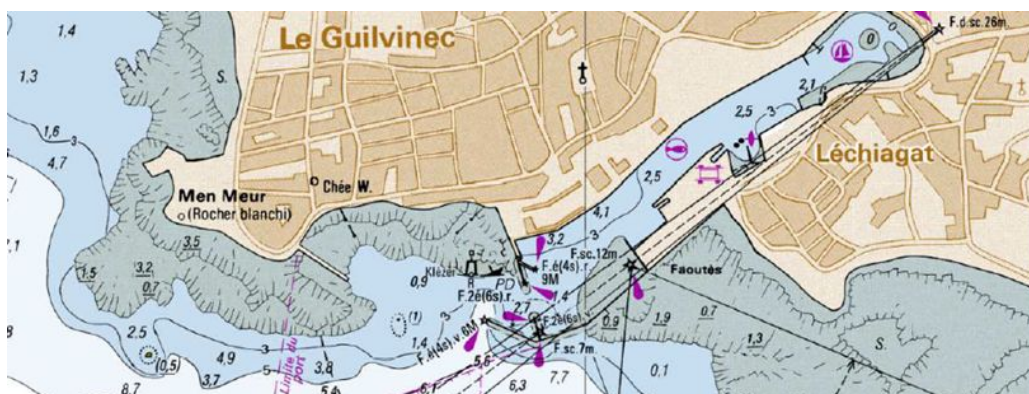


## Création d'une base de connaissances hydrographiques à partir de sources hétérogènes

**Mots-clés:** création et peuplement de bases de connaissances géoréférencées, intégration de données à références spatiales, Web sémantique géospatial, extraction d'informations à partir de textes.

**Contexte et objectifs :** Les cartes marines publiées par le [Service Hydrographique et Océanographique de la Marine](#) (SHOM) décrivent la topographie des côtes et la morphologie des fonds marins pour assurer la sécurité de la navigation de surface dans les zones côtières françaises. Ces cartes s'accompagnent de textes appelés *instructions nautiques* qui complètent les informations apportées par les cartes en détaillant les pratiques de navigation et leurs particularités locales : manoeuvres d'accès aux ports, réglementation, équipements et services, etc.



### 2.2.4. Guilvinec-Léchiagat

Carte 6646.

Le port départemental de Guilvinec-Léchiagat (vue 2.2.4.4.) est fréquenté essentiellement par les navires de pêche. L'accès en est difficile par forts vents d'Ouest à WSW qui lèvent un ressac sensible. Par forts vents de SE, le chenal devient dangereux pendant les deux dernières heures du jusant, la mer déferlant sur la Basse aux Herbes et Pen Men Ham, rochers couverts respectivement de 1,8 et 4,3 m d'eau.

#### 2.2.4.1. Accès

L'entrée du chenal est située à 2,5 M au SW de l'entrée du port. L'axe du chenal est matérialisé par l'alignement lumineux à 053° de trois feux implantés sur la rive Est du port. Les feux antérieur et intermédiaire sont portés par des pylônes blancs. Le feu postérieur est porté par une tour rectangulaire blanche avec une bande médiane verticale rouge et surmontée par un cylindre rouge. De jour, la visibilité de cet alignement est améliorée par deux voyants cylindriques rouges de grandes dimensions, l'un implanté sur un support rouge à proximité du feu intermédiaire, l'autre au sommet de la tour du feu postérieur.

#### **Extrait d'une carte marine et des instructions nautiques associées (Source : SHOM)**

De nombreux travaux récents ont porté sur la formalisation de connaissances sur l'environnement hydrographique marin et la navigation pour développer des services d'aide à la navigation à base de raisonnement : caractérisation automatique d'éléments du relief sous-marin à partir d'isobathes et de points de sondage (Yan et al., 2014) et (Yan et al., 2015), production automatique de cartes marines (Yan et al., 2017), description de trajectoires de navigation côtière (Laddada et Saux, 2017) ou encore génération automatique d'instructions nautiques (Haralambous et al., 2014) (Sauvage-Vincent et al., 2015) (Haralambous et al., 2017). Les approches proposées s'appuient sur une base de connaissances sur l'environnement hydrographique marin qui n'existe pas à ce jour. Les informations nécessaires à sa création sont disséminées dans différentes sources : ontologies construites manuellement, instructions nautiques en langage naturel, référentiels de données géolocalisés, etc.

