
Plateforme logicielle pour l'intégration et la composition de services géospatiaux

Meriem Sabrina Halilali¹, Eric Gouardères¹,
Florent Devin^{1,2}, Mauro Gaio¹

1. LIUPPA, Université de Pau et des Pays de l'Adour, 64013 Pau Cedex,
France

meriem-sabrina.halilali@univ-pau.fr,eric.gouarderes@univ-pau.fr,mauro.gαιο@univ-pau.fr

2. EISTI, 2 boulevard Lucien Favre, 64075 Pau Cedex, France

fd@eisti.eu

RÉSUMÉ. Au cours des dernières années de grands progrès ont été réalisés sur les systèmes d'information géographique et sur le World Wide Web. Les services géospatiaux se sont progressivement développés pour fournir un ensemble de fonctionnalités et de données issues de sources hétérogènes. La question centrale est : comment faciliter, la découverte, l'intégration et l'exploitation de services géospatiaux, susceptibles d'être hétérogènes, complémentaires voire conflictuels, pour en assurer l'interopérabilité et fournir des données métiers pour les géosciences. L'objectif de cet article est la présentation de notre démarche pour la conception et le développement d'une plateforme logicielle qui puisse répondre à cette question. Ce travail s'intègre dans le cadre du projet ANR CHOUCAS pour l'aide à la localisation de victimes en montagne.

ABSTRACT. During last years great progress were realized on the geographical information systems and on the World Wide Web. Geospatial services have been gradually developed to provide a set of features and data from heterogeneous sources. The key question is: how to facilitate the discovery, integration and exploitation of geospatial services, likely to be heterogeneous, complementary or even conflicting, to ensure interoperability and provide business data for geosciences. The purpose of this article is the presentation of our approach for the design and the development of a software platform that can answer this question. This work is part of the ANR CHOUCAS project for help in locating victims in the mountains.

MOTS-CLÉS : Service géospatial, Taxonomie, Composition de services Web, Web sémantique, Intégration de données géographiques hétérogènes.

KEYWORDS: Geospatial Service, Taxonomy, Web services composition, Semantic Web, Integration of heterogeneous geographic data.

1. Introduction

Ce travail s'inscrit dans le cadre du projet ANR CHOUCAS (Olteanu-Raimond *et al.*, 2017). Ce projet implique plusieurs équipes de recherches interdisciplinaires qui contribuent chacune à la création d'une géo-application pour l'amélioration de l'aide à la localisation des victimes en montagne. Un des objectifs est de proposer des méthodes et des outils permettant de développer des processus métiers pour favoriser la prise de décision. Dans notre contexte, un processus métier est une suite de traitements qui vise à constituer et enrichir des données géographiques issues de sources hétérogènes (gazetiers, base de données cartographiques, sites Web de randonnées...) afin de fournir des données métiers aux secouristes et permettre une localisation plus rapide des victimes en montagne.

Les travaux ici présentés proposent la conception et le développement d'une plateforme logicielle mettant à la disposition des partenaires un ensemble de services géospatiaux. Il s'agit de programmes modulaires, ou composants logiciels, conçus pour faciliter la découverte, l'accès et la composition de fonctions de traitement d'informations géospatiales sur Internet. En particulier, cette plateforme intégrera, sous forme de services, différentes sources de données pour en assurer l'interopérabilité et fournir les données métiers. La plateforme est construite autour d'une infrastructure de type SOA (Service-Oriented Architecture). Ce choix a été motivé par les raisons suivantes : favoriser le découpage en services pour la mise en commun des ressources informationnelles ; éviter la multiplication des environnements de déploiement ; éviter l'hétérogénéité des applications et regrouper au sein d'une même application un ensemble de service afin de proposer une vue cohérente ; et enfin disposer du même point d'entrée sans avoir à se soucier de la localisation des centres de traitement et/ou stockage.

