren. Indien we het niet kunnen reconstrueren, dan rest ons enkel de link naar het origineel (de website van de aanbieder van het XHTML-document). Is die website verdwenen of herschikt zodanig dat de link niet meer klopt, dan hebben we een verlies aan gegevens. Meer nog: het archief kan dan niet langer gebruikt worden als een soort back-up van bestaande projecten, wat nochtans een interessante service is.

Het enige wat we hier durven overwegen is het volgende: als er een div-element volledig opgeslagen is in een tabel van een specifiek metadatum, dan kunnen we in het XHTML-document dit div-element vervangen door een kort PHP-script dat dit div-element reconstrueert met de gegevens uit de metadata-tabellen, wanneer het XHTML-document naar een bezoeker van het archief wordt gestuurd. Dat kan op langere termijn een noemenswaardige besparing aan geheugenruimte opleveren. Maar het is tegelijk een methode die oorzaak kan zijn van malfunctions van het systeem. En een probleem oplossen kan veel duurder zijn dan geheugenruimte.

```
<html>
<head></head>
<body>

<?php
script dat metadata uit DC.title tabel haalt
?>

<div>
originele gegevens die (nog) niet in een specifieke tabel werden opgeslagen
</div>

<?php
ander script dat metadata uit bijvoorbeeld DC.Subject haalt
?>

etcetera

</body>
</html>
```

Ook bij deze laatste optie beschikken we dus over verschillende metadata-tabellen waarvan alle records gerelateerd zijn aan hun source-id en wat meer is, de XHTML-documenten met PHP scripts bewaren de originele volgorde van de metadata in het origineel. Een perfecte reconstructie van het origineel blijft dus mogelijk.

Geheugenbesparing is ook mogelijk door het verwijderen van CSS uit head van aangeboden XHTML-documenten. Het archief moet slechts één CSS hebben dat toepasbaar is op alle XHTML-documenten.

VI.2 Harvest: oogsten van metadata

We belichten nu de procedure die men onzes inziens best zou kunnen gebruiken om de metadata aan te bieden bij het archief. Aan de toeleverancier van metadata (XHTML-documenten) wordt gevraagd een XML-document te maken waarin alle URL's (webadressen) worden opgesomd van de XHTML-documenten die het archief moet opslaan.

Indien er slechts één XHTML-document wordt aangeboden dan is het XML-document uiterst simpel.

```
<?xml version="1.0" ?>
<harvest>

<contributor>
<username>wdejonge</username>
<password>x?kkk!!123ml</password>
<name>W. De Jonge</name>
<email>walter.dejonge@UGent.be</email>
</contributor>

<folder>
<source
type="body">http://cdonet.UGent.be/project_x/indicator_one.htm</source>
</folder>

</harvest>
```

Het XML-document bestaat dus uit een root-element 'harvest' (we 'oogsten' metadata) met twee grote

delen: contributor en folder. Contributor bevat een aantal gegevens over de aanbieder van de metadata die eventueel gecontacteerd kan worden als er iets misloopt. Naam en email zijn dus wel het minimum, maar we voorzien verder in het gebruik van nog meer contributor-gegevens. Merk op dat we ook een username en een paswoord megeven via dit xml-document. Dit document bewaart men dus best op een veilige plek.

Folder is het element (een map) waarin we de bronnen opsommen waaruit het archief metadata moet oogsten. In bovenstaand voorbeeld is dat slechts één bron, in het bijzonder het xhtml-document met url of web-adres http://cdonet.UGent.be/project_x/indicator_one.htm

In het attribuut **type** van het source-element staat hier body, wat betekent dat het archief metadata moet oogstens uit het body-element van het xhtml-document.

Indien de waarde van het type-attribuut 'head' is, dan worden metadata geogst uit de <meta ...>-elementen in de head van het xhtml-document.

Is de waarde 'both', dan wordt er zowel uit de head als uit de body geogst.

Men kan dus ook een source opgeven die geen xhtml-document is zoals beschreven in hoofdstuk IV, maar wel relevante metadata bevat, bijvoorbeeld van de homepage van de website van de organisatie. Verder wordt duidelijk waarom we dat mogelijk maken.

In elk geval wordt hier dus gesteld dat het archief ook metadata uit de head van een gewoon html-document zal kunnen oogsten, ook uit documenten die niet gestructureerd zijn volgens het standaard xhtml-document.

Wel is het zo dat een document, dat met attribuut-waarde 'head' wordt aangeboden, niet volledig in de database van het archief zal worden opgeslagen.

Stel dat we volgende source vermelden:

```
<source type="head">http://cdonet.rug.be</source>
```

..dan wordt in het archief de inhoud van het document opgeslagen maar alle tekst binnen het element <body> wordt vervangen door

```
<a href="http://cdonet.rug.ac.be">http://cdonet.rug.be</href>
```

Dus enkel de URL van het origineel blijft over, naast de metadata-inhoud van head (waaruit overigens ook CSS en scripts kunnen worden geschrapt aangezien we enkel geïnteresseerd zijn in de metadata).

Indien men wenst dat de body bewaart wordt, dan moet men het attribuut de waarde body of both geven.

Een volledig contributor-element ziet er als bijvoorbeeld als volgt uit

```
<contributor>
<username>wdejonge</username>
<password>x?kkk!!123ml</password>
<name>W. De Jonge</name>
<name>S. Deconinck</name>
<email>walter.dejonge@UGent.be</email>
<email>stefan.deconinck@UGent.be</email>
<website>http://cdonet.UGent.be</website>
<tel>+ 32 (0)9 264 82 10</tel>
<fax>+ 32 (0)9 264 83 90</fax>
<date>2003-12-31</date>
</contributor>
```

Hoewel er slechts één username en slechts één paswoord kan vermeld worden, kunnen er van de elementen name, email, website, tel en fax nul, één of meerdere voorkomen. Date is de datum waarop deze harvest-file laatst werd bijgewerkt (niet noodzakelijk de datum waarop men deze naar het archief stuurt, want die datum wordt automatisch door het archief geregistreerd). Ook het element date kan meerdere keren voorkomen, zodat de contributor een lijst kan bijhouden van datums waarop deze harvest file werd bijgewerkt.

In elk geval wordt de toeleverancier (contributor) uitgenodigd om voldoende namen en contactgegevens te vermelden zodat men -bij afwezigheid van de ene persoon- desnoods iemand anders kan bereiken.

We focussen nu verder op het folder-element.

Dit element kan source-elementen bevatten, meta-elementen of folder-elementen

Volgend voorbeeld moet de aanpak verduidelijken. Een folder met een sub-folder::

```
<folder>

<meta name="dc.title" content="metadata van enkele indicatoren" />
<meta name="dc.subject" content="sustainable development, social" />

<folder>
<source
type="body">http://cdonet.UGent.be/project_x/indicator_1.htm</source>
<source
type="body">http://cdonet.UGent.be/project_x/indicator_2.htm</source>
<source
type="body">http://cdonet.UGent.be/project_x/indicator_3.htm</source>
</folder>

</folder>
```

We kunnen dus aan de folder een aantal metagegevens toevoegen die de inhoud van de folder beschrijven.

Merk op dat we de metadata van de indicatoren in een subfolder stoppen omdat we wensen dat ze in niveau dieper in de hiërarchie van metadata zitten.

- dc.title: metadata van enkele indicatoren
- dc.subject: sustainable development, social

- metadata van indicator_1
- metadata van indicator_2
- metadata van indicator_3

Stel nu dat we twee projecten lopen hebben, dan kan volgende harvest-file ontstaan:

```
<?xml version="1.0" ?>
<harvest>
<contributor>...</contributor>

<folder>

<folder>
<meta name="issd.resource" content="organisation" />
<source type="head">http://cdonet.UGent.be</source>
</folder>

<folder>
<meta name="issd.resource" content="indicatorset" />
<meta name="dc.title" content="IDO Project X" />
<folder>
<source
type="body">http://cdonet.UGent.be/project_x/indicator_1.htm</source>
<source
type="body">http://cdonet.UGent.be/project_x/indicator_2.htm</source>
<source
type="body">http://cdonet.UGent.be/project_x/indicator_3.htm</source>
```

```

</folder>
</folder>

<folder>
<meta name="issd.resource" content="indicatorset" />
<meta name="dc.title" content="IDO Project Y" />
<folder>
<source
type="body">http://cdonet.UGent.be/project_y/indicator_1.htm</source>
<source                                type="body">http://cdonet.UGent.be
/project_y/indicator_2.htm</source>
<source                                type="body">http://cdonet.UGent.be
/project_y/indicator_3.htm</source>
</folder>
</folder>

</folder>
</harvest>

```

We hebben nu een folder met drie deel-folders.

In de eerste folder stoppen we in feite de metadata van de homepage van onze website. De metadata die daar in de head zitten, zijn een Dublin Core beschrijving van onze organisatie.

De twee andere folders bevatten een collectie metadata van indicatoren, telkens ingeleid met metadata die de verzameling indicatoren kunnen toelichten.

Indien we echter al een document hebben waarin metadata over de collectie van indicatoren staan dan kunnen we ook daaruit oogsten:

```

<folder>
<source type="body">
http://cdonet.UGent.be/project_y/introduction.htm</source>

<folder>
<source type="body" resource="indicator">
http://cdonet.UGent.be/project_y/indicator_1.htm</source>
<source type="body" resource="indicator">
http://cdonet.UGent.be/project_y/indicator_2.htm</source>
<source type="body" resource="indicator">
http://cdonet.UGent.be/project_y/indicator_3.htm</source>
</folder>

</folder>

```

Het komt er dus op neer dat men verwijst naar metadata in andere bronnen en daar eventueel nog enkele metadata aan toevoegt.

Het feit dat men een sub-folder moet creëren om een verdere stap in de hiërarchie te forceren is een punt waar sommige leken kunnen over struikelen.

Als men het niet doet is dat evenwel geen drama. De data zullen geogst worden, maar een beschrijving van een indicatorcollectie zal dan in een 'output' op dezelfde hiërarchische hoogte staan als een beschrijving van een indicator.

Een alternatief is volgend model:

```

<?xml version="1.0">
<harvest>

<contributor>idem als hierboven</contributor>

<meta level="1" name="issd.resource" content="organisation" />
<source level="1" type="head">http://cdonet.UGent.be</source>

```

```
<meta level="2" name="issd.resource" content="indicatorset" />
<meta level="2" name="dc.title" content="IDO Project X" />

<source level="3"
type="body">http://cdonet.UGent.be/project_x/indicator_1.htm</source>
<source level="3"
type="body">http://cdonet.UGent.be/project_x/indicator_2.htm</source>
<source level="3"
type="body">http://cdonet.UGent.be/project_x/indicator_3.htm</source>

<meta level="2" name="issd.resource" content="indicatorset" />
<meta level="2" name="dc.title" content="IDO Project Y" />

<source level="3"
type="body">http://cdonet.UGent.be/project_y/indicator_1.htm</source>
<source level="3"
type="body">http://cdonet.UGent.be/project_y/indicator_2.htm</source>
<source level="3"
type="body">http://cdonet.UGent.be/project_y/indicator_3.htm</source>

</harvest>
```

Hier wordt de hiërarchie geduid met het attribuut 'level'. Dit laatste zou het manuele beheer van de harvest-file mogelijks een stuk kunnen vergemakkelijken, want met de folder-constructie riskeert men heel wat problemen als men met cut en past de inhoud gaat herschikken in een zeer lange harvest-file die niet meer geheel op het scherm kan bekeken worden. Het laatste model met level-attributen kan beter gelezen worden als een boek.

Het folder-model is echter wel gemakkelijk om te verwerken met software omdat men er recursieve script-code kan op los laten, maar we gaan er dus vanuit dat velen de harvest-file manueel zullen aanmaken en niet met software, vandaar dat we voorlopig de tweede optie met level-attributen weerhouden.

VI.3 Archiveren

Wanneer men metadata in het archief wil opbergen, start men dus met het verzenden van een harvest-file naar het archief. Het archief presenteert dus een form waarmee men de harvest-file kan uploaden.

Deze harvest-file wordt ook in een databasetabel van harvest-files opgeslagen. Die tabel bevat records met volgende velden:

id
file
valid
date
email

id is de key van de record

file bevat dus het xml-document

valid bevat het resultaat van een validering van het xml-document (yes of no)

date bevat datum en uur van de upload

email bevat de boodschap die eventueel verstuurd is op de vermelde email-adressen in het element contributor

Het harvest-veld in de tabel met xhtml-documenten, bevat dus de key (id) van de harvest-record

De relaties binnen de database zijn dus boomvormig

- xml-harvest-record 1
 - xhtml-record 1-1
 - metadata-record 1-1-1
 - metadata-record 1-1-2
 - xhtml-record 1-2
 - metadata-record 1-2-1
 - metadata-record 1-2-2
 - metadata-record 1-2-3
 - metadata-record 1-2-4

Alle metadata die een organisatie aanbiedt kunnen dus geplaatst worden binnen een bredere context en men kan de zoekoperaties complementeren met de optie om hoger in het hiërarchisch niveau te kijken. waarbij eventueel een output wordt geproduceerd op basis van de xml-harvest-gegevens.

Elke organisatie die metadata aanbiedt kan nu de volledige controle bewaren over de metadata die ze aanbiedt en de manier waarop die in de database hiërarchisch worden geordend.

Het spreekt vanzelf dat een toeleverancier van metadata eerst kan bekijken hoe het archief de harvest-file heeft verwerkt, en de kans krijgt om de oogst te laten doorgaan (bevestiging) of om ze ongedaan te maken.

Achteraf kan men de inhoud van het archief nog steeds bijsturen door het verzenden van een nieuwe harvest-file met identieke username en paswoord.

VI.4 Open Source

Het is de bedoeling om bovenstaand databeheer-systeem te implementeren middels 'vrije software' of 'open source software'.

Vrije software is een verzamelnaam voor software waarvan de broncode¹ beschikbaar is voor iedereen, meestal onder de voorwaarde dat als men de software verder aanpast aan specifieke noden, deze bijkomende code ook beschikbaar wordt voor anderen.

De idee van Vrije Software gaat niet alleen over software op zich, maar ook over elke andere vorm van *informatie* ("content"), zoals bijvoorbeeld cursusmateriaal, afbeeldingen, presentaties, enz. Vrije Software wordt in het Engels vaak ook 'Free Software' (<http://www.fsf.org/>) of 'Open Source software' (<http://www.opensource.org/>) genoemd.

De term 'Free' heeft in eerste instantie de betekenis van *vrij* in de zin van *eigendom* (d.w.z., de software is eigendom van niemand), en pas in tweede instantie van *kostprijs* (d.w.z., de software is gratis). Het kostprijs-aspect is een *gevolg* van het vrij-zijn, en niet omgekeerd.

In het Engels verwoordt men dit vaak met de volgende plastische uitspraak van Richard M. Stallman: "*Free as in speech, not as in beer.*" (Stallman - <http://www.stallman.org/> - is de oprichter van de Free Software Foundation, en een onvermoeibare pleitbezorger voor Vrije Software.)

Vrije Software producten dragen een licentie² die vrij kopiëren toelaat en zelfs aanspoort. Vrije Software licenties promoten dus een zo groot mogelijke verspreiding van software en informatie, met enerzijds de juridische garantie dat geen enkel bedrijf of persoon het geproduceerde materiaal kan inpikken of de verspreiding ervan kan belemmeren, en anderzijds de zekerheid dat de creatieve initiatiefnemers steeds de waardering en krediet krijgen waarop ze recht hebben.

Enkele gebruikelijke Open Source Software licenties zijn:

- BSD (Berkeley Software Distribution)
- GPL (GNU General Public License)
- MPL (Mozilla Public License)

Een overzicht van licenties: <http://www.opensource.org/licenses/index.php>

Het besturingssysteem GNU/Linux is waarschijnlijk het meest bekende voorbeeld van Open Source Software. Daarnaast lopen er tal van projecten die op één of ander software-pakket focussen. Enkele van de meest gekende projecten zijn:

- <http://www.apache.org/> - Apache (dat meer dan 60% van de web-servers wereldwijd aanstuurt),
- <http://www.php.net/> - PHP (een scripttaal die voor dynamische web-pagina's zorgt),
- <http://www.mysql.com/> - MySQL en - <http://www.postgresql.org/> - PostgreSQL (de gegevensbanken die vaak in combinatie met PHP en Apache achter een webstek zitten),
- <http://www.latex-project.org/LaTeX-home.html> - LaTeX (een tekstverwerker van zeer hoge kwaliteit, en met talloze modulaire uitbreidingsmogelijkheden, zoals onder andere voor computerpresentaties),

¹ Je computer draait een programma door één voor één instructies uit zijn geheugen op te halen en uit te voeren. Deze instructies zijn verschillend van processor tot processor. Maar bijna niemand gebruikt deze basis-instructies rechtstreeks in zijn programma. Sinds vele tientallen jaren zijn er immers "hoog-niveau"-programmeertalen beschikbaar: Pascal, C, C++, Java, Ada, enz. Eén instructie uit een dergelijke hoog-niveau programmeertaal wordt door een compiler omgezet in een reeks van de hogervermelde laag-niveau instructies (de zogenaamde assembler. Open software programma's worden altijd beschikbaar gesteld in één van die hogere programmeertalen, en dit is de broncode van het programma. Linux zelf is geschreven in C. Bron: <http://people.mech.kuleuven.ac.be/~bruyninc/etos/>

² Een licentie is een stukje tekst dat met een programma meekomt, en dat de juridische aspecten van het gebruik en de verspreiding van de software verduidelijkt. Commerciële software heeft meestal een zeer restrictieve licentie; bijvoorbeeld: de software mag maar op één computer gebruikt worden, en niet door anderen gekopieerd worden. Open Software licenties beogen precies het tegenovergestelde: de software mag door iedereen gebruikt, gekopieerd en verdeeld worden. Het was Stallman die in de jaren 70(?) voor het eerst dit soort licentie gebruikte, onder de vorm van de GPL.

- <http://www.openoffice.org/> - OpenOffice.org (een zeer uitgebreide verzameling van dektop-software),
- <http://www.gimp.org/> - Gimp (een veelzijdig bitmap- en foto-bewerkingsprogramma),
- <http://www.gnome.org/> de Gnome en - <http://www.kde.org/> - KDE grafische gebruikersschillen voor Linux,
- <http://www.octave.org/> - Octave (uitgebreid reken-pakket), enz.

Omdat alle succesvolle Vrije Software projecten wereldwijd gebruikt worden zijn alle webstekken die hierboven vermeld werden in het Engels. De software zelf echter, heeft vaak Nederlandstalige of Franstalige versies.

Vrije Software is helemaal niet beperkt tot het GNU/Linux-besturingssysteem: de vrije beschikbaarheid van broncode en documentatie heeft er voor gezorgd dat de meeste programma's zoals de Apache-server ook draaien op andere Operating Systems zoals Windows, MacOS, UNIX, enzovoort.