L'objectif principal de la thèse sera de construire et peupler cette base de connaissances. Pour cela, on explorera les potentialités du Web de données (Atemezing et al., 2014) (Kuhn et al., 2014) pour représenter et intégrer les connaissances hydrographiques nécessaires. On fera de plus appel à des techniques d'extraction automatique d'information spatiale à partir de textes (Kergosien et al. 2014) (Kergosien et al., 2017) (Zenasni et al., 2018) pour la peupler et ainsi produire une base de connaissances hydrographiques opérationnelle.

**Approche :** L'environnement hydrographique marin est constitué d'entités géographiques de surface ou sous-marines, artificielles ou naturelles<sup>1</sup>. Il englobe également des éléments de contexte géographique visibles depuis la mer tels que des villes ou des bâtiments divers.

La thèse consistera à :

- proposer les ontologies adéquates pour représenter les différents types d'entités hydrographiques, leurs propriétés, leurs géométries et leurs relations en s'appuyant sur les standards de l'Organisation Hydrographique Internationale<sup>2</sup> (IHO) (SHOM) ;
- représenter les référentiels disponibles (SHOM, [IGN](#), OpenSeaMap, etc.) conformément à ces ontologies ;
- enrichir cette première base de connaissances par reconnaissance et résolution automatiques des entités hydrographiques mentionnées dans les instructions nautiques. Ceci supposera de prendre en compte des critères d'évaluation de similarité dont la disponibilité pourra varier selon les informations de contexte pouvant être extraites des textes (nom, nature ou fonction de l'entité hydrographique, localisation relative, relations de voisinage, etc.);
- vérifier la cohérence, qualifier et enrichir les informations intégrées à la base de connaissances à l'aide d'inférences. Il pourra être nécessaire d'explicitier certains types d'entités hydrographiques à partir des connaissances de bathymétrie disponible.

**Direction de thèse, encadrement de thèse:**

Directeur de thèse : [Eric Saux](#) (Irenav - Ecole Navale, [eric.saux@ecole-navale.fr](mailto:eric.saux@ecole-navale.fr))

Co-directrice de thèse : [Nathalie Abadie](#) (LaSTIG - COGIT/Strudel, IGN, [nathalie-f.abadie@ign.fr](mailto:nathalie-f.abadie@ign.fr))

Encadrant : [Eric Kergosien](#) (Geriico - Université de Lille, [eric.kergosien@univ-lille3.fr](mailto:eric.kergosien@univ-lille3.fr))

**Contrat doctoral:** Le contrat doctoral, d'une durée de trois ans, ouvre droit à une rémunération d'environ 1680 € brut (hors contribution aux frais de transports). Le contrat doctoral peut inclure pour l'ensemble de la durée de la thèse un service complémentaire d'enseignement, de diffusion de l'information scientifique et technique, de valorisation ou d'expertise.

**Localisation:** Equipe [LaSTIG/Strudel](#), [Institut national de l'information géographique et forestière](#) (IGN), Saint-Mandé (métro 1, station Saint Mandé).

**Début de la thèse:** automne 2018

**Profil recherché:** Master 2 ou diplôme d'ingénieur en informatique : représentation de connaissances, Web sémantique, sciences de l'information géographique, extraction d'information à partir de textes.

**Candidature :** Envoyer par e-mail à l'ensemble des encadrants (1) votre CV, (2) une lettre de motivation adaptée au sujet proposé, (3) vos relevés de notes des deux dernières années d'études, (4) l'avis du directeur de master (ou équivalent) et le cas échéant des lettres de recommandations.

