



# Thomas FRAN CART

---

**Date de naissance:** 22/09/1980 | **Nationalité:** Française | **Numéro de téléphone:**

(+33) 671112597 (Tél. mobile) | **Adresse électronique:**

[thomas.francart@sparna.fr](mailto:thomas.francart@sparna.fr) | **Site web:** <http://sparna.fr> | **Site web:**

<http://blog.sparna.fr> | **LinkedIn:** <http://www.linkedin.com/in/thomasfrancart> |

**Adresse:** 5 rue Georges Courteline, 37540, Saint-Cyr-Sur-Loire, France  
(Domicile)

## ● EXPÉRIENCE PROFESSIONNELLE

---

07/2012 – EN COURS France

### CONSULTANT SENIOR WEB DE DONNÉES SPARNA

---

Fondateur de la société **Sparna** pour valoriser une expertise dans le web de données : consulting, formations et développements logiciel utilisant les technologies du web de données, les systèmes d'organisation des connaissances, les bases de graphes sémantiques, les métadonnées documentaires.

Clients: (voir ci-dessous pour les descriptions détaillées)

- **Organisations Internationales** : Parlement Européen, OMC, Office des Publications de l'UE, Portail CORDIS, Taskforce ELI, Centre International de l'ISSN, UNESCO
- **Institutions et organisations nationales Françaises** : Archives Nationales, Bibliothèque Nationale, LexisNexis France, Agence du Numérique en Santé, INSEE, Réseau Canopé, etc.
- Humanités Numériques : Cité de la Musique/Philharmonie de Paris, MSH Val-de-Loire, Consortium MASA (archéologie), Huma-Num
- Startups (e-Santé, IA, etc.)

Formations: animation de plus de 50 sessions de formation sur le web de données.

Logiciels open-source:

- [Sparnatural](#) : constructeur visuel et configurable de requêtes SPARQL
- [SKOS Play](#) : service gratuit et ouvert de visualisation de thesaurus SKOS
- [SHACL Play](#) : service gratuit et ouvert de validation de données RDF en SHACL, et de génération de documentation à partir de modèles SHACL
- [xls2rdf](#) : service gratuit et ouvert de conversion de tableaux Excel en RDF

Spécialisation dans les Systèmes d'Organisation des Connaissances : thesaurus, ontologies, modèles de bases de données, classifications statistiques, instruments de recherche archivistiques, notices bibliographiques, descriptions patrimoniales, vocabulaires contrôlés, données patient, etc.

10/2007 – 09/2012

### DIRECTEUR TECHNIQUE MONDECA

---

Responsable d'ITM (Intelligent Topic Manager), une solution de gestion de systèmes d'organisation de connaissances basée sur les technologies du web sémantique, éditée par Mondeca (Paris, 22 personnes).

Définition de la stratégie et de la roadmap produit, déblocage des verrous technologiques, analyse des besoins client et nombreuses participations en tant qu'architecte technique à des projets nationaux et internationaux et à des projets de recherche. Forte implication dans le processus d'avant-vente. Mise en place de processus qualité avec l'équipe de développement (CI / CD).

Conception d'architectures d'intégration entre des systèmes de gestion de connaissances/ontologies (Mondeca), des systèmes de text mining (Temis, Arisem, Gate), des moteurs de recherche (SolR, Antidot, Exalead), et des triplestores RDF (Sesame, Oracle RDF).

02/2003 – 09/2017

### ARCHITECTE LOGICIEL JEE MONDECA

---

Gestion d'une équipe de 4 développeurs chez Mondeca, pour développer ITM et d'autres logiciels. Analyses fonctionnelles et techniques, conception d'architectures JEE, gestion d'équipe, intégration de composants externes (text-mining, moteurs de recherche). Chef de projet technique pour les premiers projets internationaux de Mondeca (USA, UK, Belgique). Contribution à divers projets d'ingénierie documentaire pour l'édition, l'industrie, l'e-tourisme et la défense.