Een publicatie van de Leuvense professor Herman Bruyninckx over de 'democratisering van de ICT' bevat een diepgaand pleidooi voor de toepassing van de Vrije Software in het onderwijs. Het document is aan te raden literatuur voor wie de essentie wil begrijpen van wat ICT in het onderwijs op dit ogenblik voorstelt, en van wat het zou kunnen zijn.

<http://people.mech.kuleuven.ac.be/~bruyinnc/ictvisie.html>

Bruyninckx geeft een opsomming van een aantal praktische argumenten zijn van toepassing op zowel Free Software als Open Source software.

1. *Vrij-beschikbare software van hoge kwaliteit biedt een tegengewicht tegen de dominante positie van Microsoft in onze scholen. Dit leidt tot kritische keuzevrijheid en de bijhorende competitieve klant-gerichtheid, wat op zijn beurt een ICT-opleiding van hogere kwaliteit tot gevolg heeft. Want de objectieve feiten zijn moeilijk te ontkennen: het Onderwijs fungeert op dit ogenblik als een gratis opleidingsinstituut voor Microsoft, en zulke verregaande vorm van mono-cultuur is nooit gezond op lange termijn.*
2. *Het gratis zijn van de software biedt de mogelijkheid aan scholen om hun budget voor informatica onder controle te houden, en om hun leerlingen verschillende alternatieven voor eenzelfde probleem of taak te laten uitproberen. (Dat geldt dus ook voor ngo's en openbare instellingen, wj en sd)*
3. *Vrije Software is bij uitstek geschikt voor de ontwikkeling van educatieve software: de openheid biedt de zekerheid dat iedereen die potentieel kan bijdragen (leraars, leerlingen, ouders, hogescholen en universiteiten) ook effectief de kans krijgt om bij te dragen, en dat die inspanningen voor iedereen beschikbaar zullen blijven in de toekomst....*
4. *Vrije Software ontwikkeling sluit een commerciële samenwerking en distributie helemaal niet uit: uitgever kunnen een meerwaarde leveren aan educatieve Vrije Software projecten door uitgaven te verzorgen van de software, tesamen met extra multimediaal materiaal, maar vooral door het beschikbare materiaal te selecteren, te bundelen, te filteren op kwaliteit, en aan te passen aan specifieke doelgroepen. Dit vereist wel een wijziging in hun bedrijfsmodel: het moet gebaseerd zijn op dienstverlening rond de software, en niet op eigendom op de software.*
5. *Vrije Software heeft, per definitie eigenlijk, steeds de belangen van, en het nut voor, de gebruiker voor ogen, en beschikt dus over een groot aantal werktuigen en tradities die commerciële software-producenten om strategische redenen niet aanbieden:*
 - *krachtige tekst-editors die op alle mogelijke computers draaien en de gebruikers dus toelaten om hun efficiëntie te verhogen aangezien zij altijd met hetzelfde werktuig met verschillende software-pakketten kunnen interageren.*
 - *absolute eerbied voor internationale afspraken en standaarden, terwijl commerciële aanbieders er alles aan doen om hun gebruikers aan zich te binden via gesloten, bedrijfseigen ("proprietaire") bestandsformaten.*
 - *gebruikers kunnen het voor hen meest geschikte hardware-platform kiezen, omdat ze indien nodig de beschikbare software kunnen hercompileren. Dit is echter in het algemeen geen eenvoudige opdracht, maar de kans is zeer groot dat meer competente mensen ergens anders in de wereld dit reeds voor eigen gebruik gedaan hebben en het resultaat ter beschikking stellen.*
 - *het gevaar dat bepaalde software niet langer ondersteund wordt is onbestaande: eens gebruikers over de software beschikken kan geen enkel bedrijf of organisatie hen het*

gebruiksrecht erop ontzeggen, of hen verhinderen om hun software te upgraden. De onafhankelijkheid van de gebruiker is dus absoluut gewaarborgd.

- *het internet is een onmisbaar instrument gebleken in de spectaculaire groei van Vrije Software gedurende de laatste vijf jaren. Niet alleen voor de ontwikkelaars van de software is het internet een ongeëvenaard instrument, maar zeker ook voor de gebruikers die on-line enorm veel documentatie en ondersteuning vinden bij andere gebruikers. De "helpdesk" van de Vrije Software werkt 24 uur op 24, 365 dagen per jaar.*
- 6. *De mogelijkheid om in de programma-code zelf te gaan kijken biedt aan alle informatica-leerkrachten en studenten een educatieve meerwaarde die geen enkel gesloten commercieel product kan evenaren.*
- 7. *Vrije Software besturingssystemen (zoals GNU/Linux of OpenBSD) zijn veel gemakkelijker te onderhouden en te beveiligen dan MacOS of Windows. Dit zou moeten leiden tot een lagere Total Cost of Ownership (TCO); in de praktijk hangt die TCO echter voor het grootste deel af van de competenties van de systeem-administrators. Bovendien is TCO erg moeilijk te meten: Hoeveel kost opleiding? Hoeveel kost overgang naar andere systemen? Hoeveel kost (gebrek aan) interoperabiliteit? ..."*

Open Source en duurzame ontwikkeling

Als men in het kader van duurzame ontwikkeling aandacht gaat besteden aan de zogenaamde digitale kloof (participatie), dan krijgt vrije software waaraan bovendien geen prijskaartje hangt, een bijkomende belangrijke betekenis.

Hieronder enkele sites die de '**digitale kloof**' viseren, maar dan wel alle aspecten van informatie en communicatietechnologie (ICT) belichten en niet uitsluitend het gebruik van open source software:

- Computers for Africa: <http://www.computers4africa.org>
- United Nations Information Technology Service: <http://www.unites.org>
- Digital Divide Network: <http://www.digitaldividenetwork.org>
- Digital Opportunity Task Force: <http://www.dotforce.org/>
- Africa ICT policy monitor: <http://www.apc.org/english/rights/africa/campaigns.shtml>
- E-inclusionsite.org - <http://www.e-inclusionsite.org>

De wereldtop over de informatiemaatschappij die nu wordt voorbereid en in 2005 wordt afgerond, besteed eveneens veel aandacht aan de digitale kloof:

- <http://www.itu.int/wsis/>
- <http://www.tunis2005.org/>
- <http://www.geneva2003.org/home/index01.htm>
- <http://www.wsis.info/>

Ook met het oog op milieuvriendelijkheid wordt veel voor het gebruik van Linux en Open Source gepleit. Tiny Linux bijvoorbeeld is een kleine Linux distributie, bedoeld voor oudere gerecycleerde PC's: <http://tiny.seul.org/>

Een andere compacte distributie is mulinux: <http://mulinux.sunsite.dk/>

Computerbank Australia is bijvoorbeeld een organisatie die oude computers verzamelt, desnoods herstelt, voorziet van vrije software en dan verdeelt onder minder kapitaalkrachtige mensen of organisaties (<http://www.computerbank.org.au/victoria/>)

"Computerbank Australia believes that GNU/Linux is the best choice of operating system as it is free of license issues and is able to run efficiently on older hardware"

De sociale economie heeft eveneens Open Source Software erkend als nuttig instrument.

Zie: <http://www.esnet.be/> - Sociale en solidaire economie versus informatie- en communicatietechnologie

En: <http://www.upcase.be/> - Upcase - UPCASE is een project dat een aantal ICT-toepassingen wil ontwikkelen specifiek voor de sociale economie. Het project wordt gefinancierd door de Federale overheid binnen het programma ter ondersteuning van de informatiemaatschappij.

"De informatiemaatschappij wordt meer en meer een realiteit. Nieuwe informatie- en communicatie-technologie (ICT) spelen hierbij een essentiële rol. De sociale economie kan uiteraard niet aan de kant blijven staan bij deze technologische revolutie en doet dit ook niet (bv. de opborrelende initiatieven rond de 'digitale' ongelijkheid). De kansen en mogelijkheden van ICT-toepassingen integreren in de visie en activiteit van de eigen organisatie is als speler binnen de sociale economie echter niet altijd gemakkelijk".

Enkele projectprincipes:

UPCASE biedt organisaties de kans om zulke ICT-toepassing uit te bouwen volgens volgende principes :

- **harmonie** wordt gezocht tussen de technologische toepassing en de organisatorische noden en behoeften;
- **inspraak en betrokkenheid** van de gebruikers (personeel, klant, lid, doelgroep, ?) is een belangrijke voorwaarde om deze harmonie te bewerkstelligen;
- **deskundige ondersteuning** wordt niet alleen verschaft over het technologisch aspect, maar ook over de organisatorische veranderingen die de ICT implementatie zou kunnen veroorzaken;
- er wordt gewerkt met **open software**. Dit is de verzamelnaam van alle software die als gezamenlijke eigenschap hebben dat ze de broncode ('source') vrijgeeft. Dit is in scherp contrast met commerciële software-bedrijven, die hun broncode angstvallig geheim houden, zodat vrije toepassing of open gebruik niet mogelijk is. Het Internet en Linux zijn waarschijnlijk de twee meest bekende voorbeelden van open software;
- als deelnemende organisatie treedt u op als **de voortrekker van een geheel**. Dit betekent enerzijds dat uw organisatie optreedt als proefpilot voor een toepassing die andere gelijkaardige organisaties ook kunnen gebruiken, anderzijds dat u wordt opgenomen in een stimulerend netwerk, dat als klankbord tijdens het project zal optreden.

AMP: Apache – MySQL – PHP

Om het metadata-archief zoals hierboven beschreven te organiseren zullen we ons beroepen op een zogenaamde open source AMP-distributie waarvan men verschillende versies gratis, of voor een redelijke prijs, van het Internet kan downloaden.

AMP staat voor de combiantie

- Apache – een server
- MySQL – een database
- PHP – een scriptaal

Heeft men het over een LAMP dan is de combinatie Linux-Apache-MySQL-PHP, een WAMP staat voor de combinatie Windows-Apache-MySQL-PHP

Een AMP kan men zowel voor Linux, voor Windows als voor andere Operating Systems verkrijgen. We sommen er een aantal op:

<http://www.foxserv.net/> - FoxServ is een open en vrije WAMP

<http://www.nusphere.com/> - NuSphere is een commerciële distributie met zijn eigen PHP IDE PHPEd. Bestaat uit Apache, MySQL, PHP en Perl (eveneens een scripttaal). NuSphere's kost 300 tot 500 dollar, en draait zowel op Linux als Windows.

<http://sourceforge.net/projects/phptriad> - PHPTriad, beroemd, maar naar verluidt niet langer up-to-date (een WAMP met Apache, MySQL, PHP and Perl, dus enkel voor Windows).

<http://www.sokkit.net/> - Sokkit is de commerciële versie van PHPTriad (20 tot 30 dollar)

<http://www.easypHP.org/?lang=en> - EasyPHP, een Franse WAMP

<http://firepages.com.au/> - Firepages phpdev (Apache, MySQL, PHP en Perl for Windows).

Consulteer ook overzichten zoals:

http://www.hotscripts.com/PHP/Software_and_Servers/Installation_Kits/

Er bestaan nog andere verwante acronyemen:

SWAMP: (Secure Windows Apache Mysql Php) is a fully configured, installation of Apache (mod_ssl), PHP, MySQL, OpenSSL and related administrative tools for Windows(TM) NT/2000/XP - <http://swamp.sourceforge.net/>

Last but not least:

XAMPP - <http://www.apachefriends.org/xampp-en.html> - zowel voor Linux als voor Windows

“The distribution for Linux system (tested for SuSE, RedHat, Mandrake and Debian) contains: Apache, MySQL, PHP & PEAR, Perl, ProFTPD, phpMyAdmin, OpenSSL, GD, Freetype2, libjpeg, libpng, gdbm, zlib, expat, Sablotron, libxml, Ming, Webalizer, pdf class, ncurses, mod_perl, FreeTDS, gettext, mcrypt, mhash, Turck MMCache, SQLite and IMAP C-Client.
XAMPP for Windows”

“The distribution for Windows 98, NT, 2000 and XP. This version contains: Apache, MySQL, PHP + PEAR, Perl, mod_php, mod_perl, mod_ssl, OpenSSL, phpMyAdmin, Webalizer, Mercury Mail Transport System for Win32 and NetWare Systems v3.32, JpGraph, FileZilla FTP Server, mcrypt, Turck MMCache, SQLite, and WEB-DAV + mod_auth_mysql.”

“XAMPP is free of charge”

XAMPP is een boeiend pakket dat als volgt wordt beschreven:

The Apache HTTP Server Project - At least from today's technical view: the mother of most web sites.
- <http://httpd.apache.org/>

The MySQL database

The middle letter of XAMPP. The fast and stable MySQL database.

<http://www.mysql.org/>

PHP: Hypertext Preprocessor

Maybe the best language to create dynamic generated web pages. But how to define best?

<http://www.php.net/>

Perl - Much older than PHP and in my opinion not the optimal language for generating web pages. But it's very, very popular. - <http://www.perl.org/>

ProFTPD - A fine FTP daemon. Easy and powerful. Much smaller and easier than the good old wuftpd. (only in the Linux version) - <http://www.proftpd.org/>

OpenSSL: The Open Source toolkit for SSL/TLS. No security without SSL. And without OpenSSL no SSL. - <http://www.openssl.org/>

Turck MMCache - The best PHP accelerator I know. Great software from Russia. Many thanks to Dmitry Stogov for this piece of software - <http://turck-mmcache.sourceforge.net/>

The FreeType Project - Freetype makes it possible to use TrueType fonts in GD.

<http://www.freetype.org/>

GD Graphics Library - And GD brings graphics into PHP -<http://www.boutell.com/gd/>

Independent JPEG Group - This graphics could be JPEGs.- <http://www.iijg.org/>

libpng.org - Or PNGs - <http://www.libpng.org/>

gdbm - Nice Unix database library used by PHP. I like this oldfashioned type of databases.
<http://www.gnu.org/software/gdbm/gdbm.html>

zlib - A basic compression library - <http://www.gzip.org/zlib/>

phpMyAdmin - phpMyAdmin makes it easy to use MySQL - <http://www.phpmyadmin.net/>

Webalizer - The famous Apache log file statistic tool: Webalizer - <http://www.mrunix.net/webalizer/>

Sablotron, libxml and expat - These three bring XML into your PHP: expat for event-based XML parsing, libxml for DOM and Sablotron for XSLT. (only in the Linux version)

<http://www.gingerall.com/> (Sablotron)

<http://xmlsoft.org/> (libxml)

<http://www.jclark.com/xml/expat.html> (expat)

Ming - Make your own Flash movies with Ming and PHP. (only in the Linux version)

<http://ming.sourceforge.net/>

pdf class - Open source and free PHP class to create on-the-fly PDF documents.

<http://www.ros.co.nz/pdf/>

Tenslotte vermelden we nog "New York PHP" - <http://nyphp.org/>

"New York PHP is a group of professionals dedicated to the development of PHP and AMP Technology. With over 1000 members worldwide, it's the largest such group in the United States. NYPHP actively support the PHP/AMP community through Open Source projects, sponsoring of networking events and meetings, and technical mailing lists. Developing partnerships with New York's business community and national PHP companies, New York PHP is working hard to promote and strengthen AMP Technology around the country".

VII. Producing XML- or XHTML-input through MS-Word

One of the main problems encountered when introducing XML or XHTML in an organisation or community is the existing 'habits' of data-exchange. Many people are rather conservative and stick to the procedures they have been using for many years. You can't just tell them that there are more efficient and effective procedures, which they can learn quickly and apply with a minimum of extra software. The main problem in convincing them is to find a good balance between the extra effort they have to make at the start, and what they gain by it later.

We have been looking for a technique that allows many people to produce XML- or XHTML-input without the installation of specific software. This procedure is based on a specific request by a researcher who had to produce documents which had to circulate between many authors. Those authors must have the possibility to change content and mark their changes with colours or strike-through. Traditionally they use a word processor like MS Word that supports these actions. The authors circulate the documents -their drafts- with email.

A draft could look like this:

"I is a fact that climate change will lead to longer summers in our country. Are you sure. I've heard it's going to rain more (johnny). We will discuss this in ~~chapter 5~~ section 4."

One editor marks his contributions (comments, changes) green and another editor used red to indicate changes. When all contributors are consulted and have made their changes in the text, one editor starts to clean up the document, ending up with a text that has no more markers (color, strike-through). For example:

"Some say the summers will be longer, others expect more rain. We will discuss this in section 4."

One could argue whether this is a good procedure. Fact is however, that those involved do not want to change. Sometimes it is impossible to change the procedure, while the only software shared by all editors is MS-Word and email. Some authors are not even allowed to install other software on the desktops of the administration or company they work for, and/or they will not do it just for one or two documents they are consulted for.

This practice is very common in Belgium especially when policy-planning documents are produced and the authors want to ask experts to comment on it. The production of the Flemish Environmental Indicators Reports (MIRA) is based on such an exchange of Word-documents between experts that are allowed to add, change or strike-through pieces of text.

When the text is finalised it is often cut and pasted in a more structured format (database-records or HTML/XML). This step is not a big job if only a few documents were produced, but what if hundreds of them are made? How many times are we going to cut and past?

How can this be avoided? Here we'll make a proposal that works fine in a context described (exchanging MS-Word-documents by email, and allowing marking changes with colors).

The final product is an XML- or XHTML-file, that can be exchanged between different platforms (MacOS, Linux, ...), but we focus on the production of content with MS-Windows and MS-Word, because that is the most widely distributed desktop-system in Belgium.

First example

Suppose that -in the end- you want to produce XML-documents with the following structure (a simple Dublin Core Metadata-set and an Administrative Core)

```
<?xml version="1.0" encoding="iso-8859-1" ?>  
<metadata_issd_dc_simple>
```

```

<AC>
<AC.name></AC.name>
<AC.email></AC.email>
<AC.contact></AC.contact>
<AC.activity></AC.activity>
<AC.date></AC.date>
<AC.dateRange></AC.dateRange>
<AC.rights></AC.rights>
</AC>

<DC>
<DC.Title></DC.Title>
<DC.Identifier></DC.Identifier>
<DC.Subject></DC.Subject>
<DC.Description></DC.Description>
<DC.Creator></DC.Creator>
<DC.Contributor></DC.Contributor>
<DC.Publisher></DC.Publisher>
<DC.Language></DC.Language>
<DC.Date></DC.Date>
<DC.Type></DC.Type>
<DC.Format></DC.Format>
<DC.Source></DC.Source>
<DC.Coverage></DC.Coverage>
<DC.Rights></DC.Rights>
<DC.Relation></DC.Relation>
</DC>
</metadata_issd2_dc_simple>

```

The first thing we do is hiding these mark-up tags in a Word-document that must be circulated.

- Step 1: create tables with one column and many rows .
- Step 2: write the XML-code in table-cells, leaving one cell open where contributors must write content. Colour the XML-cells grey.
- Step 3: Save this document as the 'template' (basic 'empty' form).

