

CoreBook et la Cyber-carothèque nationale : deux outils pour bancariser des métadonnées riches lors des opérations de carottage scientifique (continental, marin, glaciaires)

Fabien Arnaud^{1*}, Cécile Pignol¹, Xavier Crosta², Isabelle Billy², Bruno Galabertier¹, Pascal Calvat³, Julien Penguen³, Elodie Godinho⁴, Karim Bernardet⁴

¹EDYTEM, CNRS, Université Savoie Mont-Blanc, France

²EPOC, CNRS, Université de Bordeaux, France

³OASU, CNRS, Université de Bordeaux, France

⁴DT-INSU, CNRS, France

La bancarisation des données demeure défi majeur pour la recherche scientifique. Dans les domaines des géosciences basées sur l'étude d'échantillons, un retard considérable a été pris pour de multiples raisons. Ce problème est comparable à celui des « Longues traînes » de données, ces données produites de manières disparates par un grand nombre d'acteurs déconnectés les uns des autres et qui constituent à la fois, la grande majorité des activités de recherche dans le monde et un défi pour la bancarisation des données. Profitant des moyens financiers et de la visibilité nationale offerts par le projet d'équipement d'excellence CLIMCOR (2013-2019), les laboratoires EPOC et EDYTEM, la DT-INSU et l'OASU, ont joint leurs efforts pour proposer une réponse communautaire et « FAIR » à ce défi.

Nous poserons les bases conceptuelles d'un dispositif visant à bancariser des données riches issues d'opérations de prélèvement d'échantillons géologiques de type « carottes ». Nous insisterons notamment sur l'articulation de cette base avec des systèmes tiers et montrerons que la collecte de données de terrain offre un socle fondamental pour bancariser ultérieurement toutes les données acquises sur ces échantillons.

Nous insisterons sur la présentation opérationnelle des deux outils actuellement disponibles : i) CoreBook, une application Android dédiée à la capture des informations de carottage et ii) Cyber-carothèque nationale, à la fois entrepôt de données, visualiseur et distributeur de données et outil de recherche. Ces outils fonctionnent quel que soit l'opérateur institutionnel de terrain : labos, Flotte Océanographique Française, partenaires étrangers, ou encore prestataires privés.

Nous insisterons enfin sur les aspects d'interopérabilité, notamment avec le système d'identification IGSN. Ces outils et l'effort consenti par notre consortium s'inscrivent ainsi dans l'effort mondial pour la bancarisation des données et contre la crise de la reproductibilité.

Mots-Clés : Carottages scientifiques ; Métadonnées de terrain ; Échantillons ; FAIR ; IGSN ; CoreBook ; Cyber-carothèque nationale