## ÉDUCATION ET FORMATION

---

2000 – 2001

**INGÉNIEUR INFORMATIQUE** University of Pennsylvania

---

Echange universitaire d'une année entre l'UTC et l'University of Pennsylvania (Engineering school et Wharton Business school).

1998 – 2003

**INGÉNIEUR INFORMATIQUE** UTC - Université de Technologie de Compiègne

---

1998

**BACCALAURÉAT MENTION TRÈS BIEN**

---

## COMPÉTENCES LINGUISTIQUES

---

Langue(s) maternelle(s): **FRANÇAIS**

Autre(s) langue(s):

|                 | COMPRÉHENSION       |                      | EXPRESSION ORALE            |                   | ÉCRIT |
|-----------------|---------------------|----------------------|-----------------------------|-------------------|-------|
|                 | Compréhension orale | Compréhension écrite | Expression orale en continu | Interaction orale |       |
| <b>ANGLAIS</b>  | C1                  | C2                   | C1                          | C1                | C1    |
| <b>ALLEMAND</b> | A1                  | A2                   | A1                          | A1                | A1    |

---

Niveaux: A1 et A2: utilisateur de base; B1 et B2: utilisateur indépendant; C1 et C2: utilisateur expérimenté

## COMPÉTENCES NUMÉRIQUES

---

### Standards du web de données

RDF | OWL | SPARQL | SHACL | SKOS | PROV

### Modèles conceptuels et ontologies

FRBRoo | CIDOC-CRM | Wikidata | ORG | Records-In-Context | schema.org | Dublin Core | FOAF | EDM (Europeana) | XKOS | ELI (European Legislation Identifier)

### Outils et API du Web de Données

Apache Jena | RDF4J | Virtuoso | GraphDB | Oracle Semantic Technologies

### Développement web et JEE / Ingénierie logicielle

Design Patterns | SQL | JEE (de la base de données aux interfaces web) | HTML | CI / CD | XSLT | CSS | Spring framework | Linux Bash | Maven | Javascript | Typescript | JSP | Git | Eclipse | XML | Jira

## ● INFORMATIONS SUPPLÉMENTAIRES

---

### PROJETS

09/2021 – EN COURS

**Parlement Européen** La division KMS (*Knowledge Management Services*) du Parlement Européen est notamment en charge de la division des données Open Data du PE, et à ce titre doit modéliser et produire les jeux de données à diffuser sur le portail open data de l'institution.

Dans une première phase de preuve de concept j'ai été en charge de convertir des jeux de données décrivant les députés européens, les commissions du parlement, les résultats de votes.

Une fois cette preuve de concept réalisée, montrant que la constitution d'un graphe de connaissances connecté était possible, la seconde phase a consisté à modéliser, sur la base de l'ontologie ELI+ELI-DL, et à convertir les notices de métadonnées documentaires issues du système actuel.

La troisième phase a consisté en la modélisation et la conversion des agendas parlementaires et des résultats de votes détaillés, amendements par amendements.

La phase actuelle implique des travaux sur la reprise de l'antériorité des résultats de vote, et la conversion du règlement du Parlement

**Liens** <https://data.europarl.europa.eu/fr/developer-corner> | <https://github.com/europarl/open-data-beta-testing>

07/2021 – 04/2022

**Organisation Mondiale du Commerce (OMC)** L'OMC s'occupe des règles du commerce international entre ses états membres. La division LDSD (*Language and Documentation Services Division*) de l'OMC est en particulier responsable de l'application [Documents OnLine \(DOL\)](#) qui permet de rechercher et d'accéder aux publications officielles de l'institution.

LDSD souhaitait remplacer l'actuel DOL par un moteur de recherche sémantique, basé sur un graphe de connaissances. Le projet a consisté à nettoyer, structurer et convertir les vocabulaires contrôlés déjà existants, et à formaliser d'autres vocabulaires contrôlés.