At this stage you have a document that is operational, but not yet user friendly. You can however test the procedure. If the authors write text in the cells without XML, and the final version is saved as plain text (**txt instead of word-format doc**), the result will be a well-formed XML-document

The basic idea is that the Word-format (doc) is used during editing and circulation between different contributors, but once the text is finished, it is saved as plain 'text' and this txt-document is well-formed XML.

The first table below is a Word-document that could be used to experiment with. On the next page you can find a document that is more user-friendly. Using **XML-comments**, we've added information to instruct the authors. The **XML-code is hidden**. Consult your Word-help files to find out how to hide (and display) pieces of text (you must select the text that you want to hide and change the 'character-type' in hidden).

<?xml version="1.0" encoding="iso-8859-1" ?> <metadata_issd2_dc_simple>
--

<AC><AC.name>

</AC.name><AC.email>

</AC.email><AC.contact>

</AC.contact><AC.activity>

</AC.activity><AC.date>
</AC.date><AC.dateRange>
</AC.dateRange><AC.rights>
</AC.rights></AC>

<DC><DC.Title>
</DC.Title><DC.Identifier>
</DC.Identifier><DC.Subject>
</DC.Subject><DC.Description>
</DC.Description><DC.Creator>
</DC.Creator><DC.Contributor>
</DC.Contributor><DC.Publisher>
</DC.Publisher><DC.Language>
</DC.Language><DC.Date>
</DC.Date><DC.Type>
</DC.Type><DC.Format>
</DC.Format><DC.Source>
</DC.Source><DC.Coverage>
</DC.Coverage><DC.Rights>
</DC.Rights><DC.Relation>
</DC.Relation></DC></metadata_issd2_dc_simple>

Click on display/hide button of Word in order to view hidden code

Please, do not edit in grey cells. they contain hidden code to support data-management. !!! The first table with the administrative core can be duplicated (copy and past) if necessary. Do not use following special characters: &, < and > . If you need them, type: & (ampersand) or < (lower then) or > (greater then).

Name of author
Email of author
Contact (phone, fax, address) of author

Activity of author (What have you done with text?)
Date of activity
DataRange (optional)
Copyrights for metadata

Title (name) of information resource
Identifier (URL, ISBN,..) of information resource
Subject (keywords)
Description (abstract, contents)
Creator (author, editor)
Contributor
Publisher
Language of resource
Publishing date (yyyy-mm-dd)
Type (book, article, ...)
Format (xml, pdf, word, xml,...)
Source (larger context)
Coverage (geographic and temporal)
Copyrights
Related resources

Questions and Answers:

Does this procedure replace database management systems?

No. We are not discussing the management of files here, only the production of content in an environment that is 'primitive' from an Information-Technical perspective . And we can not consider using distributed authoring and versioning (WebDav)¹, because many contributors do not know about this.

Is it robust?

It's as robust as MS-Word is. Our experience is that it works better than Word-forms if the documents have to be exchanged between different versions of Word. It is however important to tell the authors what you're doing and to give clear instructions. Do not send them a template without a good definition of the content that should be written in the cells. For the example demonstrated above, the authors should of course know what a Dublin Core and an Administrative Core is. Using comment-tags, we can of course insert the definitions of the DC- and AC-elements in the 'Word-form'.
Back up the files regularly.

Does it work anywhere in the world?

Japanese people and many others that use languages with special characters have to adjust the encoding-value of the xml-instruction. Here we've used :

```
<?xml version="1.0" encoding="iso-8859-1" ?>
```

iso-8859-1 is the set of Latin/Western characters (Dutch, English, French, German, Spanish,...).

But this is not a fundamental problem.

Can you support editing of attribute values too?

Yes. For example: editing an anker could be supported with following table:

<!-- URL --><a href="
"><!-- Title of link -->

When XML code is hidden the table looks like this:

URL
Title of link

¹ <http://www.webdav.org/> - Briefly: WebDAV stands for "Web-based Distributed Authoring and Versioning". It is a set of extensions to the HTTP protocol which allows users to collaboratively edit and manage files on remote web servers.

It is important to notice that this option might become tricky when you allow duplication of elements. In our opinion one should first produce xml-files where no attributes are edited by the contributors. Elements are used to store data for attributes and later these xml-files can be transformed (with XSLT for example) to the structure that is preferred. One might try to have it all in one step, but that is probably not always possible.

There is a comment inside the a-element, but this does not cause problems. The following document is valid:

```
<?xml version="1.0" ?>
<!DOCTYPE xmltest [
<!ELEMENT xmltest (a)>
<!ELEMENT a (#PCDATA)>
<!ATTLIST a href CDATA #REQUIRED>
]>

<xmltest>
<!-- URL --><a href="
"><!-- Title of link -->
</a>
</xmltest>
```

I want to provide selection lists with controlled vocabularies; how can I do that in Word?

One might consider developing macro's (visual basic) attached to the Word-template. Is this robust? We do not know. There is of course the option to exploit all the possibilities that MS-Office has to offer. On the other hand a metadata-author can also simply cut and past words (from a thesaurus or controlled vocabulary) from a list, that is available as another Word document.

The vocabulary could also be placed in another cell and contained in comment tags. The author should then copy and past his selected words to the 'white' cells.

Secondly, the editing of certain elements could be supported with a HTML-form and the appropriate ECMA-scripts (Javascrpts) that supports the production of controlled vocabularies. The metadata-creator must then past some text from his browser to the Word document.

Remember, the basic goal here was to find a procedure that allows circulation of drafts, where many authors can add, change or delete the content of a document and mark these changes with colours. This procedure does not replace database-management systems or client-server systems with HTML-forms etc...The starting point is: most people can only work with 'things' like Word; they can not install xml-software (some of them are still surfing with IE 4.0) , they do not even know what XML is; they can not install professional editors like (Emacs)².They are used to write some text and mail it to an other person who can change it a bit, mail it again, etcetera....These are the users we are considering here. They do exist.

There are certainly limits to this technique. In many cases it is better to work with html-forms and database-support. But one should notice that many database-systems do not support the use of colours and other markers.

One should also take notice of the fact that this procedure works offline.

Another example: XML document to be transformed with XSLT

This example is only demonstrating the fact that xml can be produced for further processing with XML-specific tools like XSLT.

² <http://www.gnu.org/software/emacs/emacs.html>

The following pages demonstrate the xml-code, xslt-code, Word-style, etc. used for the production of some metadata of favorite websites. The layout of the Word document is different from the previous example (it is indeed a matter of taste).

bookmarks_input_template.xml

Empty template of bookmarks input-file. The file below has one AC-element (administrative core) en two DC-elements (Dublin Core). The DTD accepts one AC-element and one or more DC-elements.

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<!-- Bookmark metadata -->
<DC.Input>
<AC>
<!-- Your name -->
<AC.Name></AC.Name>
<!-- Your Email-address -->
<AC.Email></AC.Email>
<!-- Other contact (phone,...) -->
<AC.Contact></AC.Contact>
<!-- Date: yyyy-mm-dd -->
<AC.Date></AC.Date>
</AC>

<DC >
<!-- URL (web-address) -->
<DC.Identifier></DC.Identifier>
<!-- Title (name) -->
<DC.Title></DC.Title>
<!-- Other title (acronym) -->
<DC.Title.alternative></DC.Title.alternative>
<!-- Description -->
<DC.Description></DC.Description>
<!-- Keywords -->
<DC.Subject></DC.Subject>
<!-- Language -->
<DC.Language></DC.Language>
<!-- Publisher -->
<DC.Publisher></DC.Publisher>
<!-- bookmark-folder-->
<DC.Subject scheme="DCMIDCSV">bookmarkfolder= </DC.Subject>
</DC>

<DC >
<!-- URL (web-address) -->
<DC.Identifier></DC.Identifier>
<!-- Title (name) -->
<DC.Title></DC.Title>
<!-- Other title (acronym) -->
<DC.Title.alternative></DC.Title.alternative>
<!-- Description -->
<DC.Description></DC.Description>
<!-- Keywords -->
<DC.Subject></DC.Subject>
<!-- Language -->
<DC.Language></DC.Language>
<!-- Publisher -->
<DC.Publisher></DC.Publisher>
<!-- bookmark-folder-->
<DC.Subject scheme="DCMIDCSV">bookmarkfolder= </DC.Subject>
```

```
</DC>
<!-- End of input -->
</DC.Input>
```

bookmarks_input.dtd

Document Type Definition (DTD) of bookmarks_input.xml file.
Tested with HTML-kit, Cooktop 2.200³ and Internet Explorer .

```
<!DOCTYPE DC.Input [
<!ELEMENT DC.Input (AC, DC*)>
<!ELEMENT AC (AC.Name, AC.Email, AC.Contact, AC.Date)>
<!ELEMENT DC (DC.Identifier | DC.Title | DC.Title.alternative |
DC.Description | DC.Subject | DC.Publisher | DC.Language)*>
<!ELEMENT AC.Name (#PCDATA)>
<!ELEMENT AC.Email (#PCDATA)>
<!ELEMENT AC.Contact (#PCDATA)>
<!ELEMENT AC.Date (#PCDATA)>

<!ELEMENT DC.Identifier (#PCDATA)>
<!ELEMENT DC.Title (#PCDATA)>
<!ELEMENT DC.Title.alternative (#PCDATA)>
<!ELEMENT DC.Description (#PCDATA)>
<!ELEMENT DC.Subject (#PCDATA)>
<!ELEMENT DC.Language (#PCDATA)>
<!ELEMENT DC.Publisher (#PCDATA)>
<!ATTLIST DC.Subject scheme CDATA #FIXED "DCMIDCSV">
]>
```

bookmarks_input_example.xml

In this example the comments (<!-- --> have been removed.

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<DC.Input>
<AC>
<AC.Name></AC.Name>
<AC.Email>walter.dejonge@rug.ac.be</AC.Email>
<AC.Contact></AC.Contact>
<AC.Date></AC.Date>
</AC>

<DC>
<DC.Identifier>http://cdonet.rug.ac.be</DC.Identifier>
<DC.Title>Centrum voor Duurzame Ontwikkeling - Universiteit Gent -
België</DC.Title>
<DC.Title.alternative>CDO-UG</DC.Title.alternative>
<DC.Description>
Het Centrum voor Duurzame Ontwikkeling heeft tot doel op een
multidisciplinaire wijze duurzame ontwikkeling als beleidskader verder uit
te bouwen.
De werking is gericht op onderzoek naar en ondersteuning van rationele
besluitvorming en doelmatigheid van beleidsinstrumenten, in functie van
langetermijndoelstellingen, in een geïntegreerd perspectief met maximale
aandacht voor participatie en communicatie.
```

³ <http://www.xmlcooktop.com/> - Cooktop is an editor and development environment for XML, DTD, and XSLT documents. Cooktop is a Windows application. Best of all, it's free!

In de multidisciplinaire aanpak dienen ondermeer volgende vijf hoofdlijnen aanwezig te zijn : ecologische, sociale, economische, institutionele en ethische facetten.

```

</DC.Description>
<DC.Subject>
indicatoren, leefbare steden, stad, lokaal, ketenanalyse, productbeleid,
verkeersveiligheid, grondstoffen, voorraadbeheer,
milieugebruiksruimte, lokale agenda21, agenda 21, integraal ketenbeheer
</DC.Subject>
<DC.Language>nl</DC.Language>
<DC.Publisher>CDO-UG</DC.Publisher>
<DC.Subject scheme="DCMIDCSV">bookmarkfolder=1</DC.Subject>
</DC>

<DC>
<DC.Identifier>http://cdonet.rug.ac.be</DC.Identifier>
<DC.Title>
Centre for Sustainable Development - University of Ghent -
Belgium</DC.Title>
<DC.Title.alternative>CSD-UG</DC.Title.alternative>
<DC.Description>
The Centre for Sustainable Development (CSD) is an interdisciplinary
research group drawing on the knowledge and expertise of nine departments
at Ghent University.
The aim of the centre is to use an interdisciplinary methodology for the
elaboration on the sustainable development as a social framework. This
approach includes, but is not limited to, research in the following five
keyfields: social, economic, ecological, institutional and ethical.
The CSD undertakes three kinds of tasks in order to achieve its aims:
research, education and provision of a social service.
</DC.Description>
<DC.Subject>
indicators, viable cities, city, chain analysis, resource management,
environmental utilisation space
</DC.Subject>
<DC.Language>en</DC.Language>
<DC.Publisher>CSD-UG</DC.Publisher>
<DC.Subject scheme="DCMIDCSV">bookmarkfolder=1</DC.Subject>
</DC>
</DC.Input>
    
```

bookmarks_input.doc (input template MS-Word: xml is visible here)

```

<?xml version="1.0" encoding="iso-8859-1" ?>
<!-- Bookmark metadata -->
    
```

<pre> <DC.Input><AC><!-- Your name --> <AC.Name> </pre>		<pre> </AC.Name> </pre>
<pre> <!-- Your Email-address --> <AC.Email> </pre>		<pre> </AC.Email> </pre>
<pre> <!-- Other contact (phone,...) --> <AC.Contact> </pre>		<pre> </AC.Contact> </pre>
<pre> <!-- Date: yyyy-mm-dd -- > <AC.Date> </pre>		<pre> </AC.Date></AC> </pre>

<DC >		
<!-- URL (web-address) --> <DC.Identifier>		</DC.Identifier>
<!-- Title (name) --> <DC.Title>		</DC.Title>
<!-- Other title (acronym) --> <DC.Title.alternative>		</DC.Title.alternative>
<!-- Description --> <DC.Description>		</DC.Description>
<!-- Keywords --> <DC.Subject>		</DC.Subject>
<!-- Language --> <DC.Language>		</DC.Language>
<!-- Publisher --> <DC.Publisher>		</DC.Publisher>
<!-- bookmark-folder--> <DC.Subject scheme="DCMIDCSV">bookmarkfolde r= </DC>		</DC.Subject>

```
<!-- End of input --></DC.Input>
```

bookmarks_input.doc (Input template MS-Word with hidden xml-code).

Third column can be restricted to minimum-size and table-cells that are not used for input can be colored black.

Bookmark metadata

Your name		
Your Email-address		
Other contact (phone,...)		
Date: yyyy-mm-dd		

URL (web-address)		
Title (name)		
Other title (acronym)		
Description		
Keywords		
Language		
Publisher		
bookmark-folder		

URL (web-address)		
Title (name)		
Other title (acronym)		
Description		
Keywords		
Language		

Publisher	
bookmark-folder	

End of input

bookmark_input.xsl⁴

XSLT-file to transform bookmark-input.xml file to html. This file does not classify bookmarks in specific bookmark folders (the folder-info is not transformed to html).

This xsl-file was tested with Cooktop and Internet Explorer 5.5

```
<?xml version="1.0"?>
<xsl:stylesheet                                version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:output method = "html" encoding="iso-8859-1" />
<xsl:output doctype-public="-//W3C//DTD HTML 4.0 Transitional//EN" />
<xsl:output doctype-system="http://www.w3.org/TR/REC-html40/loose.dtd" />

  <xsl:template match="/">
  <html>
  <title>Bookmarks</title>
  <style type="text/css">
  body {margin-left: 10%;
        margin-right; 10%}
  h2 {font-family: arial;
      font-size: 14 pt;
      color: navy;}
  span.name {font-family: arial;
            font-size: 11 pt;
            font-weight:bold;
            color: navy;}
  span.content {font-family: arial;
               font-size: 11 pt;
               color: black;}
  a {font-family: arial;
     font-size: 11 pt;
     color: blue;}
  </style>
  <body>
  <h2>Bookmarks</h2>
  <xsl:for-each select = "DC.Input/DC">
  <span class="name">URL (www-addresss): </span>
  <a>
  <xsl:attribute name="href">
  <xsl:value-of select="DC.Identifier"/>
  </xsl:attribute>
  <xsl:value-of select="DC.Identifier"/>
  </a>
  <br />
  <span class="name">Language: </span>
```

⁴ XSLT is a strong tool but it takes some time to learn it. If you do not have to develop XSLT-code regularly, you'll probably forget how it works and you'll have to learn it again.

XSLT is an interesting tool, but has not the same broad application-'power' as writing scripts with the Document Object Model. If one has to choose what to learn first: XSLT or scripting with the DOM, the latter might be the best option, especially if one is already producing scripts (PHP, Perl, Javascript) for other applications.


```

    <span          class="content"><xsl:value-of          select="DC.Language"
/></span><br/>
    <span class="name">Title: </span>
    <span class="content"><xsl:value-of select="DC.Title" /></span><br/>
    <span class="name">Alternative Title: </span>
    <span class="content"><xsl:value-of select="DC.Title.alternative" />
</span><br/>
    <span class="name">Description: </span>
    <span          class="content"><xsl:value-of          select="DC.Description"
/></span><br/>
    <span class="name">Subject (keywords): </span>
    <span          class="content"><xsl:value-of          select="DC.Subject"
/></span><br/>
    <hr />
</xsl:for-each>
</body>
</html>
</xsl:template>
</xsl:stylesheet>

```

HTML result after transformation

```

<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN"
"http://www.w3.org/TR/REC-html40/loose.dtd">
<html>
<title>Bookmarks</title>
<style type="text/css">
  body {margin-left: 10%; margin-right: 10%;}
  h2 {font-family: arial; font-size: 14 pt; color: navy;}
  span.name {font-family: arial; font-size: 11 pt; font-weight:bold;
color: navy;}
  span.content {font-family: arial; font-size: 11 pt; color: black;}
  a {font-family: arial; font-size: 11 pt; color: blue;}
</style>
<body>
<h2>Bookmarks</h2>
<span class="name">URL (www-addresss): </span>
<a href="http://cdonet.rug.ac.be">http://cdonet.rug.ac.be</a><br />
<span class="name">Language: </span><span class="content">nl</span><br />
<span class="name">Title: </span>
<span class="content"> Centrum voor Duurzame Ontwikkeling - Universiteit
Gent - België</span><br />
<span class="name">Alternative Title: </span>
<span class="content"> CDO-UG </span><br />
<span class="name">Description: </span>
<span class="content">Het Centrum voor Duurzame Ontwikkeling heeft tot doel
op een multidisciplinaire wijze duurzame ontwikkeling als beleidskader
verder uit te bouwen. De werking is gericht op onderzoek naar en
ondersteuning van rationele besluitvorming en doelmatigheid van
beleidsinstrumenten, in functie van lange termijn doelstellingen, in een
geïntegreerd perspectief met maximale aandacht voor participatie en
communicatie. In de multidisciplinaire aanpak dienen ondermeer volgende
vijf hoofdlijnen aanwezig te zijn : ecologische, sociale, economische,
institutionele en ethische facetten. </span><br />
<span class="name">Subject (keywords): </span>
<span class="content">indicatoren, leefbare steden, stad, lokaal,
ketenanalyse, productbeleid, verkeersveiligheid, grondstoffen,
voorraadbeheer, milieugebruiksruimte, lokale agenda21, agenda 21, integraal
ketenbeheer</span><br />
<hr />
<span class="name">URL (www-addresss): </span>
<a href="http://cdonet.rug.ac.be">http://cdonet.rug.ac.be</a><br />

```

```
<span class="name">Language: </span><span class="content">en</span>
<br /><span class="name">Title: </span>
<span class="content">Centre for Sustainable Development - University of
Ghent - Belgium</span><br />
<span class="name">Alternative Title: </span>
<span class="content"> CSD-UG</span><br />
<span class="name">Description: </span>
<span class="content">The Centre for Sustainable Development (CSD) is an
interdisciplinary research group drawing on the knowledge and expertise of
nine departments at Ghent University. The aim of the centre is to use an
interdisciplinary methodology for the elaboration on the sustainable
development as a social framework. This approach includes, but is not
limited to, research in the following five keyfields: social, economic,
ecological, institutional and ethical.
The CSD undertakes three kinds of tasks in order to achieve its aims:
research, education and provision of a social service.</span><br />
<span class="name">Subject (keywords): </span>
<span class="content">indicators, viable cities, city, chain analysis,
resource management, environmental utilisation space</span><br />
<hr />
</body>
</html>
```

XHTML-example

Producing XML first, to transform it with XSLT to (X)HTML is an option, but a better option is to produce XHTML in the first place.