**Date limite de candidature:** 5 septembre 2018

---

<sup>1</sup> Baies, rochers, bancs de sables, estuaires, ports, phares, bouées, câbles sous-marins, épaves, obstructions, etc.

<sup>2</sup> <https://www.iho.int/srv1/index.php?lang=en>

## Bibliographie

- Atemezing, G.A., N. Abadie, R. Troncy and B. Bucher (2014)** Publishing Reference Geodata on the Web: Opportunities and Challenges for IGN France. Terra Cognita 2014, 6th International Workshop on the Foundations, Technologies and Applications of the Geospatial Web. In Conjunction with the 13th International Semantic Web Conference.
- Haralambous, Y., Sauvage-Vincent, J. and Puentes, J., 2014.** INAUT, a controlled language for the French coast pilot books instructions nautiques. In International Workshop on Controlled Natural Language (pp. 102-111). Springer, Cham.
- Haralambous, Y., Sauvage-Vincent, J. and Puentes, J., 2017.** A hybrid (visual/natural) controlled language. Language Resources and Evaluation, 51(1), pp.93-129.
- Kergosien E., Laval B., Roche M., Teisseire M. 2014.** OPILAND : identification de la perception des territoires par la fouille de texte. Revue des Nouvelles Technologies de l'Information, Hermann, 2014, MASHS'2014: Fouille de Données et Humanités Numériques, RNTI-SHS-2, pp.185-212. <lirmm-01054916>
- Kergosien E., Sallabery C., Bessagnet M.-N., Le Parc-Lacayrelle A., Chaudiron S. 2017.** Using a GIR tool in a Business Intelligence Context: the case of EGC conferences. Proceedings of the 7th International Conference on Information Systems and Economic Intelligence , May 2017, Al-Hoceïma, Morocco. <hal-01590207>
- Kuhn, W., Kauppinen, T. and Janowicz, K., 2014, September.** Linked data-a paradigm shift for geographic information science. In *International Conference on Geographic Information Science* (pp. 173-186). Springer, Cham.
- Laddada, W. and Saux, É., 2017.** Description formelle de trajectoire de navigation en environnement maritime côtier. Revue Internationale de Géomatique, 27(2), pp.179-202.
- Sauvage-Vincent, J., Haralambous, Y. and Puentes, J., 2015.** Sentence ordering in electronic navigational chart companion text generation. In Proceedings of the 15th European Workshop on Natural Language Generation (ENLG) (pp. 66-70).
- SHOM, France.** Interest of semantic Web standards for S-100 developments. Paper for consideration by S-100 WG. En ligne, consulté le 01/04/2018 : [https://www.iho.int/mtg\\_docs/com\\_wg/S-100WG/S-100WG2/S100WG2-9.3B\\_SemanticWeb.pdf](https://www.iho.int/mtg_docs/com_wg/S-100WG/S-100WG2/S100WG2-9.3B_SemanticWeb.pdf)
- Yan, J., Guilbert, E. and Saux, E., 2014.** An ontology for the generalisation of the bathymetry on nautical charts. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2(2), p.1.
- Yan, J., Guilbert, E. and Saux, E., 2015.** An ontology of the submarine relief for analysis and representation on nautical charts. The Cartographic Journal, 52(1), pp.58-66.
- Yan, J., Guilbert, E. and Saux, E., 2017.** An ontology-driven multi-agent system for nautical chart generalization. Cartography and Geographic Information Science, 44(3), pp.201-215.
- Zenasni S., Kergosien E., Roche M., Teisseire M. 2018.** Spatial Information Extraction from Short Messages. Expert Systems with Applications, Elsevier, 2018, 95, pp.351 - 367. <IF 3.93> . <10.1016/j.eswa.2017.11.025> . <hal-01706936>