À travers cette infrastructure, on s'attache à lever deux verrous scientifiques. Le premier verrou concerne l'intégration de différentes sources de données sous forme de services géospatiaux et la structuration de ces services qui sont susceptibles d'être hétérogènes, complémentaires ou conflictuels. La section 2 présente notre vision pour garantir cette intégration dans un catalogue de services en s'appuyant sur une taxonomie. Le second verrou concerne l'automatisation de la composition des services et la réutilisabilité de ces services composites. L'hétérogénéité et le manque de sémantique dans ces services constituent un obstacle qu'il s'agit de lever. La section 3 explique comment nous envisageons d'utiliser les principes du Web sémantique, notamment la description de taxonomie, pour automatiser le processus de composition de services représentant

des processus métiers. Enfin nous concluons l'article avec quelques perspectives présentées dans la section 4.

2. Intégration des services géospatiaux dans un catalogue de services

Il existe une grande quantité de services géospatiaux en ligne qui fournissent diverses données et fonctions géospatiales. La question centrale est de garantir une intégration transparente des sources de données et des fonctions dans notre catalogue, que celles-ci soient internes (des services qui seront développés dans le cadre du projet CHOUCAS), fournies par les partenaires du projet (des services conformes aux standards de l'OGC (Open Geospatial Consortium)¹, WMS (Web Map Service), WMTS (Web Map Tile Service), WFS (Web Feature Service)) disponibles sur le portail IGN², ou externes en provenance du Web. Afin de faciliter la structuration et l'interrogation de ces derniers, nous utilisons les technologies du Web sémantique. Celles-ci visent à offrir de meilleures méthodes d'intégration, de découverte et d'accès aux données. Elles permettent, en particulier, aux traitements d'exploiter la description sémantique (exprimée selon des standards) des données géospatiales, des services et des chaînes de services. Elles jouent un rôle important également dans la découverte et la structuration des services géospatiaux aux moyens des ontologies et des taxonomies. Nous utilisons ces dernières afin de structurer notre domaine en définissant les concepts importants qui le caractérisent et les relations entre ces concepts.

Dans les travaux qui portent sur les taxonomies pour la description et la classification de services Web géospatiaux nous constatons la présence de deux approches. La première consiste à concevoir des taxonomies génériques basées sur des considérations générales (Bai *et al.*, 2009). C'est le cas de la norme internationale ISO 19119 (19119:2016(fr), s. d.) pour la classification des services Web de l'OGC qui propose une classification s'appuyant sur les principes des architectures logicielles en technologies de l'information (interaction, modèle, traitement, communication...). L'avantage avec ce type de taxonomie, est de procurer un bon référentiel standard et réutilisable. Cependant, elle reste trop générale et souffre de l'absence de description sémantique formelle directement exploitable par une machine. La seconde consiste plutôt à concevoir des taxonomies spécifiques basées sur des considérations métiers (Al-Areqi *et al.*, 2016; Yue *et al.*, 2011). Une taxonomie spécifique est en général très bien adaptée au contexte métier auquel elle s'adresse. Sa spécificité la rend malheureusement très peu réutilisable dans d'autres contextes.

1. OGC Web services standards, <http://www.opengeospatial.org/standards>

2. portail IGN Institut national de l'information géographique et forestière <http://geoservices.ign.fr>

Pour les raisons citées, et pour aborder le premier verrou nous avons opté pour l'intégration du Web sémantique (Berners-Lee *et al.*, 2001 ; McIlraith *et al.*, 2001 ; Li *et al.*, 2015) pour la description de taxonomies géospatiales. Nous préconisons de produire deux taxonomies distinctes, une pour les services et une pour les données. La taxonomie des services s'appuiera sur une caractérisation fonctionnelle en respectant les normes internationales. Cette taxonomie sera utilisée pour décrire formellement et sémantiquement les services géospatiaux. La taxonomie de données contribuera en particulier à la conception d'un DSL (Domain Specific Language) pour annoter sémantiquement les données d'entrées/sorties des services géospatiaux. Ce DSL sera une spécialisation de langages d'annotation standards tels que TEI³ et ISO-Space (Pustejovsky, 2017).