Following document produces XHTML-files which can be used immediately on a website. The XHTML-tags are hidden in grey area. It's a document used in a project where Dutch content was produced.

CDO projectfiche

Naam van het element:	Titel	
Taalcode:	nl	
Schema:	-	
Inhoud (volledige titel):		

Naam van het element:	Auteurs	
Taalcode:	-	
Schema:	-	
Inhoud (naam van de auteurs, onderzoekers):		

Naam van het element:	Beschrijving	
Taalcode:		
Schema:	-	
Inhoud (volledige projectbeschrijving):		

--	--	--

Naam van het element:	Opdrachtgever	
Taalcode:	-	
Schema:	-	
Inhoud (naam van de opdrachtgevers):		

Naam van het element:	Datum	
Taalcode:	-	
Schema:	W3CDTF	
Inhoud (yyyy-mm-dd):		

Naam van het element:	WWW-adres	
Taalcode:	-	
Schema:	URI	
Inhoud:http://		
Herhaal inhoud:http://		

Naam van het element:	Taal	
Taalcode:	-	
Schema:	RFC3066	
Inhoud (taal van het document, volgens code):		

Naam van het element:	Format	
Taalcode:	-	
Schema:	-	
Inhoud (code voor documentformat):		

Naam van het element:	Metadata-auteur	
Taalcode:	-	
Schema:	-	
Inhoud (initialen van degene die deze fiche heeft ingevuld):		

Naam van het element:	Invuldatum	
Taalcode:	-	
Schema:	W3CDTF	
Inhoud (datum waarop deze fiche is ingevuld/bewerkt, yyyy-mm-dd):		

Hidden XHTML-code

```
<!-- CDO projectfiche -->
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title> </title>
<meta http-equiv="Content-type" content="text/html; charset=iso-8859-1" />
<link rel="stylesheet" type="text/css" href="metadata_body1.css" />
</head>
<body>
<div class="element"><div class="name" lang="en"> dc.title
</div>
<div class="local_name"><!--Naam van het element: -->
Titel
</div>
<div class="lang"><!--Taalcode: -->
nl
</div>
<div class="scheme"><!--Schema: -->
-
</div>
<div class="content" lang="nl"><!--Inhoud (volledige titel): -->
</div></div>
<div class="element"><div class="name" lang="en">dc.creator
</div>
<div class="local_name"><!--Naam van het element: -->
Auteurs
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
<div class="scheme"><!--Schema: -->
-
</div>
<div class="content" lang="nl"><!--Inhoud (naam van de auteurs,
onderzoekers):-->
</div></div>
<div class="element"><div class="name" lang="en">dc.description
</div>
<div class="local_name"><!--Naam van het element: -->
Beschrijving
</div>
<div class="lang"><!--Taalcode: -->
</div>
<div class="scheme"><!--Schema: -->
-
</div>
<div class="content" lang="nl"><!--Inhoud (volledige projectbeschrijving):
-->
</div></div>
<div class="element">
<div class="name"><!--Naam van het element: -->
Opdrachtgever
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
```

```

<div class="scheme"><!--Schema: -->
-
</div>
<div class="content" lang="nl"><!--Inhoud (naam van de opdrachtgevers):-->
</div></div>
<div class="element"><div class="name" lang="en">dc.date
</div>
<div class="local_name"><!--Naam van het element: -->
Datum
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
<div class="scheme"><!--Schema: -->
W3CDTF
</div>
<div class="content" lang="nl"><!--Inhoud (yyyy-mm-dd): -->
</div></div>
<div class="element"><div class="name" lang="en">dc.identificer
</div>
<div class="local_name"><!--Naam van het element: -->
WWW-adres
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
<div class="scheme"><!--Schema: -->
URI
</div>
<div class="content" lang="nl"><!--Inhoud: http://--><a href="http://
">
<!--Herhaal inhoud:--> http://
</a></div></div>
<div class="element"><div class="name" lang="en">dc.language
</div>
<div class="local_name"><!--Naam van het element: -->
Taal
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
<div class="scheme"><!--Schema: -->
RFC3066
</div>
<div class="content" lang="nl"><!--Inhoud (taal van het document, volgens
code):-->
</div></div>
<div class="element"><div class="name" lang="en">dc.format
</div>
<div class="local_name"><!--Naam van het element: -->
Format
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
<div class="scheme"><!--Schema: -->
-
</div>
<div class="content" lang="nl"><!--Inhoud (code voor documentformat):-->
</div></div>
<div class="element"><div class="name" lang="en">ac.name

```

```

</div>
<div class="local_name"><!--Naam van het element: -->
Metadata-auteur
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
<div class="scheme"><!--Schema: -->
-
</div>
<div class="content" lang="nl"><!--Inhoud (initialen van degene die deze
fiche heeft ingevuld):-->
</div></div>
<div class="element"><div class="name" lang="en">ac.date
</div>
<div class="local_name"><!--Naam van het element: -->
Invuldatum
</div>
<div class="lang"><!--Taalcode: -->
-
</div>
<div class="scheme"><!--Schema: -->
W3CDTF
</div>
<div class="content" lang="nl"><!--Inhoud (datum waarop deze fiche is
ingevuld/bewerkt, yyyy-mm-dd):-->
</div></div>
<hr />
<p>
&copy; cdo<br />

</p>
</body>
</html>

```

Take notice of the fact that the code includes a 'link' to an external CSS-file (Cascading Style Sheet) 'metadata_body1.css'. In some cases it might be a better option to include CSS in document, or to include the full URL of the CSS-file: http://cdonet.UGent.be/.../metadata_body.ccs.

VIII. Using frames on an indicator website

A lot of information about sustainable development and indicators is available online (www). So, it is obvious to add the www-address or Uniform Resource Locator of this information to the metadata. Often this address (something like http://www.org.org/folderx/indicator_x.htm) is used as a unique identifier of the information-resource.

The identifier of the metadata can become a problem if it is a URL of a page that is presented in a website with frames. Frames are often used when a project leads to the production of many separate documents, for example: documents that describe indicators. In such cases a frame is build with -for example- on the left side a list of the indicators, and on the right side the indicator that is selected in the list.

Figure 1 is a frameset with two frames. Often there is a third top-frame (header) with title and logo and sometimes there are even more frames. But our problems and solutions here will be explained for a two-frame situation (actually the solution works for any number of frames).

list of indicators • ind1 • ind2 • ind3	document describing selected indicator
--	--

Figure 1.

To create a frameset like figure 1 in a browser, two basic html-documents are needed:

- A document with the frame-structure (we will call that "**frameset.htm**")
- A document with the list of indicators (we will call that "**indicatorlist.htm**")

Next we have the documents describing the indicators are **ind1.htm**, **ind2.htm**, **ind3.htm**, etcetera. Finally there is a **default.htm**, a page presented in the right frame when no specific indicator is selected (that is when the visitor 'arrives' at the scene).

The basic content of **frameset.htm** could be:

```
<html>
<head>
<title>Home of a framed site</title>
</head>
<frameset cols="30%,70%">
<frame src="indicatorlist.htm">
<frame src="default.htm" name="main">
</frameset>
</html>
```

In order to focus on the frame-issue here we have not written tidy XHTML and not included a no-frame-option, nor meta-elements (Dublin Core) in the frameset-page, but it is not a bad idea to think about that too when you'll adjust your framed-pages.

The content of **default.htm** is a simple message:

```
<html>
<head>
<title>default.htm</title>
```

```
</head>
<body>
<h1>Hello World!</h1>
<p>Select an indicator in the list. Click with your left-mouse button on
one of the indicators.</p>
</body>
</html>
```

The **indicatorlist.htm** document contains a list of links to the indicator-documents (ind1.htm, ...) with the target-frame **main** (name of the right frame, in the frameset-page).

```
<html>
<head>
<title>indicatorlist.htm</title>
</head>
<body>
<ul>
<li><a href="ind1.htm" target="main">Indicator 1</a></li>
<li><a href="ind2.htm" target="main">Indicator 2</a></li>
<li><a href="ind3.htm" target="main">Indicator 3</a></li>
<li><a href="ind4.htm" target="main">Indicator 4</a></li>
</ul>
</body>
</html>
```

An indicator-file looks like this: **ind1.htm**

```
<html>
<head>
<title></title>
</head>
<body>
<h1>Indicator 1</h1>
<p>Here comes the description of indicator 1</p>
</body>
</html>
```

The same structure can be used for **ind2.htm**, **ind3.htm** and **ind4.htm**

If you put all these files in the same directory and click on **frameset.htm**, we get a picture like figure 2 (if you're reading a word document, the links will not work):

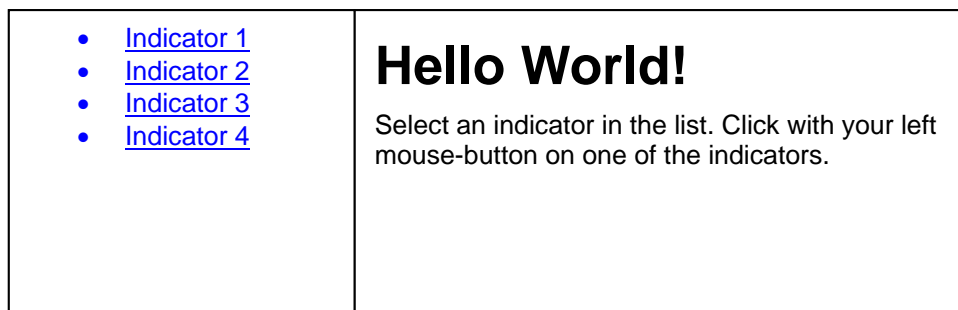
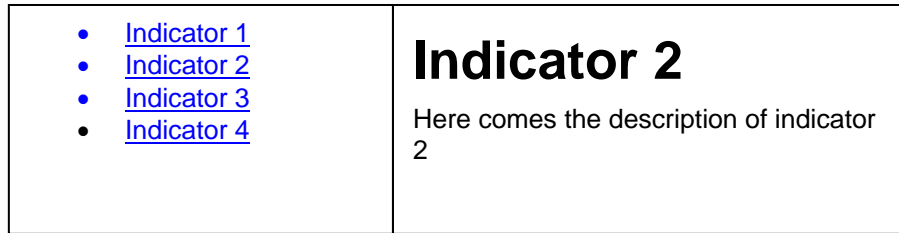


Figure 2

When the link of indicator 2 is clicked, the output on screen changes (figure 3):

**Figure 3**

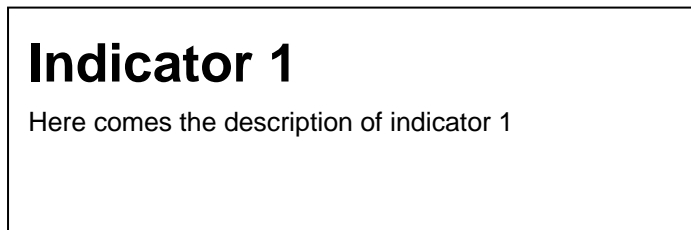
So far so good. This website will function perfectly...if it wants to remain an island on the huge internet.

The problem:

Suppose that we want to make metadata of the indicators 1,2, 3 and 4. Of course we would like to add the filenames of the indicators to these metadata. The basic metadata would then be, for example:

Title: Indicator 1
Identifier: <http://organisation.org/indicators/ind1.htm>

However if someone uses this identifier and types it in the address-field of his browser, the response will look like this (figure 4):

**Figure 4**

No frames, no context...The ind1.htm file is now an 'orphan'. If the user of metadata was only interested in this file, he/she will be happy, but what if the user wants to know more about the context?

One option is to add context-information to the metadata:

Title: Indicator 1
Identifier: <http://organisation.org/indicators/ind1.htm>
Source: <http://organisation.org/indicators/frameset.htm>

It's a good option and metadata-providers should use it, but in that case one must assume that all users understand this metadata-set. It is very possible that a visitor selects the identifier, and gets frustrated with the orphan because he/she did not see the source-link.

Another option is to add links to the frameset in all the indicator-documents (as we have done with **ind4.htm**).

```
<html>
<head>
<title></title>
</head>
<body>
<a href="frameset.htm">To homepage of this project</a><br/>
<h1>Indicator 4</h1>
<p>Here comes the description of indicator 4</p>
```

```
</body>  
</html>
```

[To homepage of this project](#)

Indicator 4

Here comes the description of indicator 4

Figure 5

It's a bad solution. Things go wrong when **ind4.htm** is already framed, that is when **ind4.htm** has been selected in the **indicatorlist.htm** and is not approached from the metadata. If -in that case- the 'homepage-link' is clicked, the left frame will be split up in two new frames.

This will not happen in our examples, since we've added a deframing JavaScript in the **frameset.htm** file. This script is very brief:

```
<script language="JavaScript" type="text/javascript">  
<!--  
// deframing from other frames  
  
if (top != self) {  
    top.location = self.location;  
}  
//-->  
</script>
```

This latter script also deframes when our frame is visited from another external frame. However, for the **indx.htm** documents, there is a better general solution that replaces the link to the frameset- or home-page: Javascripts that analyse how a file like **ind4.htm** is approached, either from the 'internal' frame (**indicatorlist**) or from an 'external' link as in our metadata-document.

Two Javascripts must be written, one for the **frameset.htm** document and another that must be used in all the **indx.htm** documents.

Those scripts are classic tools for professional website developers, and they need only small adaptations to the local situation.

First we present the script **orphan.js** that must become part of the **indx.htm** documents

The content of **orphan.js**:

```
<!--  
  
//orphan script that is inserted in  
//pages which must be exposed in a frame  
//the value of the variable parent_url  
//must be the url of the page containing  
//frameset scriptcode  
  
//so one might need to change "frameset.htm"  
//into something like "http://domain.org/folderx/frameset.htm"  
  
if (top == self) {  
    //URL of IISD-frameset page  
    var parent_url = "frameset.htm";
```

```
//URL of this page
var orphan_url = self.location.href;

//new address for reframing this page
var reframe_url = parent_url + "?" + orphan_url;

//Send user to that address
location.href = reframe_url;
}
// -->
```

This script is inserted in the **indx.htm** documents (in their body):

```
<html>
<head>
<title></title>
</head>
<body>
<script type="text/javascript" src="orphan.js">
</script>
<h1>Indicator 3</h1>
<p>Here comes the description of indicator 3</p>
</body>
</html>
```

So each **indx.htm** document needs only more one line of code.

The **frameset.htm** file needs a complete rebuild:

```
<html>
<head>
<title>Home of a framed site</title>
</head>
<!-- FIRST OUR DEFRAMING SCRIPT -->
<script language="JavaScript" type="text/javascript">
<!--
// deframing from other frames
if (top != self) {
    top.location = self.location;
}
//-->
</script>

<!-- NOW THE ORPHAN-REFRAMING SCRIPT -->

<script type="text/javascript">
<!--
var frame_source;
//Is there a search component to this page's address?
if (location.search) {
//if so extract the string, which is the address of an orphaned page
frame_source = location.search.substring(1);
}
else {
//otherwise, enter the filename of the default page
frame_source = "default.htm";
}
//write the frameset code
document.write('<frameset cols="30%,70%">');
document.write('<frame src="indicatorlist.htm">');
document.write('<frame SRC="' + frame_source + '" NAME="main">');
```

```
document.write('</frameset>');  
// -->  
</script>  
</html>
```

We are not going to explain how these scripts actually work (some hints are given in the //script-comments). Main point is that if someone uses a metadata-identifier like 'http://organisation.org/path/ind3.htm', he/she will not only see the ind3.htm file, but he/she will see it in the frame, in it's context. And this will save a lot of frustration.

Many people are often frustrated by websites with frames and they are right, because often the frames are build from an isolated or stand-alone viewpoint and not as a part of a larger Word Wide Web. Of course, writing Javascripts can be frustrating too :-)

Make sure you're webmaster is in a happy mood before you ask for it.

References:

McFedries (2001), *Special Edition, Using JavaScript*, Que
Paul McFedries' site: <http://www.mcfedries.com/JavaScript/>

IX. A simple DC-editor using client side Javascript

This metadata-editor (only a small Dublin Core) is an HTML-form that can be used with an Internet Explorer (Microsoft).

Users of editor type content in text fields and convert this content to xml- and/or xhtml-file.

This editor is for demonstrative purposes only, and is not meant to be perfect.

- Works with client side javascript
- Uses 'save snapshot' function of IE-explorer
- Not platform-independent: only for use on MS-Windows with IE-Explorer

The user can select a background color (just to demonstrate that some enduser-preferences can be provided)

The user is asked to give a name tot the file (DC-record) that he produces. This filename is stored in the content of the XML and/or XHTML -file.

The user can interrupt his job and save his draft-file with the 'save-as' function in the file menu of the browser (a name must be given to his draft snapshot-file and it should be saved as a html-file with htm- or html-extension).

If the enduser wants to continue to work on the file, the snapshot file that is saved should be opened with the browser, and the user will find his draft text in the txt-fields of the form.

This only works in Internet Explorer. It uses a MS-specific save-snapshot property in the style sheet.

The content of the form-fields can be converted to an XML-document and an XHTML-document and the XHTML-output can be viewed on screen.

Special characters can be converted to the standard XHTML-codes.

The user should cut and past the final source code from the final text-field into a separate text-editor (like notepad) and save it as a file with xml- or html-extension.

During the project we have tried many versions based on client side javascripts that could be used offline. The main problem remained: if one tries tot produce something hat works perfect in Internet Explorer of MS-Windows, it will probably not function with another browser.

Finally we've concluded that such tools can only be used locally, that is within organisations that share the same platform. If one wants to work in cooperation with other organisations a system with server side scripts should be preferred (using for example an open source architecture with **A**pache (server), **M**ySQL (database) and **P**HP or **P**erl (scripts) – the AMP-architecture, described in this report.

One should notice however that client side scripts could mean: only for use when online.