- formalisation d'un thésaurus spécifique au commerce international pour l'OMC (environ 900 concepts, traduits, sourcés, partitionnés, hiérarchisés)
- Mise au point d'une taxonomie de produits alignée sur le Harmonized System de la WCO a été formalisée.
- Modélisation de la structure du graphe de connaissances (documents, agents, évènements, concepts)
- Rédaction d'un appel d'offre pour un outil de gestion des vocabulaires contrôlés

05/2020 – EN COURS

**Agence du Numérique en Santé : modélisation de terminologies médicales et conversion de données** L'Agence du Numérique en Santé (ANS) a pour mission de promouvoir l'interopérabilité et les technologies numériques dans le domaine de la santé en France.

L'ANS contribue au renforcement de l'efficacité des politiques de santé et à l'amélioration de la coordination, de l'efficience et de la qualité des soins. Elle détient un large portefeuille de projets, notamment pour créer un cadre pour le développement de l'e-santé ; dans ce contexte, l'ANS vise à assurer l'interopérabilité des systèmes informatiques de santé, en publiant un [serveur multi-terminologies \(SMT\)](#).

J'assiste l'ANS sur la modélisation ontologique et la conversion de données de grandes terminologies médicales diffusées sur le serveur multi-terminologie :

- [Référentiel national d'interopérabilité du médicament](#) (ANSM + CIP + Medicabase)
- [CCAM](#) (codification des actes médicaux)
- [LOINC](#) (terminologie des tests de laboratoire)
- [CIM-10](#) et [CIM-11](#) (classification internationale des maladies de l'OMS)

**Lien** [https://smt.esante.gouv.fr/explorer-les-concepts/terminologie-ref\\_interop\\_med](https://smt.esante.gouv.fr/explorer-les-concepts/terminologie-ref_interop_med)

09/2015 – 12/2018

**Service Central de Législation du Luxembourg : Consulting Legilux / Casemates** Le [Service Central de Législation](#) (SCL) du Luxembourg est chargé de la publication du Journal Officiel du Grand-Duché de Luxembourg et de la consolidation de la législation. Dans le prolongement de l'initiative ELI, il a souhaité mettre en place un nouveau portail de diffusion de la loi : [legilux.public.lu](http://legilux.public.lu). Ce portail est entièrement basé sur des données sémantiques en RDF, distribuées via la plateforme [Casemates](#).

La phase 1 du projet a consisté à spécifier la nouvelle architecture pour la diffusion des données législatives à travers Casemates. Le système comprend un dépôt de documents (HTML, PDF et XML des textes législatifs) associé à un triplestore RDF, basé sur un modèle dérivé de FRBR et proche de l'ontologie ELI. Le livrable de cette phase était le cahier des charges de l'appel d'offre pour la réalisation de Casemates.

Suite au développement de Casemates, j'ai accompagné le SCL en phase 2 dans la spécification des fonctionnalités attendues dans le nouveau portail Legilux. L'architecture finalement retenue est un couplage RDF+ElasticSearch.

La phase 3 a consisté à mettre en place la plateforme [Skosmos](#) pour la diffusion de l'ensemble des vocabulaires contrôlés utilisés dans la description des lois luxembourgeoises.

Lien <https://legilux.public.lu/>

08/2015 – EN COURS

**Taskforce ELI : modélisation et expertise technique sur le web sémantique** L'initiative européenne [ELI \(European Legislation Identifier\)](#) est une application des principes du web de données pour la diffusion et la mise en relation des lois publiées par les États membres de l'UE. L'objectif est d'améliorer la visibilité et l'accès à la législation sur le web.