Notre approche pour la description sémantique de données et de services géospatiaux vise à améliorer la découverte des services géospatiaux et faciliter la composition automatique de chaînes de services en assurant un bon équilibre entre les taxonomies génériques pour soutenir la réutilisation et spécifiques pour répondre à nos besoins métiers.

3. Composition de services géospatiaux pour la création de processus métier

Après l'intégration dans un catalogue de service, nous disposerons de divers services Web géospatiaux (dont le nombre peut être important). Ces services sont susceptibles d'interagir entre eux dans le but de répondre à une requête utilisateur qui correspond à un besoin métier. Considérons par exemple une requête utilisateur qui vise à attribuer une géolocalisation à des entités spatiales dans un texte. S'il existe un service atomique qui permet de satisfaire la requête, celui-ci est directement présenté à l'utilisateur. Dans le cas contraire, il faut construire un service composite. Dans notre exemple, on peut construire une chaîne de services par composition de trois services (Moncla, Gaio, 2017). Le premier pour interroger les gazetiers. Le second pour la désambiguïsation dans la mesure où plusieurs gazetiers sont utilisés. Le troisième pour générer la liste des toponymes.

Pour concevoir de meilleures méthodes de composition de services géospatiaux, nous nous appuyons sur les technologies du Web sémantique. Celles-ci exploitent les descriptions sémantiques de services géospatiaux afin d'en permettre la découverte automatique. La découverte de service consiste à l'émission de requêtes basées sur les concepts présents dans la taxonomie de service, pour rechercher tous les services qui répondent à une certaine fonctionnalité ou besoin métier. Dans un deuxième temps des requêtes sémantiques sont émises selon la description sémantique des données. Elles se basent sur les concepts de la taxonomie données pour assurer la compatibilité des types de données

3. TEI norme internationale pour le balisage textuel. <http://www.tei-c.org/Guidelines/P5/>

(entrées/sorties) entre les différents services sélectionnés et assurer leur interopérabilité.

Bien qu'il existe des travaux de ce type (Yue *et al.*, 2007; Al-Areqi *et al.*, 2016), ils ont été développés dans un environnement dédié qui ne favorise pas la réutilisation de services atomiques ou composites. D'autres propositions (Lemmens *et al.*, 2006) en ce sens ont été faites sur la base des architectures SDI (Spatial Data Infrastructure) et des descriptions de services améliorées OWL-S (Web Ontology Language for Services). Bien qu'intéressantes, elles ne prennent pas en considération les standards ouverts actuels pour la gestion des données géospatiales. Ces constats nous incitent à proposer un environnement ouvert qui favorise la réutilisation en s'appuyant sur les standards de l'OGC et les dernières recommandations de l'ISO en termes d'architectures et de technologies logicielles pour les services géospatiaux.

4. Conclusion

Dans cet article, nous avons présenté notre démarche pour l'intégration de services géospatiaux dans un catalogue de services et leur composition pour la création de processus métiers au sein d'une infrastructure de type SOA. Notre démarche se base sur deux taxonomies pour la description sémantique de données et de services géospatiaux. Le but de ces taxonomies est de structurer nos sources de données suivant les caractéristiques communes de ces dernières et de fournir un référentiel commun pour la découverte et la composition des services.

Dans ce contexte, nous avons initié la conception d'une taxonomie spécifique de services géospatiaux à partir des services fournis par nos collègues qui travaillent sur l'appariement de traces GPS avec des descriptions textuelles de randonnées. Sur la base de cette taxonomie, nous avons modélisé la chaîne de services présentée dans (Moncla, Gaio, 2017) en utilisant un éditeur de workflow. L'objectif est de l'implémenter sur plusieurs outils décrits dans la littérature afin d'en évaluer les possibilités. Un premier démonstrateur a été implémenté sur l'environnement jABC⁴ qui s'est avéré trop fermé pour nos besoins. Quelques plates-formes d'orchestration de services géospatiaux (ZOO Project⁵, 52° North⁶) ont été suggérées dans les travaux de (Rautenbach *et al.*, 2012). Les plates-formes ont été évaluées pour leurs capacités d'orchestration des services pour la production de cartes thématiques. Nous allons évaluer dans nos prochains travaux ces plates-formes suivant les exigences et les besoins du projet CHOUCAS.