An option is that all users install free AMP-software on their desktop. Cross-platform cooperation is possible, but it is not so easy to convince newbies that AMP-software is easy to install. Simply because it is not 'easy' if you're a beginner.

Finally we came up with an easy trick: Word-documents with hidden codes, described earlier in this report. It is of course a procedure developed for people that work with an MS-Windows platform (as most people in Belgium do). But it is a procedure that functions very well and allows exchange of draft-data by email for co-editing (see elsewhere in this report).

We did not look for an equivalent procedure that can be used with office-tools (word-processors) on a Linux platform. On Linux however, many much more options are available because of the open source philosophy and the fact that most Linux platforms will have AMP-software installed.

Part of DC-editor as presented on canvas of browser (without buttons)

Select background color dublincore.org

Text fields to insert matadata

DC Title

this is a title

DC Identifier

some text w ith special character é

...etcetera...

Output field from where xml- and or xhtml-file is produced from input.
This result should be copied into a text-editor like notepad and saved there as an xml- or html file.

```
<dc thisfile='name of file.xml'>
<title>this is a title</title>
<identifier>some text w ith special character é</identifier>
<description></description>
<subject></subject>
<creator></creator>
<publisher></publisher>
<contributor></contributor>
<date></date>
<rights></rights>
<type></type>
<format></format>
<source></source>
<language></language>
<relation></relation>
<coverage></coverage>
</dc>
```

Source code of the editor :

```

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3c.org/TR/1999/REC-html401-19991224/loose.dtd">
<HTML>
<HEAD>
<TITLE>test savesnapshot</TITLE>
<META http-equiv=Content-Type content="text/html; charset=windows-1252">
<META content=snapshot name=save>

<STYLE type=text/css>

H3 {MARGIN-BOTTOM: 0em; COLOR: black; FONT-FAMILY: verdana}
P {FONT-SIZE: 12pt; FONT-FAMILY: arial}
.saveSnapshot {BEHAVIOR: url(#default#saveSnapshot)}

</STYLE>

<SCRIPT language=JavaScript>
<!--

function makeXML(form)
{
form.xmlfield.value = "<dc thisfile='" + form.thisfile.value + ".xml'>\n" +
"<title>" + form.title.value + "</title>\n" +
"<identifier>" + form.identifier.value + "</identifier>\n" +
"<description>" + form.description.value + "</description>\n" +
"<subject>" + form.subject.value + "</subject>\n" +
"<creator>" + form.creator.value + "</creator>\n" +
"<publisher>" + form.publisher.value + "</publisher>\n" +
"<contributor>" + form.contributor.value + "</contributor>\n" +
"<date>" + form.date.value + "</date>\n" +
"<rights>" + form.rights.value + "</rights>\n" +
"<type>" + form.type.value + "</type>\n" +
"<format>" + form.format.value + "</format>\n" +
"<source>" + form.source.value + "</source>\n" +
"<language>" + form.language.value + "</language>\n" +
"<relation>" + form.relation.value + "</relation>\n" +
"<coverage>" + form.coverage.value + "</coverage>\n" +
"</dc>";
}

// character encoding in decimal unicode

function encode(xmlfile)
{
    regexp1 = /%/g;
    xmlfile = xmlfile.replace(regexp1, "%37");
    regexp1 = /â/g;
    xmlfile = xmlfile.replace(regexp1, "%226");
    regexp1 = /à/g;
    xmlfile = xmlfile.replace(regexp1, "%224");
    regexp1 = /á/g;
    xmlfile = xmlfile.replace(regexp1, "%225");
    regexp1 = /ä/g;
    xmlfile = xmlfile.replace(regexp1, "%228");
    regexp1 = /ã/g;
    xmlfile = xmlfile.replace(regexp1, "%227");
}

```

```
    regexpl = /ç/g;
    xmlfile = xmlfile.replace(regexpl, "&#231;");
    regexpl = /ê/g;
    xmlfile = xmlfile.replace(regexpl, "&#234;");
    regexpl = /é/g;
    xmlfile = xmlfile.replace(regexpl, "&#233;");
    regexpl = /è/g;
    xmlfile = xmlfile.replace(regexpl, "&#232;");
    regexpl = /ë/g;
    xmlfile = xmlfile.replace(regexpl, "&#235;");
    regexpl = /ï/g;
    xmlfile = xmlfile.replace(regexpl, "&#239;");
    regexpl = /ñ/g;
    xmlfile = xmlfile.replace(regexpl, "&#241;");
    regexpl = /ó/g;
    xmlfile = xmlfile.replace(regexpl, "&#243;");
    regexpl = /ò/g;
    xmlfile = xmlfile.replace(regexpl, "&#242;");
    regexpl = /ô/g;
    xmlfile = xmlfile.replace(regexpl, "&#244;");
    regexpl = /ù/g;
    xmlfile = xmlfile.replace(regexpl, "&#249;");
    regexpl = /ú/g;
    xmlfile = xmlfile.replace(regexpl, "&#250;");
    regexpl = /û/g;
    xmlfile = xmlfile.replace(regexpl, "&#251;");
    regexpl = /ü/g;
    xmlfile = xmlfile.replace(regexpl, "&#252;");
    regexpl = /' /g;
    xmlfile = xmlfile.replace(regexpl, "&#39;");
    regexpl = /°/g;
    xmlfile = xmlfile.replace(regexpl, "&#176;");

//filename should not be changed
    regexpl = /<dc thisfile=&#39;/;
    xmlfile = xmlfile.replace(regexpl, "<dc thisfile='");
    regexpl = /&#39;>/;
    xmlfile = xmlfile.replace(regexpl, "'>");

    editor.xmlfield.value = xmlfile;
}

// transformation from xml to xhtml

function makeXHTML(xmlfile)
{
    var stylesheet = "h1 {font-family: verdana; color: blue;}\\n div.dc
{font-family: arial; font-size:12pt; margin-left:10%; margin-right:10%;}\\n
span.element {width:20%; background: yellow; font-weight: bold;}\\n
span.content {margin-left:2em;}";

    // title first, otherwise HTML-title tag will be replaced

    regexpl = /<title>/;
    xhtmlfile = xmlfile.replace(regexpl, "<span
class='element'>Title</span><br/><span class='content'>");
    regexpl = /<\/title>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\\n");

    regexpl = /<identifier>/;
```



```

    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Identifier</span><br/><span class='content'>");
    regexpl = /<\//identifier>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<description>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Description</span><br/><span class='content'>");
    regexpl = /<\//description>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<subject>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Subject</span><br/><span class='content'>");
    regexpl = /<\//subject>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<creator>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Creator</span><br/><span class='content'>");
    regexpl = /<\//creator>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<publisher>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Publisher</span><br/><span class='content'>");
    regexpl = /<\//publisher>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<contributor>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Contributor</span><br/><span class='content'>");
    regexpl = /<\//contributor>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<date>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Date</span><br/><span class='content'>");
    regexpl = /<\//date>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<rights>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Rights</span><br/><span class='content'>");
    regexpl = /<\//rights>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<type>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Type</span><br/><span class='content'>");
    regexpl = /<\//type>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<format>/;
    xhtmlfile          =          xhtmlfile.replace(regexpl, "<span
class='element'>Format</span><br/><span class='content'>");
    regexpl = /<\//format>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<source>/;

```

```

    xhtmlfile = xhtmlfile.replace(regexpl, "<span
class='element'>Source</span><br/><span class='content'>");
    regexpl = /<\/source>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<language>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "<span
class='element'>Language</span><br/><span class='content'>");
    regexpl = /<\/language>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<relation>/;
xhtmlfile = xhtmlfile.replace(regexpl, "<span
class='element'>Relation</span><br/><span class='content'>");
    regexpl = /<\/relation>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<coverage>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "<span
class='element'>Coverage</span><br/><span class='content'>");
    regexpl = /<\/coverage>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</span><br/><br/>\n");

    regexpl = /<dc thisfile='\w*\.\w*'>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "<!DOCTYPE html PUBLIC '-
//W3C//DTD XHTML 1.0 Transitional//EN'
'http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd'>\n<html
xmlns='http://www.w3.org/1999/xhtml'>\n<head>\n<title>Dublin
Core</title>\n<style type='text/css'>\n" + stylesheet +
"</style>\n</head>\n <body>\n <h1>Dublin Core Metadata</h1>\n<div
class='dc'>\n");

    regexpl = /<\/dc>/;
    xhtmlfile = xhtmlfile.replace(regexpl, "</div>\n</body>\n</html>");

// results in xmlfield
    editor.xmlfield.value = xhtmlfile;

}

function previewXHTML()
{
    previewWindow = window.open();
    previewWindow.document.write ("DO NOT SAVE FILE IN THIS BROWSER WINDOW");
    previewWindow.document.write (xhtmlfile);
}

// -->
</SCRIPT>

<!--=====FORM STARTS HERE=== -->

<FORM name=editor action="">
<META content="MSHTML 6.00.2715.400" name=GENERATOR>
</HEAD>

<BODY>
<H1>Dublin Core editor</H1>

<P>Select background color
<SELECT onchange=document.bgColor=this.options[this.selectedIndex].value>

```

```

<OPTION value=40E0D0>Torquoise
<OPTION value=2E8B57>Sea Green<OPTION value=87CEEB>Sky   Blue
<OPTION value=F4A460>Sandy Brown<OPTION value=FFF0F5>Lavender Blush
<OPTION value=FFFFFF selected>White</OPTION>
</SELECT>
<A href="http://dublincore.org/">dublincore.org</A>
<HR>

<!--===== -->

<H3>Filename</H3>

<P>letters (a-z), digits (0-9) and underscore (_) only</P>
<INPUT onclick="alert('Filename: Name of the file where these Dublin Core
Metadata will be stored. Do not type extension (.xml, .html)')" type=button
value=Help>
<TEXTAREA class=saveSnapshot id=thisfile rows=1 cols=40></TEXTAREA>
<HR>

<!-- =====DUBLIN CORE =====>

<P>Following characters can be used:</P>
<P>from a to z, from A to Z, ? , . , % , ° , ' , â , à , á , ä , ã , ç , ê , é , è , ë ,
ï , ñ , ô , ó , ò , ü , û , ú , ù</P>

<H3>DC Title</H3>
<INPUT onclick="alert('Title: A name given to the resource. Typically, a
title will be a name by which the resource is formally known.')"
type=button value=Help>
<TEXTAREA class=saveSnapshot id=title cols=80></TEXTAREA>

<H3>DC Identifier</H3>
<INPUT onclick="alert('Identifier: An unambiguous reference to the resource
within a given context. Recommended best practice is to identify the
resource by means of a string or number conforming to a formal
identification system. Example formal identification systems include the
Uniform Resource Identifier - URI - including the Uniform Resource Locator
- URL, the Digital Object Identifier -DOI- and the International Standard
Book Number -ISBN.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=identifier cols=80></TEXTAREA>

<H3>DC Description</H3>
<INPUT onclick="alert('Description: An account of the content of the
resource. Description may include but is not limited to: an abstract, table
of contents, reference to a graphical representation of content or a free-
text account of the content.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=description cols=80></TEXTAREA>

<H3>DC Subject</H3>
<INPUT onclick="alert('Subject: The topic of the content of the resource.
Typically, a Subject will be expressed as keywords, key phrases or
classification codes that describe a topic of the resource. Recommended
best practice is to select a value from a controlled vocabulary or formal
classification scheme.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=subject cols=80></TEXTAREA>

<H3>DC Creator</H3>
<INPUT onclick="alert('Creator: An entity primarily responsible for making
the content of the resource. Examples of a Creator include a person, an
organisation, or a service. Typically, the name of a Creator should be used
to indicate the entity.')" type=button value=Help>

```

```
<TEXTAREA class=saveSnapshot id=creator cols=80></TEXTAREA>

<H3>DC Publisher</H3>
<INPUT onclick="alert('Publisher: An entity responsible for making the
resource available. Examples of a Publisher include a person, an
organisation, or a service. Typically, the name of a publisher should be
used to indicate the entity.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=publisher cols=80></TEXTAREA>

<H3>DC Contributor</H3>
<INPUT onclick="alert('Contributor: An entity responsible for making
contributions to the content of the resource. Examples of a Contributor
include a person, an organisation, or a service. Typically, the name of a
Contributor should be used to indicate the entity.')" type=button
value=Help>
<TEXTAREA class=saveSnapshot id=contributor cols=80></TEXTAREA>

<H3>DC Date</H3>
<INPUT onclick="alert('Date: A date associated with an event in the life
cycle of the resource. Typically, Date will be associated with the creation
or availability of the resource. Recommended best practice for encoding the
Date value is defined in a profile of ISO 8601 and follows the YYYY-MM-DD
format.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=date cols=80></TEXTAREA>

<H3>DC Rights</H3>
<INPUT onclick="alert('Rights: Information about rights held in and over
the resource. Typically, a rights element will contain a rights management
statement for the resource, or reference a service providing such
information. Rights information often encompasses Intellectual Property
Rights -IPR, Copyright, and various Property Rights. If the Rights element
is absent, no assumptions can be made about the status of these and other
rights with respect to the resource.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=rights cols=80></TEXTAREA>

<H3>DC Type</H3>
<INPUT onclick="alert('Type: The nature or genre of the content of the
resource. Type includes terms describing general categories, functions,
genres, or aggregation levels for content. Recommended best practice is to
select a value from a controlled vocabulary (for example, the working draft
list of Dublin Core Types. To describe the physical or digital
manifestation of the resource, use the Format element.')" type=button
value=Help>
<TEXTAREA class=saveSnapshot id=type cols=80></TEXTAREA>

<H3>DC Format</H3>
<INPUT onclick="alert('Format: The physical or digital manifestation of the
resource. Typically, Format may include the media-type or dimensions of the
resource. Format may be used to determine the software, hardware or other
equipment needed to display or operate the resource. Examples of dimensions
include size and duration. Recommended best practice is to select a value
from a controlled vocabulary (for example, the list of Internet Media Types
-MIME- defining computer media formats.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=format cols=80></TEXTAREA>
<H3>DC Source</H3>
<INPUT onclick="alert('Source: A Reference to a resource from which the
present resource is derived. The present resource may be derived from the
Source resource in whole or in part. Recommended best practice is to
reference the resource by means of a string or number conforming to a
formal identification system.')" type=button value=Help>
<TEXTAREA class=saveSnapshot id=source cols=80></TEXTAREA>
```

```

<H3>DC Language</H3>
<INPUT onclick="alert('Language: A language of the intellectual content of
the resource. Recommended best practice for the values of the Language
element is defined by RFC 1766 which includes a two-letter Language Code
(taken from the ISO 639 standard), followed optionally, by a two-letter
Country Code (taken from the ISO 3166 standard. For example, en for
English, fr for French, or en-uk for English used in the United Kingdom.')"
type=button value=Help>
<TEXTAREA class=saveSnapshot id=language cols=80></TEXTAREA>

<H3>DC Relation</H3>
<INPUT onclick="alert('Relation: A reference to a related resource.
Recommended best practice is to reference the resource by means of a string
or number conforming to a formal identification system.')" type=button
value=Help>
<TEXTAREA class=saveSnapshot id=relation cols=80></TEXTAREA>

<H3>DC Coverage</H3>
<INPUT onclick="alert('Coverage: The extent or scope of the content of the
resource. Coverage will typically include spatial location (a place name or
geographic coordinates), temporal period (a period label, date, or date
range) or jurisdiction (such as a named administrative entity). Recommended
best practice is to select a value from a controlled vocabulary (for
example, the Thesaurus of Geographic Names - TGN) and that, where
appropriate, named places or time periods be used in preference to numeric
identifiers such as sets of coordinates or date ranges.')" type=button
value=Help>
<TEXTAREA class=saveSnapshot id=coverage cols=80></TEXTAREA>
<HR>

<!-- ===== -->

<H2>XML and XHTML</H2>
<P>Click each button only once</P>

<BUTTON onclick=makeXML(this.form)>Make XML</BUTTON>&nbsp;
<BUTTON onclick=encode(this.form.xmlfield.value)>Encoding</BUTTON>&nbsp;
<BUTTON onclick=makeXHTML(this.form.xmlfield.value) name=xhtml>Make
  XHTML</BUTTON>&nbsp;
<BUTTON onclick=previewXHTML()>Preview XHTML</BUTTON>
<BR><BR>
<TEXTAREA class=saveSnapshot name=xmlfield rows=20 cols=80></TEXTAREA><BR>
</FORM>
</BODY>
</HTML>

```

X. ISSD-metadata for indicators

This chapter contains some major conclusions concerning special metadata that could be used to describe indicators and collections of indicators.

It's only a first draft, that might be further developed. Please visit our website at:

<http://cdonet.UGent.be/issd/index.html> or <http://purl.oclc.org/NET/cdo>.

The Information System for Sustainable Development (ISSD) of the OSTC (<http://www.belspo.be/issd>) has already produced metadata for projects, organisations, publications and persons.

This is a system that is centralised: a database presents online forms, which have to be filled in by contributors of metadata. In this report we have looked for another complementary approach using open standards – XHTML- and XML-documents – which can be produced on the desktop and can be uploaded to a centralised archive for metadata-extraction.

We've re-evaluated the Dublin Core Metadata, which has become a de facto standard on the World Wide Web (a situation quite differently from that when the ISSD started).

In our opinion, when specific metadata for ISSD are produced, they should be complementary to the Dublin Core.

While Dublin Core metadata use the prefix DC, we will use the prefix ISSD to differentiate the specific ISSD-metadata from others.

One basic ISSD-metadata-element will be

issd.type

issd.type can have six values:

project, person, organisation, publication, indicator and indicatorset

If someone describes an indicator with Dublin Core metadata, for example:

DC.Title: Mean energy use of households

And adds following element to the DC-elements:

ISSD.type: indicator

We will know that the DC.Title is the title of a document that describes an indicator. Generally this title will be the name of the indicator.

X.1. Metadata for an indicator: further refinements

An indicator is a measurable parameter that gives relevant information about a certain aspect that asks for our attention. We measure something in order to make a judgement about something we are concerned about (see chapter II). Often, a description of an indicator therefore generally contains two major parts:

- description of what is being measured (observation)
- description of the concern or aspect we which to assess with the measurement

For example: mean family income (observation) could be related to welfare (concern).

Often the term 'goal' is used instead of concern. But measurements can give an indication for symptoms or problems also. Concerns can be expressed in positive or negative terms.

We'll use the term 'concern' as the abstract object for anything that is judged or observed by means of another object: the observation.

Both can have a name and a description, and for both – concern and observation- targets could be set. Therefore at least six metadata-elements could be considered:

Issd.isd.concern.name:

Example: poverty

Issd.isd.concern.description:

Our definition/description of poverty (a list of keywords is allowed)

Issd.isd.concern.target:

'In 2010 there are no more poor people in our country'

Issd.isd.observation.name:

Example: percentage of families with income lower than 400 Euro/month

Issd.isd.observation.description:

Description of measurement

Issd.isd.observation.target

Example: 'Zero percent'

Some more details about the observation can be added with:

Issd.isd.observation.unit: example kilowatthour/capita

Issd.isd.observation.method: who observes, when, how often

Issd.isd.observation.source: a link to primary data-resources

We could add a lot of other metadata-elements to the list, but in the end we've decided not to do so, while it could be asking too much of those that should produce metadata. It happens that fifty to hundred, or even more, indicators are produced.