J'apporte une expertise technique au groupe de travail ELI sur des aspects variés :

- Modélisation de l'[ontologie ELI](#) (extension FRBRoo).
- Modélisation de l'[ontologie ELI-DL](#) (extension CIDOC-CRM) ;
- Formation et assistance aux États membres sur le web de données et l'ELI (ateliers en Irlande, Italie, France, Luxembourg, Danemark, Finlande, Autriche, Malte).
- Développement de logiciels pour aider à la publication des métadonnées ELI : Valideur ELI, basé sur SHACL.
- Définition et proposition d'une [extension "Législation"](#) au vocabulaire schema.org, accompagnée d'un [guide de bonnes pratiques](#) pour la description des lois en utilisant schema.org.
- Rédaction de guides techniques de bonnes pratiques pour l'implémentation d'ELI et ELI-DL
- Analyse et spécifications techniques sur divers sujets : Formats XML (Akoma Ntoso), diffusion d'ELI en open-data, utilisation de JSON-LD, comparaison avec ECLI

Lien <https://eur-lex.europa.eu/eli>

10/2013 – EN COURS

**Office des Publications de l'Union Européenne : consulting et développement**

L'[Office des Publications de l'Union Européenne](#) (basée à Luxembourg) est en charge de l'édition des documents des institutions de l'Union Européenne (UE), en particulier de la publication du Journal Officiel de l'UE.

- Mission de conseil (15 jours) pour la spécification de patterns d'IRI interinstitutionnels pour les projets de loi, les événements et les documents associés (rapports, opinions, amendements, etc.).
- Mission de développement puis conseil (50 jours) pour la spécification et le développement d'un prototype pour l'identification, la description et l'accès aux métadonnées des subdivisions de la législation.
- Mission de conseil (50 jours) pour l'implémentation d'ELI (European Legislation Identifier) dans le portail EUR-Lex : analyse d'impact dans les différents systèmes informatiques, rédaction de spécifications détaillées, création de prototypes.
- Mission de développement (60 jours) sur site et hors site, pour développer une interface de visualisation et de modification des données dans le Cellar, la base centrale de métadonnées au cœur du processus de production de l'OP.
- Conversion des données du thésaurus Eurovoc en profil d'application Euvoc (30 jours).

2017 – 2020

**ISSN International Center : formation, modélisation, synchronisation avec Wikidata**

Le [Centre International de l'ISSN](#) est une organisation internationale qui attribue et gère l'identification et la description des ressources périodiques (revues scientifiques notamment).

L'ISSN souhaitait accroître ses compétences en web de données, modéliser et documenter la structure des données publiques de son nouveau portail. La mission consistait, après une formation initiale de 2 jours, à formaliser et documenter le [profil d'application de l'ISSN](#), en intégrant des propriétés de schema.org, Dublin Core, BIBO, MARC21, FOAF, et d'autres modèles. Cela incluait la spécification de la diffusion de ces métadonnées à l'aide du vocabulaire schema.org par le biais du portail.

En 2020, une analyse et le développement du logiciel [ISSNBot](#) ont été réalisés pour synchroniser les métadonnées des périodiques du registre de l'ISSN avec [Wikidata](#). L'objectif est d'améliorer la visibilité des données ISSN sur le web.

Lien <https://www.issn.org/fr/comprendre-lissn/regles-d-attribution/profil-dapplication-de-lissn-pour-les-donnees-liees/>

2015 – 2016

**UNESCO : maintenance et publication du thesaurus en SKOS** L'UNESCO dispose d'un [thesaurus](#) en 5 langues (français, anglais, espagnol, russe, arabe), de 4400 concepts. L'institution a souhaité déployer un nouveau système de gestion du thesaurus en le rendant compatible avec les standards du web de données. L'objectif est de valoriser le thesaurus, d'en permettre une réutilisation plus aisée, et de le traduire via un outil coopératif avec un workflow de validation.

Nous avons proposé une réponse au cahier des charges n'utilisant que des composants open-source (VocBench, Skosmos, Jena, Skos Play), permettant ainsi de construire un système à l'état de l'art et à moindre coût. Plus de détails peuvent être trouvés dans [cet article](#).

Lien <https://vocabularies.unesco.org>

---