4. Java Application Building Center. <http://ls5-www.cs.tu-dortmund.de/projects/jabc/>

5. ZOO Project. <http://www.zoo-project.org>

6. 52° North. www.52north.org

Concernant le travail sur les taxonomies, nous allons généraliser notre taxonomie de services pour prendre en compte la classification ISO 19119 et les services fournis par nos partenaires du projet CHOUCAS. La collaboration avec ces derniers doit nous permettre également d'identifier et de caractériser les données géospatiales dont ils ont besoin afin de concevoir notre taxonomie de données.

Bibliographie

- 19119:2016(fr) I. (s. d.). *Iso/tc 211. (2016). iso 19119:2016(fr), information géographique — services*. <https://www.iso.org/obp/ui/#iso:std:iso:19119:ed-2:v1:fr>.
- Al-Areqi S., Lamprecht A.-L., Margaria T. (2016). Constraints-driven automatic geospatial service composition: Workflows for the analysis of sea-level rise impacts. In *International conference on computational science and its applications*, p. 134–150.
- Bai Y., Di L., Wei Y. (2009). A taxonomy of geospatial services for global service discovery and interoperability. *Computers & Geosciences*, vol. 35, n° 4, p. 783–790.
- Berners-Lee T., Hendler J., Lassila O. (2001). The semantic web. *Scientific american*, vol. 284, n° 5, p. 34–43.
- Lemmens R., Granell C., Wytzisk A., de By R., Gould M., van Oosterom P. (2006). Semantic and syntactic service descriptions at work in geo service chaining. In *Proceedings of the 9th AGILE Conference on Geographic Information Science, shaping the future of geographic information science in europe*, p. 51–61. Visegrad, Hungary, University of West Hungary Press.
- Li H., Wang Y., Cheng P. (2015). Semantic description for the taxonomy of the geospatial services. *Boletim de Ciências Geodésicas*, vol. 21, n° 3, p. 515–531.
- McIlraith S. A., Son T. C., Zeng H. (2001). Semantic web services. *IEEE intelligent systems*, vol. 16, n° 2, p. 46–53.
- Moncla L., Gaio M. (2017). Services web pour l'annotation sémantique d'information spatiale à partir de corpus textuels. In *Spatial Analysis and GEomatics 2017 (Sagéo 2017)*, p. 2–14. Rouen, France. (<hal-01633342>)
- Olteanu-Raimond A.-M., Davoine P.-A., Gaio M., Gouardères E., Van Damme M.-D., Villanova-Oliver M. *et al.* (2017, novembre). Projet CHOUCAS : Intégration de données hétérogènes et raisonnement spatial pour l'aide à la localisation des victimes en montagne. In *Spatial Analysis and GEomatics 2017 (Sagéo 2017)*. Rouen, France. (<hal-01649156>)
- Pustejovsky J. (2017). Iso-space: Annotating static and dynamic spatial information. In *Handbook of linguistic annotation*, p. 989–1024. Van Godewijkstraat 30, 3311 GX Dordrecht, The Netherlands, Springer. (ISBN 978-94-024-0879-9)
- Rautenbach V., Coetzee S., Strzelecki M., Iwaniak A. (2012). Results of an evaluation of the orchestration capabilities of the zoo project and the 52^r north framework for an intelligent geoportal. *ISPRS Annals of Photogrammetry, Remote Sensing*

and Spatial Information Sciences, vol. I-4, p. 163–168. Consulté sur <https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/I-4/163/2012/>

Yue P., Di L., Yang W., Yu G., Zhao P. (2007). Semantics-based automatic composition of geospatial web service chains. *Computers & Geosciences*, vol. 33, n° 5, p. 649–665.

Yue P., Gong J., Di L., He L., Wei Y. (2011). Integrating semantic web technologies and geospatial catalog services for geospatial information discovery and processing in cyberinfrastructure. *GeoInformatica*, vol. 15, n° 2, p. 273–303.