If we know what has been measured (observed) and if we know which concerns are being assessed, we might already have sufficient information to support a good search facility in an archive.

The Dublin Core can be used to complete the metadata. For example:
dc.identifier, dc.format, dc.type, dc.language, etcetera.

If the description of measurement and the description of a 'concern' are stored in separate data-files, one could produce two metadata-records, one with `issd.isd.observation.name` and `issd.isd.observation.description`,... another with `issd.isd.concern.name` and `issd.isd.concern.description`,....

X.2 Metadata for a collection (set) of indicators

Following specific metadata-elements can be added to a Dublin Core:

Issd.isdset.methodology: Includes some information about methodology used to develop the indicator set.

Issd.isdset.participation: description of people or organisations involved in the development of the indicator collection.

Issd.isdset.frame: description of the general structure of the metadata-collection. For example: the DPSIR-frame could be used. One could also describe the fact that some decision support system is used (multi-criteria-evaluation, expert system, ...) or logframes or SWOT for example.

Issd.isdset.frequency: are indicators updated regularly?

Issd.isdset.adminstration: who is responsible for the management of the data (indicators)

Issd.isdset.status: is it a draft, a finished or obsolete set of indicators? Is the set replaced by a new version?

Again these metadata-elements could (should) be used in combination with Dublin Core

X.3 Presentation of Indicators for Sustainable Development

One of the issues raised at the start of this project concerns the presentation of Indicators for Sustainable Development. If one visits ISD-sites on the web, soon it becomes clear that it is hard to find two sites that look the same.

When the presentation of information is discussed, often a distinction is made between:

- **content;**
- **structure, and**
- **style**

These 3 issues cannot be separated strictly, but it helps to focus on these aspects when presentation is discussed. In many books about website-development or multimedia-development the 3 issues are considered and it is often advised first to think about content, then structure and finally style.

In any case, one should have the 'users' in mind when developing the presentation. Information should be presented in a user-friendly way. At the Centre for Sustainable Development, we think that one special kind of users should be considered, and that is blind or disabled people (see chapter 5 for Dutch-speaking readers)

Sometimes a fourth aspect is added: **format**. PDF (portable document file) is a format, HTML- and XML -markup are standard formats used on the web. The same information can be presented in many formats.

In fact, the markup languages used on the web are formats that allow you to manage content, structure and style separately, which means that the same content -stored in XML- or (X)HTML- files- can be presented in many different structures or styles, without touching the content.

We are not going to dig into the literature of website- or multimedia-development, where more and sometimes other terms are used to specify certain aspects of presentation.

We consider the terms content, structure, style and format as sufficient, to discuss some metadata-options for the description of the presentation.

Content

Focussing on **content** means: focussing on the 'message' . What do we want to say? Another related basic question is: Who is going to - or should- read it? What is the target group and what do we want to achieve with the information? Some readers will ask: isn't the content already defined here? The content is 'Indicators for Sustainable Development' (ISD), isn't it.

Concerning the presentation of indicators for sustainable development the content can indeed be kept quite simple:

- a list of indicators, combined with
- one separate document for each indicator

ISD-collections can be regarded as an isolated 'thing', but generally they are not isolated at all. ISD-collections are a tool, an instrument and therefore embedded in a larger context where they must function.

The content of the presentation can grow in many ways. We will sum up the main additions that can be found on many websites that present indicators for sustainable development:

Focus on the instrumental role of indicators

Some sites focus on the role of Indicators for Sustainable Development. The main functions attached to this instrument are:

- supporting planning and decision making (ex-post and ex-ante-evaluation, early warning)
- supporting the debate about sustainable development (indicators might help to define the meaning of sustainability)

Both are of course strongly linked with each other.

(Overall) Evaluation and policy-responses

Often for each indicator an evaluation is made, following the latest observations and analysing the trends over the last years.

But the separate observations that are made can also become subject of an overall evaluation. This evaluation can be a text, a conclusion that is written down in words and sentences.

If indicators are instrumental, it becomes obvious to present them in a larger context of policy-making and -implementation. This might lead to the construction of an information-system where the indicators are linked to policy-domains (decisions and actions that respond to the scores and evaluation of the indicators).

Evaluation can be followed by (optional) policy-responses, which might be based on one or more observations.

Critical view on indicators, aggregation and ranking

Most experts agree that collections of Indicators for Sustainable Development are the result of 'subjective' choices. The indicators are representing a certain view on the subject 'sustainability', an interpretation or meaning given by a group of people. Although this issue is often raised in the scientific reports about ISD, it is not often mentioned in the final presentations for the broad audience. A critical note on the subject 'indicator' could however be very informative. The reader could be invited to form his opinion about the subject, for instance via confrontation with other sets of indicators made by other groups of people. The ISD then become instrumental in education about sustainable development. If for example a city presents ISD to evaluate its policy, local schools could be invited to develop their ISD, to evaluate the sustainability of the school.

Some developers of ISD-sets produce an aggregated score, and use this score to focus attention of the media on the indicator-project.

Aggregation however, always includes 'simplification' and the aggregation-method itself can become subject of discussion, especially when the overall score is used in a ranking where -for example - countries are ranked according to their aggregated score.

The result depends on the selected indicators, the data that are used and the aggregation-method. If a country scores badly, one can be almost certain that those (politicians) who are 'responsible' for the result, will raise questions about the methods that are applied.

A public debate about the interpretation of indicators and aggregated scores can have a positive effect in a sense that the broad public 'learns' to be critical about statistics. But at the same time the final conclusion made by this broad public can be negative also. The fact that subjective choices are made, might be considered as a proof that the result has no 'scientific' value. Many people think that 'subjective' choices are of less value than objective information. Which is of course not true.

The main problem is that a set of indicators includes objective information and subjective information (choices based on human perception, interpretation of sustainability, values, preferences,...). Subjective information is not of lower quality. But this means that one should explain why certain choices are made. The reader must understand the 'logic' behind the whole process of indicator-selection, and then he/she can decide whether he/she agrees with this logic.

In many cases the indicator-collections have a hierarchical structure, where a series of steps is made between the overall concern 'sustainability' and the final measurable aspects of sustainability, the indicators.

This process is well known in multi-criteria evaluation where a complex and vague property like sustainability is split up in less complex sub-criteria.

One popular first step that is made is a split up of sustainability in ecological, economical and social aspects. This 'logic' behind this first step could be summarised as follows:

Y is sustainable

- IF the social performance of Y is good AND
- IF the environmental performance of Y is good AND
- IF the economical performance of Y is good

This implication can be understood by most people. It is a sentence that fits perfect in common daily language. Such formal constructions are used in expert systems, where knowledge is stored in terms of :

- facts (X is good, Y is bad), and
- implications (Y is good IF X is good AND Y is good)

It is somehow 'strange' that this formal technique is not often used in indicator-presentations, while it is flexible, simple and at the same time very effective. To be more precise: many reports that explain why certain indicators are selected present a certain 'reasoning'-proces, but explicit use of formal logic is rare.

One site that presents the logic behind it's selected indicators - in a brief way - can be found here:

<http://www.ee/baltic21/indicators/>

The whole reasoning-process that lies behind the selection of certain indicators can be summarized in a series of formal implications that supports the presentation of in indicators. The reader then has a quick overview of all the steps in the reasoning process and can decide for each implication if he/she agrees or not.

Example

Many websites that publish Indicators for Sustainable Development (ISD), do not only present the ISD but other information too. For example, the ISD can be placed in an educational context, or can be used as a tool to focus attention on policy-making. The set of ISD can be presented as a subject for debate, or a call for participation -in decision-making or in action. In all these cases the ISD are considered 'instrumental', a tool that can have one or more functions.

As an example we focus on a design of a website that includes a set of indicators (Sustainability Barometer) for a city. This design was used when we (Centre for Sustainable Development, University of Ghent) discussed the presentation-options with our client (administration of the city of Ghent Belgium).

It was the ambition to use the Sustainability Barometer of Ghent, not only as an evaluation-tool, but also as an instrument that:

- focuses the citizens attention on the local policy (local Agenda 21),
- can be used as a starting point for debate about sustainable development (inviting people to participate in the local debate)
- as an entrance for the discovery of many initiatives in other cities, or initiatives on a higher institutional level (national, European Union, OECD, United Nations,...)

The ambitions are therefore not only to present an evaluation tool but also to enhance participation in policy-making and to support education in this field.

The Indicators for Sustainable Development (the barometer) were explicitly presented as an instrument, that is part of a larger program, a program that is called - in this case: Local Agenda 21.

The 'homepage' of the Local Agenda was a sitemap with a link to the barometer, but also links to related subjects. On top of the list had a starting point four a brief 'tour d' horizon', that is a sequence of small texts which could be read as an introduction, explaining what the site has to offer. The visitor should be able to surf through this 'tour d' horizon' in about 15 minutes, which means that the text should indeed be brief. They describe the following items:

- Sustainable Development
- Agenda 21
- Local Agenda 21
- Indicators for Sustainable Development, with examples
- What is the function of ISD
- How can citizens participate, concerning ISD and/or the Local Agenda
- Options to develop or use ISD at school, at work or at home.

After that, the visitor could decide whether he/she wants to read more about Sustainable Development, (Local) Agenda 21, or about ISD and the local sustainability barometer.

Five ambitions can be discussed:

- to educate
- to place local activities in a national and international context
- to call for participation in action, and in decision making (including selection of indicators)
- to promote the use of ISD in organisations (school, enterprise...)
- to focus attention on (local) policy

Structure

Choosing a **structure** is like asking how the information will be split up in 'presentation-units' (pages, sections,...) and how these units will be ordered or linked with each other.

Concerning ISD the structure is mostly kept simple. Each indicator is described in a document which looks like a record of database. The collection of indicators is presented in a list, a hierarchical (tree-like) structure or a table.

So far, the structure is simple. But when the presentation includes more context (education, links with policy-domains, etc...), the structure can become complicated.

It is of course a matter of taste but many experienced website-builders do prefer structures that resemble other structures which are common in computer-practice: a structure that is hierarchical or

tree-like structure. The visitor of a website is then confronted with a structure that he/she has seen many times before. Well-known hierarchical structures are the file-management systems on desktop-computers and the bookmark-lists that can be build with the browser-software. In both cases a hierarchical structure of folders (directories and sub-directories) is build, folders that can contain an unlimited numbered of (ordered) files.

The structure of a book is also hierarchical, starting with tree main parts: front, middle and end. The first part can include a content list, a summary or introduction, acknowledgement, etc. The middle part can be split up in chapters and sections. The last part can include an index, a list of acronyms, etc...

The hierarchy looks like this:

- book
 - front
 - content list
 - summary
 - acknowledgements
 - copyright notes
 - middle
 - chapter 1
 - chapter 2
 - chapter 3
 - end
 - footnotes
 - index
 - acronyms
 - authors

Footnotes can be placed on the bottom of each page, at the end of a chapter or in the 'last part'.

Although there is some variation in the structure of books, one could speak of a certain standardisation. No reader will look for a content-list in the middle of a book. If one presents information on the internet, one can decide to follow this tradition, in order to make it easy for the website-visitor.

Good examples of this practice can be found in electronic publications or software-tutorials or programming- and mark-up manuals. See for example:

<http://www.w3.org/TR/html4/>

<http://www.python.org/doc/current/tut/tut.html>

The differences with a book is that on top or bottom of the pages, some links are available to select the next or previous section, or to consult the table of contents (toc) or an index.

Closely related are the so called **help-tools** that are distributed with software. Some people have indeed been thinking about the good structures for the presentation of documentation that must help the software-user. And the conclusion seems to be a preference for hierarchical (tree-like) structures.

More information on this subject can be found here:

HTML Help Start Page at Microsoft:

<http://msdn.microsoft.com/library/default.asp?url=/library/en-us/htmlhelp/html/vsconHH1Start.asp>

<http://www.winwriters.com/restools.htm>

<http://java.sun.com/products/javahelp/>

When this structure is used on websites one often uses 'frames'. When ISD are presented: the list of indicators can be shown in the left pane or frame, while a selected indicator can be shown in the right pane- or frame.

Framing can however become a problem when metadata-archives are made with links to documents in a framed information-structures. For those webmasters that prefer to use frames for the presentation of their ISD and related context, see also section 9.3 about frames.

Style

Style includes the selection of colours, font-types, the use of logo's, etcetera. Style helps in focussing attention on certain aspects or keywords, it can support the structure (different sections can have different background colours). Style is an important factor, if one wants to keep the attention of the reader.

Reading through a set of Indicators for Sustainable Development takes time. So it is not only a question of focussing the attention on the indicators, but also to keep it there. Concerning websites one should keep in mind that a visitor can leave the site with one mouse-click. Pages which need more time for downloading (including many pictures), pages which are difficult to read, or which produce bad prints, can be arguments to leave the site.

Style is however a matter of taste and we prefer not to discuss it in detail here. In one of the following sections we will focus on the Web Accessibility Initiative (WAI). This initiative focuses on the accessibility of information for disabled people. Taking the guiding lines of WAI into consideration, can have a huge impact on style and structure of the presentation.

Metadata about presentation

Concerning the presentation of ISD we came up with the following lists of optional metadata-elements that focus on this issue.

The first list is quite detailed and includes ten metadata-elements.

Metadata for presentation-content:

Five optional metadata-elements: one to describe the audience (target groups that should read the message), one to describe educational ambitions, one to describe links with decision-making and policy-responses, and two to focus on participation: participation in the debate on indicators and sustainable development, another about participation in response (action).

[issd.isdset.audience] Audience: What is the audience - the target- group that should be reached with the ISD?

[issd.isdset.education] Education: Are ISD used as an educational tool, to learn more about sustainable development? Do you teach about ISD and their functions? Are references made to other ISD-sets? Are readers invited to develop their own ISD, for example on enterprise-level?

[issd.isdset.response] Decision-making and Policy-response: Are the indicators used in decision-making processes: ex-ante-evaluation of policy-options. Are the indicators linked to policy-responses or actions that should have an impact on the scores that are observed.

[issd.isdset.debate] Participation in debate: Is there an explicit call for participation in the debate about sustainable development and the selection of ISD?

[issd.isdset.action] Participation in action: Are ISD used to call for action?

Metadata for presentation-structure:

An optional metadata-element: one that allows to give a general description on data-management and presentation, and a second one which focusses attention on the use of structures (analytical frames) that are used in Decision Support Systems or management.

[issd.isdset.structure] Structure-Description: Brief description of the structure. Do you use a database or XML to build a dynamic site? Does it look like a book? Is it a tree-like (hierarchical) structure? Can the ISD be classified in different structures: for example a thematic classification and a sector classification?

Style

Again two optional metadata-element, one for general description of the style, and a second one that focuses attention on Accessibility.

[issd.isdset.style] Style-Description: Did you use a style that is adapted to a certain audience (children)? Did you use graphical presentations like a dashboard from a car? Is it a multi-media presentation (sound, video, ...) .

[issd.wai] Web Accessibility Initiative: Are style (and structure) based on WAI-guidelines?

The second list is simple and has only four elements:

[issd.isdset.context] Context: Description of the context of the presentation of the ISD. Are ISD used as an educational tool, to learn more about sustainable development? Do you teach about ISD and their functions? Are references made to other ISD-sets? Are readers invited to develop their own ISD, for example on enterprise-level?

Are the indicators used in decision-making processes: ex-ante-evaluation of policy-options. Are the indicators linked to policy-responses or actions that should have an impact on the scores that are observed.

Is there an explicit call for participation in the debate about sustainable development and the selection of ISD? Are ISD used to call for action?

[issd.isdset.structure] Structure: Description of the structure of the information. Do you use a database or XML to build a dynamic site? Does it look like a book? Is it a tree-like (hierarchical) structure? Are the ISD classified in more than one structure: for example a thematic classification and a sector classification?

Is the structure of the ISD-presentation based on an analytical frame like DPSIR (driving-forces, pressure, state, impact, response), or a decision aids like Multi Criteria Evaluation, Expert System, or tools like SWOT-analysis, value-trees, logframes?

[issd.isdset.style] Style: Description of the style. Did you use a style that is adapted to a certain audience (children) ? Did you use graphical presentations like a dashboard from a car? Is it a multi-media presentation (sound, video,...) . Are style (and structure) based on WAI-guidelines?

[issd.isdset.audience] Audience: What is the audience - the target- group that should be reached with the ISD?

These lists can be used as food for discussion.

The authors of this report concluded that the second list might be sufficient and could be used as a flexible set that gives a lot of freedom to the metadata-author.

On the other hand, one could also decide to focus attention on only a few aspects of the presentation if one prefers to.

Perhaps the label 'content' could be replaced with the term '**context**' or '**presentation-context**', because that is in fact the issue being described here.

XI. Slideshow:

Indicators for Sustainable Development (ISD) metadata: ambitions

Indicators for Sustainable Development (ISD) metadata: ambitions

- Simple/Flexible
 - Local data-needs first
 - Hierarchical data-structures
- Cross-Browser-Functional
- Web Accessibility Initiative
- Dublin Core Oriented
- XML/XHTML
- To fit in broader range of metadata (person, project, publication, indicator, indicator-frame,...)

