

Résumé

Le projet "Open Data Portal for European and Belgian GNSS Reference Station Data Collections, Built Upon FAIR Guiding Principles" (FAIR-GNSS), financé dans le cadre du programme BRAIN-be 2.0, s'est déroulé du 15 décembre 2020 au 15 mars 2023. Il a été coordonné par l'équipe GNSS de l'Observatoire royal de Belgique (ORB), avec le soutien de l'Université de Gand (UGent).

L'ORB gère depuis de nombreuses années des dépôts de données d'observation GNSS (Global Navigation Satellite Systems, comme GPS et Galileo), provenant de stations de référence belges et européennes, y compris le dépôt de données du réseau permanent EUREF (EPN), accessible au public. Ces données sont essentielles pour de nombreuses applications scientifiques et opérationnelles : surveillance des déformations du sol et des changements climatiques, étude de la météo spatiale, et amélioration des prévisions numériques du temps.

Malgré leur large utilisation, l'accès aux données GNSS RINEX de l'ORB était complexe. Les obstacles incluaient une faible découvrabilité des données, l'absence de métadonnées lisibles par machine, un manque d'informations sur la provenance et les licences, ainsi qu'une absence de pratiques standardisées de citation. Cela a mis en évidence la nécessité de moderniser la gestion des données GNSS à l'ORB, en appliquant les principes FAIR (Findable, Accessible, Interoperable, Reusable).

Objectifs

Le projet FAIR-GNSS visait à :

1. Améliorer l'accès et la réutilisation des données GNSS tout en renforçant la confiance des utilisateurs ;
2. Soutenir la préservation à long terme des ensembles de données GNSS ;
3. Contribuer à la standardisation des pratiques de citation des données ;
4. Développer un portail de données ouvertes pour les données GNSS belges et européennes.

Méthodologie

Pour atteindre ces objectifs, le projet a adopté une méthodologie structurée et progressive, inspirée de feuilles de route d'implémentation FAIR :

1. Évaluation de la conformité FAIR : Une analyse des lacunes a été réalisée à l'aide d'outils tels que le questionnaire FAIR-Aware et les FAIRsFAIR Data Object Assessment Metrics, afin d'évaluer la maturité FAIR initiale des dépôts GNSS de l'ORB et d'identifier les points à améliorer.

2. Création de FAIR Digital Objects (FDOs) : Les données GNSS ont été enrichies avec des métadonnées riches et lisibles par machine, et des Persistent Identifiers (PIDs) leur ont été attribués. Cela a nécessité une concertation avec la communauté GNSS pour définir le contenu et la structure des métadonnées, garantir leur conformité avec les standards communautaires (par ex. les vocabulaires du International GNSS Service) et adapter ou étendre les schémas de métadonnées existants (par ex. DCAT, schema.org). L'utilisation de PIDs a renforcé la trouvabilité, l'accessibilité, l'interopérabilité et la réutilisabilité des données.
3. Réorganisation des dépôts : Les bases de données internes et les flux de données ont été révisés pour permettre l'intégration de métadonnées riches. Des outils ont été développés pour automatiser des tâches telles que la validation des données et l'alimentation des dépôts. L'infrastructure a été évaluée à l'aide du modèle CoreTrustSeal+FAIR CapMat, avec pour objectif de devenir un dépôt de données fiable et compatible FAIR.
4. Mise en œuvre d'un accès FAIR aux données : Des APIs web ont été développées pour permettre un accès machine-actionable aux (méta)données. Ces APIs utilisent des standards ouverts (par ex. HTTPS) et supportaient des formats comme JSON, XML et JSON-LD. Les APIs GNSS existantes ont été revues pour évaluer leur compatibilité avec les exigences FAIR et les besoins des utilisateurs.
5. Développement d'un portail de données ouvertes : Une version bêta du portail a été lancée comme point d'accès central pour la recherche et la consultation des données et métadonnées GNSS. Ce portail intègre l'accès aux APIs, des outils de recherche et de la documentation. Sa conception s'est appuyée sur une analyse des bonnes pratiques issues d'initiatives européennes FAIR et GNSS.
6. Développement itératif : Le projet a suivi une progression logique, débutant par l'enrichissement des (méta)données GNSS, suivi par l'amélioration des protocoles d'accès, et culminant avec la mise en place de services orientés utilisateur tels que le portail de données.

Résultats clés

Deux propositions de schémas de métadonnées GNSS ont été développées : une pour les métadonnées spécifiques aux stations, et une pour les fichiers GNSS RINEX. La proposition pour les stations a été approuvée par le GeodesyML Task Force du International GNSS Service (IGS).

Un modèle de métadonnées a été conçu pour permettre la citation des jeux de données GNSS via DOI, en collaboration avec le GGOS DOI Working Group. Un outil logiciel et une interface web ont été créés pour permettre aux fournisseurs de données d'attribuer des DOIs à leurs jeux de données.

Les flux de données internes de l'ORB ont été réorganisés afin de documenter systématiquement les informations de provenance, améliorer la traçabilité des données et consigner les corrections apportées aux fichiers.

Le portail de données ouvertes <https://www.gnss.be/opendataportal/> et les APIs associées permettent un accès amélioré aux données GNSS validées, au bénéfice de la recherche scientifique ainsi que des applications publiques/privées de géoréférencement.

Conclusions

Le projet FAIR-GNSS a considérablement amélioré la maturité FAIR des dépôts GNSS de l'ORB. L'intégration de métadonnées standardisées, de PIDs et d'APIs ouvertes a renforcé la découvrabilité, l'accessibilité et la réutilisabilité des ensembles de données GNSS. Ces avancées profitent directement aux communautés scientifiques actives dans la surveillance de l'environnement, les sciences de la Terre et les applications géospatiales.

Recommandations

Mettre en œuvre un écosystème de données FAIR nécessite des investissements soutenus en expertise, en infrastructures, et en collaboration avec des data stewards et des experts du domaine. Une planification à long terme et des évaluations régulières de la conformité FAIR sont essentielles pour suivre l'évolution des standards et des besoins des utilisateurs. L'engagement avec les communautés GNSS et géoscientifiques au sens large est indispensable pour établir des standards partagés et garantir leur adoption. Bien que ce processus soit exigeant en ressources, chaque cycle de développement renforce la qualité des données, la confiance des utilisateurs et les services offerts.

Un enseignement majeur de FAIR-GNSS est que l'application des principes FAIR n'est pas une tâche ponctuelle, mais un processus continu générant des améliorations cumulatives. Les services GNSS de l'ORB répondent désormais mieux aux besoins des utilisateurs belges, européens et internationaux, tout en posant les bases d'une future certification de confiance.