ISD-metadata

- A. (Re)produce your data in a simple (X)HTML-format and publish on www
 - = Cross Browser Functional
 - = WAI
 - = ready for processing
- B. Map your data onto ISD-metadata
- C. Call robot to harvest your metadata
- D. Consult metadata-archive
- E. Option: produce your own metadata-collection

Part 1: XML

eXtended Markup Language

Example 1: xml-document

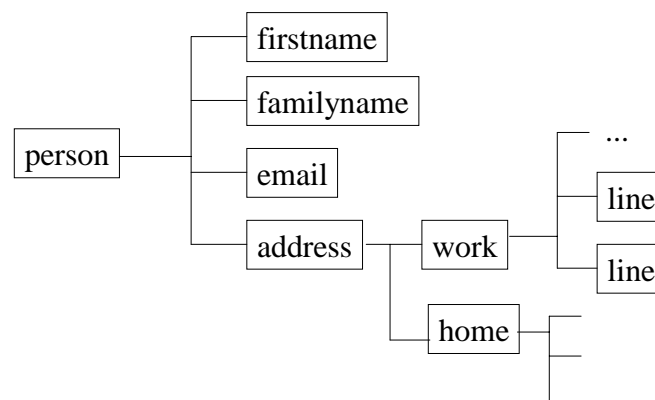
```
<?xml version="1.0"?>
<person>
  <firstname>Stefan</firstname>
  <familyname>Deconinck</familyname>
  <email></email>
  <address>
    <work>
      <line>CDO</line>
      <line>Poel 16</line>
      <line>9000 Gent</line>
      <line>Belgium</line>
    </work>
    <home></home>
  </address>
</person>
```


Element: data (content) + tags



- Element must have start- **and** end-tag
- Start- and end-tag must have the **same name**
- End-tag has **forward slash /**
- Case-sensitive: `<person>` is not equal to `<PersonN>`

Elements are nested in tree-like structure



This is wrong: `<person>...<email>...</person></email>`
 Person- and email-element are not properly nested

Parents and children

- Person is parent of firstname, familyname, email and address
- Work and home are children of address
- Root-element (person):
one element must be parent of all other elements

Document **MUST** start with a **prolog**

Minimum: `<?xml version="1.0"?>`

Meaning: This is an xml-document (xml-version 1.0).
Today, there is only one xml-version

Arguments argumentname="value"

Examples:

```
<email where="work">walter.dejonge@UGent.be</email>  
<date code="yyyy-mm-dd">2001-04-03</date>  
<title lang="dutch">Indicatoren voor Duurzame Ontwikkeling</title>  
<person id="cdo_wdj" version="1">....</person>
```

- Arguments have a name and a value.
- They appear only in start-tag
- Generally value is selected from a fixed set of possible values.
- Arguments support data-management
- Arguments are generally metadata about element

Empty element : to insert special objects

example: horizontal line

```
<hor_line></hor_line>
```

also: <hor_line/>

example: picture

```
<image ref="graphic_15.gif"></image>
```

also: <image ref="graphic_15.gif"/>

Comment: data for authors use only

```
<!-- John, correct this please if wrong -->
```

White space between elements is not relevant

```
<?xml version="1.0"?><person> <firstname>Walter</firstname>
  <familyname>De Jonge</familyname>
<email>walter.dejonge@rug.ac.be</email><address><work>
  <line>CDO</line><line>Poel 16</line><line>9000 Gent</line>
  <line>Belgium</line> </work>
<home></home> </address></person>
```

This xml-document is **well-formed**:

1. There is a prolog
2. There is a root-element
3. All elements are properly nested (tree-structure)
4. Start- and end-tags have same name

Document Type Definition (DTD)

Describes structure of an XML-document

```
<!DOCTYPE person [  
<!ELEMENT person (firstname, familyname, email*, address)>  
<!ELEMENT firstname (#PCDATA)>  
<!ELEMENT familyname(#PCDATA)>  
<!ELEMENT email(#PCDATA)>  
<!ELEMENT address (work, home?)>  
<!ELEMENT work (line*)>  
<!ELEMENT home (line*)>  
<!ELEMENT line (#PCDATA)>  
>
```

? means 'optional' (zero or one)
* means 'zero or more'
+ means 'one or more'

PCDATA =
parsed character data
~ 'text'

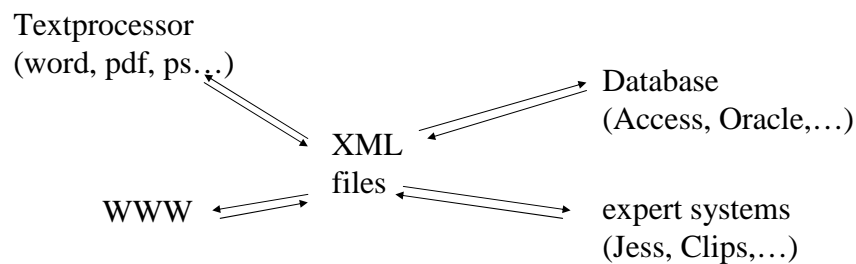
Valid XML-document

XML-document is valid if it has a the structure which is described in definition type document.
DTD is included in XML-document, or there is a reference to an external dtd-file (which can be on another computer)

```
<?xml version="1.0"?>  
<!DOCTYPE person SYSTEM "person.dtd">  
  
<person>  
....  
</person>
```

Why XML ?

- A standard to store and transmit data
- To control data-structure (well-formed, valid ?)
(important in group-work)
- Exchange of data between different applications,
different platforms (Linux, Windows,...), etc.



Advantages of XML

- world wide accepted standard, no licence needed (www.w3c.org)
- platform independent
- no special software needed to read and write an XML-file (notepad is enough). Most people will need only two basic tools: XML-editor (read/write) and XML-browser (datatransfer from servers: http, ftp, email...)
- structure of documents is very explicit
- tagname can have a 'meaning' (<title>This is the title</title>)
- accepted standard for Internet (e-commerce, news-channels, e-books...)
- all programming languages have modules to work with XML (Java, Perl, Python, PHP...)
- all big software-vendors adapt their tools to XML (Microsoft, Sun, IBM, Adobe, ...) and provide transformation-software (from XML to PDF,...)
- open source initiatives work with XML (Apache, Mozilla, Linux, ...)

XML is not an application

- XML does not replace Database Management Systems (DBMS)
- XML facilitates exchange between different DBMS
- XML does not replace Word, Adobe Acrobat (PDF), Latex or other Desktop Publishing-tools
- XML facilitates exchange between formats
- XML does not replace spreadsheet, calculator, filemanagers...
- XML is a format, like pdf or rtf

Example

Situation today:

Infosystem A asks to fill in a form with info about our organisation

Infosystem B asks to fill in another form with same kind of info

Infosystem C has another form

OK, we fill in three forms (copy-paste, copy-paste, ...)

• But if address of our organisation changes ...

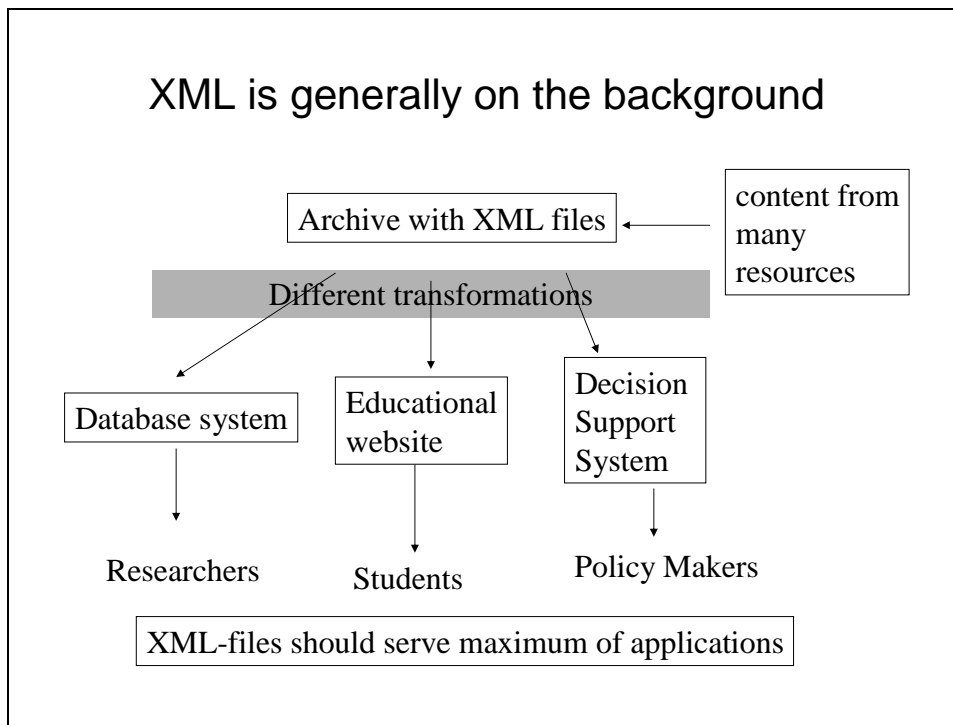
we have to adjust info in three different datatypes

Tomorrow:

• Each organisation makes (and updates) an XML-file with info and mails it to different datatypes who can do with it what they want

Datatypes download latest version regularly

• Or: XML-files are collected in a central repository where all infosystems can collect (download) data they need.



Software-development

- **General:** <http://www.w3.org/XML/>
<http://www.xml.com/>
<http://www.xml.org/>
<http://www.oasis-open.org/>
<http://xml.coverpages.org/>
- **Python:** Special Interest Group for XML-processing
<http://python.org/sigs/xml-sig/>
- **Apache server:** <http://xml.apache.org/>
- **Zope server:** <http://www.zope.org/Wikis/zope-xml/FrontPage>
- **Java:** <http://java.sun.com/xml/>
- **Microsoft:** <http://msdn.microsoft.com/xml/>
<http://www.microsoft.com/net/>
- **IBM:** <http://www-106.ibm.com/developerworks/xml/>
- **Linux:** http://www.stti-usa.com/Linux_XML.htm

Part 2: metadata

Dublin Core Metadata Initiative (DCMI)
<http://dublincore.org/>

Dublin Core

- 15 elements

DC.Title = "Sustainability Monitor Ghent"

DC.Coverage = "Ghent (Belgium); 1999"

DC.Identifier = "http://cdonet.UGent.be/mg/a21/smg.htm"

etcetera

- qualifiers (element refinements): 34 elements in total

DC.Title.Alternative = "SMG"

DC.Coverage.Spatial="Ghent (Belgium)"

DC.Coverage.Temporal="1999"

DC.Relation.IsPartOf = "http://cdonet.UGent.be/mg/a21/index.htm"

etcetera

DC on the www

```

<html>
<head>
<meta name="DC.Title" lang="en"
content="Sustainability Monitor Ghent" />

<meta name="DC.Title" lang="nl"
content="Duurzaamheids Monitor Gent" />

<meta name="DC.Identifier" scheme="URI"
content="http://cdonet.rug.ac.be/mg/a21/sdmg.htm" />

<meta name="DC.Coverage.Temporal" scheme="W3CDTF"
content="1999-03-21" />

<meta name="DC.Subject" scheme="DCMIDCSV" lang="en"
content="ISSD.indicator.scope=city" />
</head>
</body>...</body>
</html>

```

Example: Environment Australia

<http://www.ea.gov.au/>

```

<HEAD>
<TITLE>World Summit on Sustainable Development - Discussion Paper on World Summit on Sustainable
Development (2-12 September 2002)</TITLE>
<LINK REL="schema.AGLS" href="http://www.naa.gov.au/recordkeeping/gov_online/agls/1.2">

<META NAME="DC.Title" LANG="en" CONTENT="World Summit on Sustainable Development - Discussion
Paper on World Summit on Sustainable Development (2-12 September 2002)">
<META NAME="DC.Creator" SCHEME="GOLD" CONTENT="jurisdiction=Commonwealth of Australia;
corporateName=Environment Australia">
<META NAME="DC.Subject" SCHEME="Environment Australia Thesaurus http://www.ea.gov.au/sdd/library/ea-
thesaurus.html" CONTENT="International relations">
<META NAME="DC.Publisher" SCHEME="GOLD" CONTENT="jurisdiction=Commonwealth of Australia;
corporateName=Environment Australia">
<META NAME="DC.Type.documentType" SCHEME="ea-document" CONTENT="Web Page">
<META NAME="DC.Format" SCHEME="IMT" CONTENT="text/html">
<META NAME="DC.Language" SCHEME="RFC1766" CONTENT="en">
<META NAME="DC.Coverage.jurisdiction" LANG="en" CONTENT="Commonwealth">
<META NAME="DC.Coverage.spatial" LANG="en" CONTENT="Australia">
<META NAME="DC.Rights" SCHEME="URI" CONTENT="http://www.ea.gov.au/about/copyright.html">
<META NAME="DC.Date.valid" SCHEME="DCMIPeriod" CONTENT="start=2001-12-20; end=2004-12-20">
<META NAME="DC.Identifier" SCHEME="URI"
CONTENT="http://www.ea.gov.au/cooperation/international/wssd/discussion-paper.html">
<META NAME="DC.Date.created" SCHEME="ISO 8601" CONTENT="2001-09-01">
<META NAME="DC.Date.modified" SCHEME="ISO 8601" CONTENT="2001-12-20">
</HEAD>
(agls = australian government location service)

```

ISSD-DC options

- Combine DC with A-Core (Administrative Core):
metadata about metadata, who produced metadata and when?

```
<meta name="AC.Name" content="Walter De Jonge" />  
<meta name="AC.Email" content="walter.dejonge@UGent.be" />  
<meta name="AC.Contact" content="telephone=+32(0)92648207;  
address=CDO, Poel 16, 9000 Ghent, Belgium" />  
<meta name="AC.Date" scheme="W3CDTF" content="2001-01-19" />
```

- Adding more metadata to satisfy needs of Information System for Sustainable Development (ISSD)

prefix ISSD

```
<meta name="ISSD.audience" content="citizens" />
```

Dublin Core: not domain specific

- DC is used for all kinds of information resources
- Books, websites, databases, CD-rom, ... indicators
- Metadata of indicators can become part of a broader metadatasystem that uses same basic DC-set
- Dublin Core is simple, flexible (extensions possible), not domain-specific, more and more accepted as the basic minimum that each metadata-system should offer
- <http://dublincore.org>

Actors using Dublin Core

- **e- Publishers:**

PRISM: Publishing Requirements for Industry Standard Metadata (Adobe, Reuters, Time, KPMG..., International Press Telecommunications Council)
<http://www.prismstandard.org>

- **Environment:**

Bridges: Minnesota's Gateway to Environmental Information
<http://bridges.state.mn.us/>

DCMI Environment Special Interest Group (EEA??)

<http://www.dublincore.org/groups/environ/>

- **Education**

Gateway to Educational Materials (GEM-US): <http://www.geminfo.org>

<http://www.thegateway.org/>

EdNA (Education Network Australia): <http://standards.edna.edu.au/metadata/>

Information on Education (Germany): <http://www.bildungsserver.de/>

European Schoolnet:

http://www.en.eun.org/eun.org2/eun/en/etb/entry_page.cfm?id_area=14

- **Metadata Observers:**

Metadata Observatory: <http://www.sub.uni-goettingen.de/ssgfi/observatory/>,
<http://www.cenorm.be/iss/Workshop/MMI-DC/Default.htm>

Other observer: <http://www.schemas-forum.org/>

XHTML = HTML and XML

- XHTML is HTML

- Most browsers can interpret it

- XHTML is XML

- Documents can be processed with XSLT, DOM
- Documents can be validated

- XHTML supports transformation to semantic web

- No tabula rasa policy but step by step
- Old HTML and new X(HT)ML can be mixed on one site

- Free robust tools available to transform HTML to 'Tidy HTML' ... to XHTML

- WAI (Web Accessibility Initiative) promotes use of XHTML for disabled people (PARTICIPATION!!!)

Simple XHTML-format

Name of Indicator: Annual Energy Consumption per Capita
Place in DPSIR-frame: Pressure
Facts: 2600 kWh

```
<div>
<span class="issd">resource</span>
<span class="content">indicator</span>
</div>

<div>
<span class="local">Name of indicator</span>
<span class="dc">title</span>
<span class="content">Annual Energy Consumption per Capita</span>
</div>

<div>
<span class="local">Place in DPSIR-frame: </span>
<span class="issd">dpsir</span>
<span class="content">Pressure</span>
</div>

<div>
<span class="local">Facts: </span>
<span class="content">2600 kWh</span>
</div>
```

also metadata for Google e.a.

```
<html>
<head>
<meta name="DC.Title" lang="en" content="Annual Energy Consumption per Capita" />
<meta name="ISSD.resource" lang="en" content="indicator" />
<meta name="ISSD.dpsir" lang="en" content="pressure" />
<meta name="AC.email" content="johny@hollywood" />
</head>
<body>
<div>
<span class="issd">resource</span>
<span class="content">indicator</span>
</div>
<div>
<span class="local">Name of indicator: </span>
<span class="dc">title</span>
<span class="content">Annual Energy Consumption per Capita</span>
</div>
<div>
<span class="local">Place in DPSIR-frame: </span>
<span class="issd">dpsir</span>
<span class="content">Pressure</span>
</div>
<div>
<span class="local">Facts: </span>
<span class="content">2600 kWh</span>
</div>
</body>
</html>
```

Part 3: ISSD-organisation

ISSD = Information System for
Sustainable Development
Metadata Exchange

Input

- Cities, schools, corporations, administrations or other actors make indicator-files and publish them on the web
- If these files are edited in a 'clever' way, then it becomes easy to extract metadata from these files.
- Clever means: XHTML with necessary attributes
- XHTML is a WWW-standard. All browsers can handle XHTML-files.
- XHTML is good for publishing on WWW and good for metadata-exchange at the same time

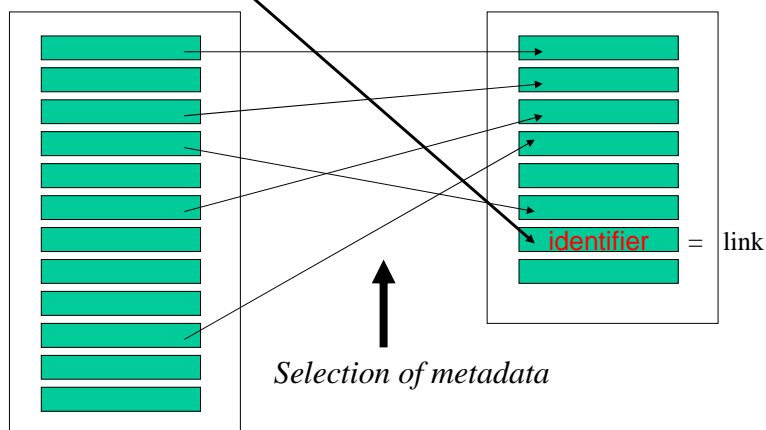
Principles

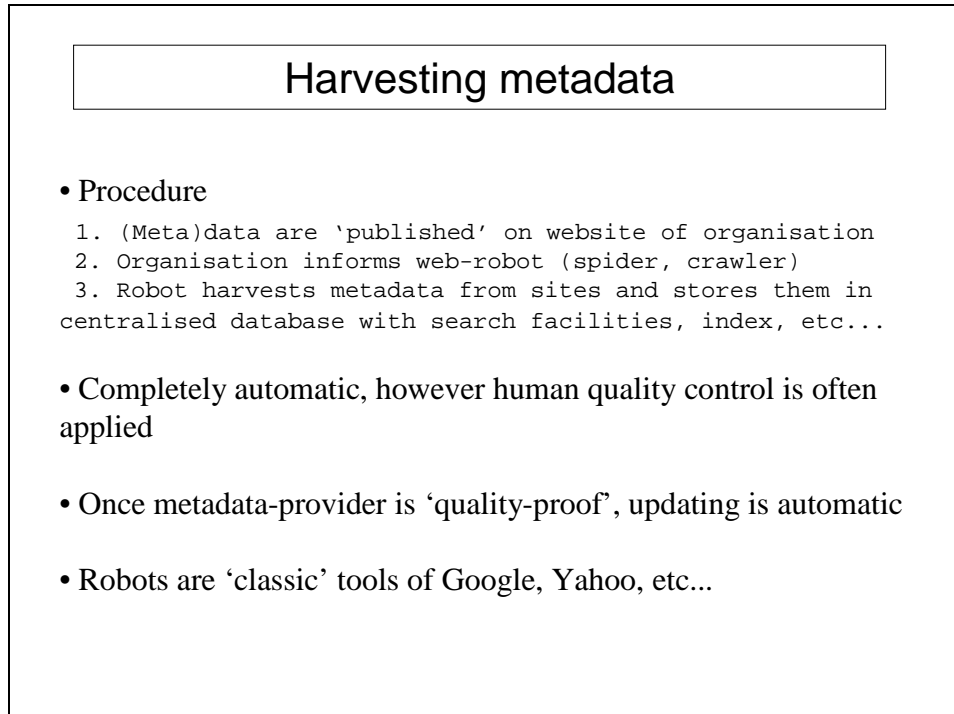
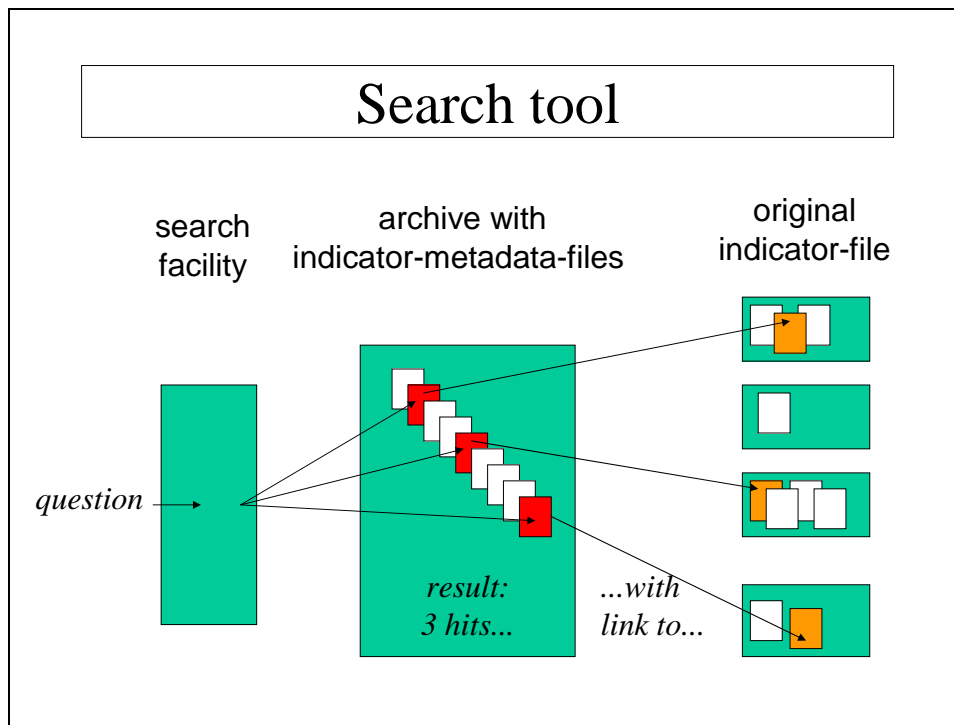
- Each organisation can still develop indicator-documents that suite their specific needs
- Only mapping to ISSD is required. Locally, one can use its own elements ('record-fields').
- Basic ISSD will be kept simple and flexible
- X(HT)ML will be used for exchange of data
- All ISSD Data should be accessible with all browsers (Explorer, Navigator, Opera, Amaya,...)
- Copyright: free use
- WAI: pages can be 'viewed' by blind people

Mapping

INDICATOR DOCUMENT OF ORG. X
at <http://x.org/ind/34.htm>

ISSD-metadata





Instructions for robot

harvest file

```
<?xml version="1.0"?>

<harvest>
<AC.Email>walter.dejonge@rug.ac.be</AC.Email>
<tree>
<folder type="indicatorset">
<metadata>http://cdonet.UGent.be/bg/index.htm</metadata>
<folder type="indicator">
<metadata from="head">http://cdonet.UGent.be/bg/indicator1.htm</metadata>
<metadata from="body">http://cdonet.Ugent.be/bg/indicator2.htm</metadata>
</folder>
</folder>
</tree>
</harvest>
```

Hierarchy or tree-like data-superstructures

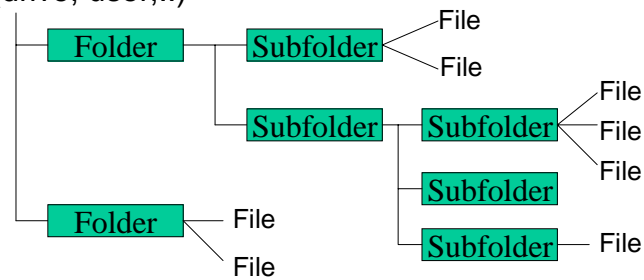
Goal: once we have x(ht)ml-documents (basic structured data about projects, organisations, indicators, etc...), we need a superstructure to organise these documents. This superstructure should give the end-user a good overview of the documents and their main relations with other documents.

Proposal: there are many options to accomplish this, but perhaps the hierarchical or tree-like structure is the 'easiest' one for the broader public. It has been proven to be sufficient in many situations.

Example 1: file-management-systems

All desktop-systems (Linux, Windows,...) have a 'file-manager' to order the files in a tree (hierarchy) of directories and subdirectories (folders and subfolders).

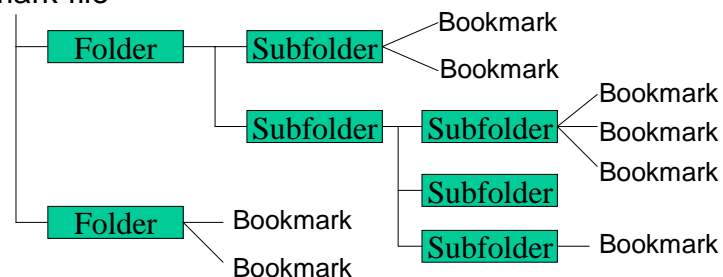
Root (drive, user,..)



Example 2: bookmark-management-systems

All www-browsers have a manager to order the favorite www-addresses (bookmarks, links...) in a tree (hierarchy) of folders and subfolders.

Bookmark-file

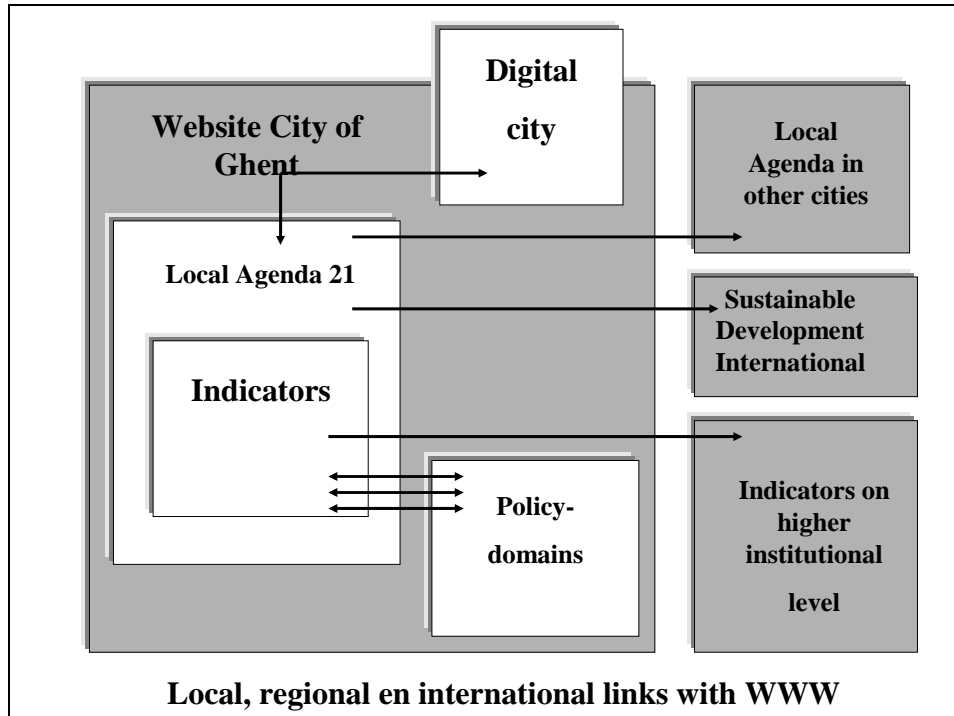
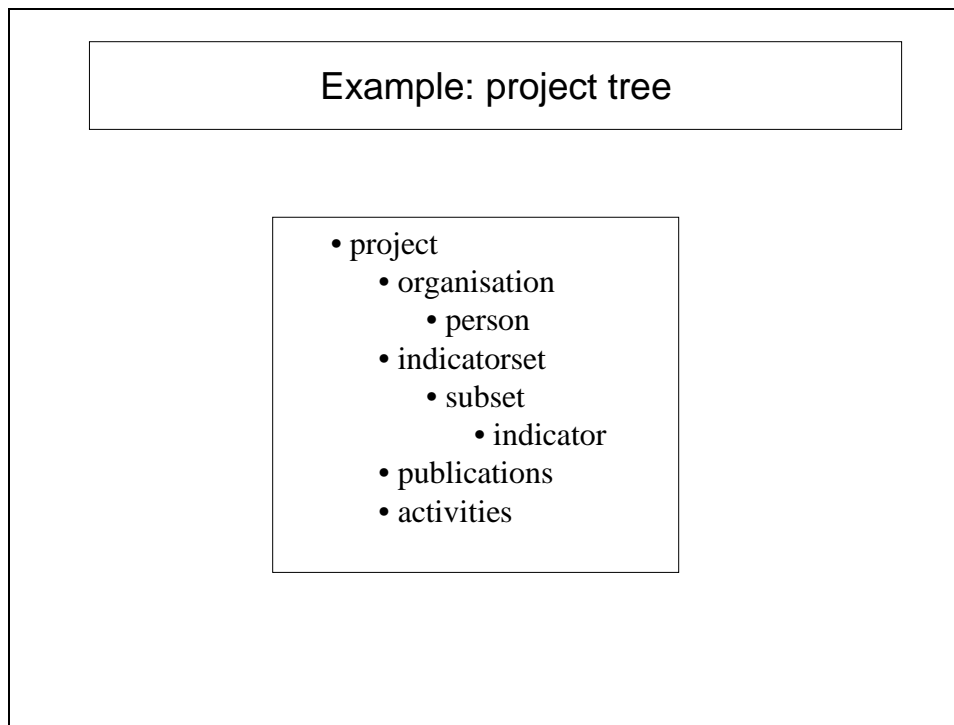


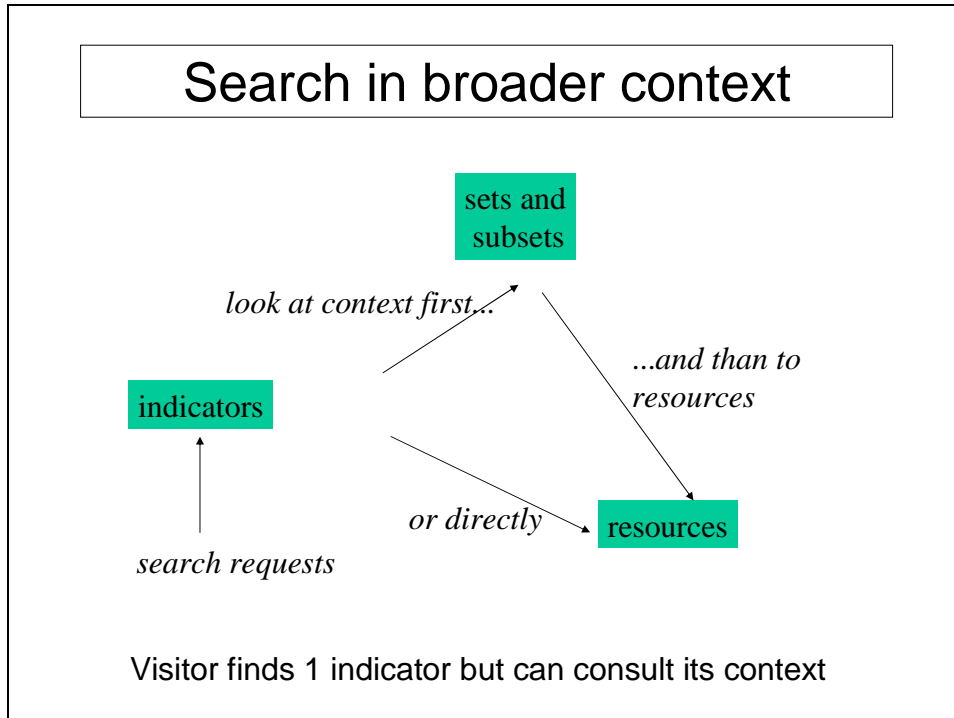
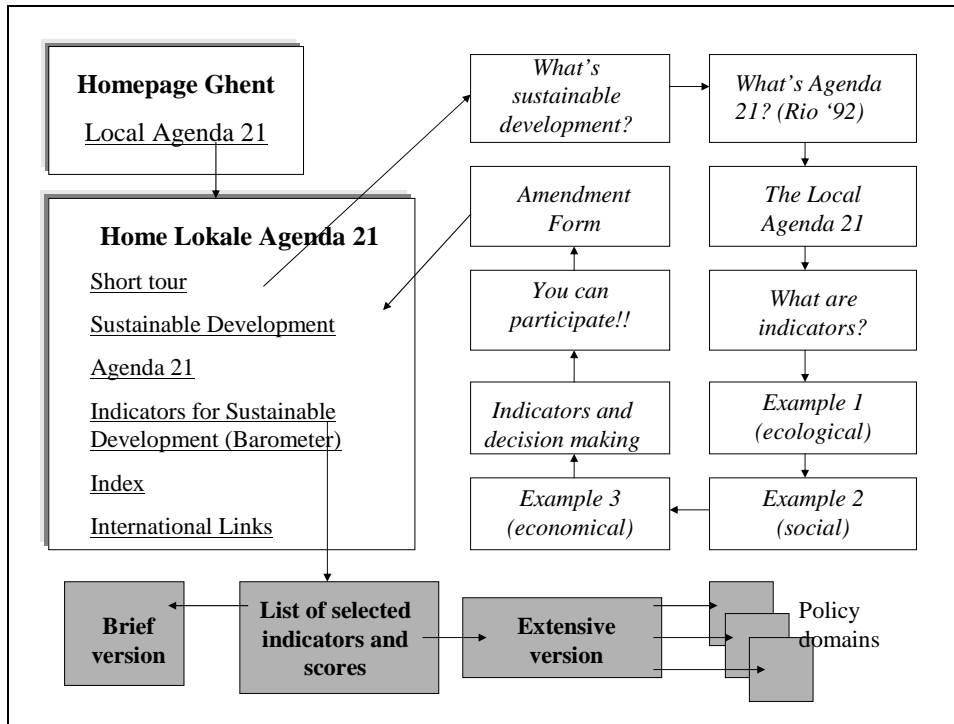
Other examples

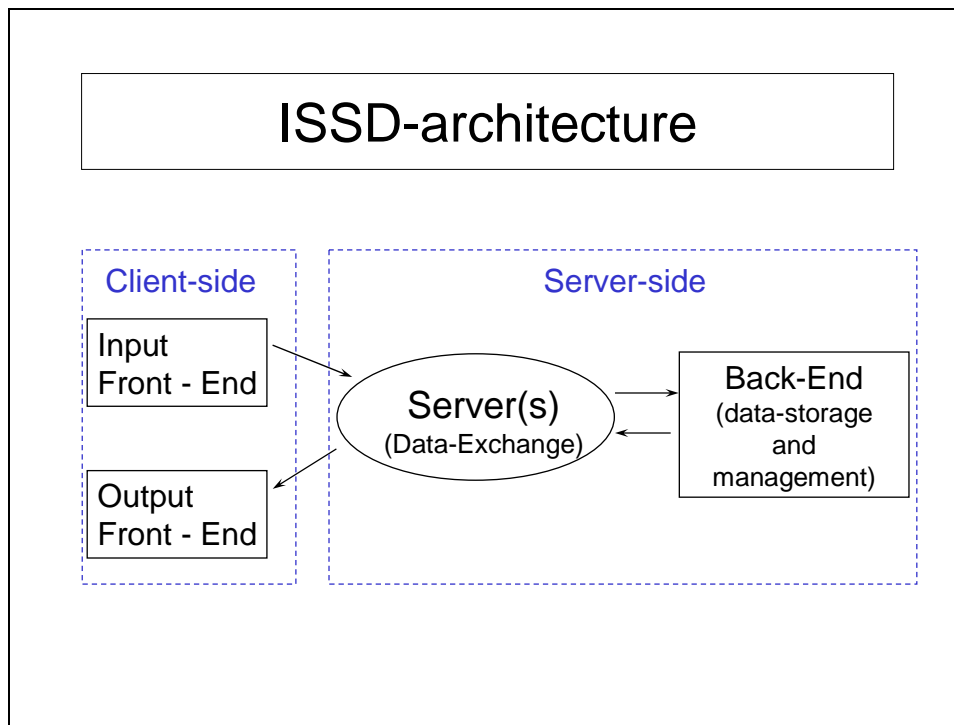
- Books have generally a tree like structure (chapters are ‘folders’ and sections in chapters are ‘subfolders’)
- The ‘best’ websites have a tree-like site-map
- Most classification-systems are hierarchical (also in the domain of sustainable development.)
- Help-files for software and online tutorials are ordered in a tree. Take note of the fact that tree-like help files have always an **index, search-facility** (and **lexicon**) as complementary tools.

Advantages of tree-like structure

- **Simple** (most end-users understand it): a tree is a list of lists of lists....(acyclic graph).
- **Flexible** (new folders can be inserted anywhere)
- Trees can be **virtual**. Data are not physically stored in a tree (files on harddisc are not really stored in folders). Trees can be changed, without really moving the data. Same data can be ordered in many trees.
- **Easy to visualise** (graphical user interfaces available for almost all programming languages: Visual Basic, Java, Tk,...)
- **Fits perfectly in XML-structure** which is always a tree of elements (‘folders’) and their children (‘subfolders’).
- Matrix-like structures (tables) can be converted into trees







Procedure 1: single XHTML-document

- Content-provider produces XHTML-document
- Validation of XHTML-document
- Content-provider 'uploads' his document in database-management system
- Server provides front-end: tree, search, index

Procedure 2: Robot harvests XML-document-collections

- Content-provider produces XHTML-documents
- Validation of XHTML-documents
- Content-provider makes his documents available online (FTP or HTTP-server)
- Content-provider produces XML-harvest-document and validates it
- Content provider 'asks' robot to visit his site
- ISSD-back-end harvests XHTML-documents
- Back-end stores metadata in database
- ISSD-server provides front-end: XHTML-files, tree, search, index

Software

• Client-side

End-user (content-provider and/or user)

- editors
- browsers
- email-client
- other 'office' -tools
- viewers

• Server-side

Data-managers and e-service-developers

- webserver (apache)
- emailserver
- database (MySQL)
- XSLT, DOM
- script (PHP, perl)