

Semantic_forms

Un couteau suisse
pour le Web Sémantique

orienté
édition des données
et construction d'applications

J.M. Vanel

V 4.0

Semantic_forms en bref

- application générique de navigation du **LOD (Linked Open Data)**
 - les données et modèles de données s'appuient sur les recommandations du **Web Sémantique du W3C**.
- annotation, édition de données structurées
- cadre (framework) pour construire des **applications d'entreprise** centrées sur les formulaires et SPARQL
 - Cf Ruby on Rails, Django, Symfony, etc
 - un pas vers le Système d'Information Sémantique, qui va remplacer les PGI (**Progiciel de Gestion Intégré**, ERP)

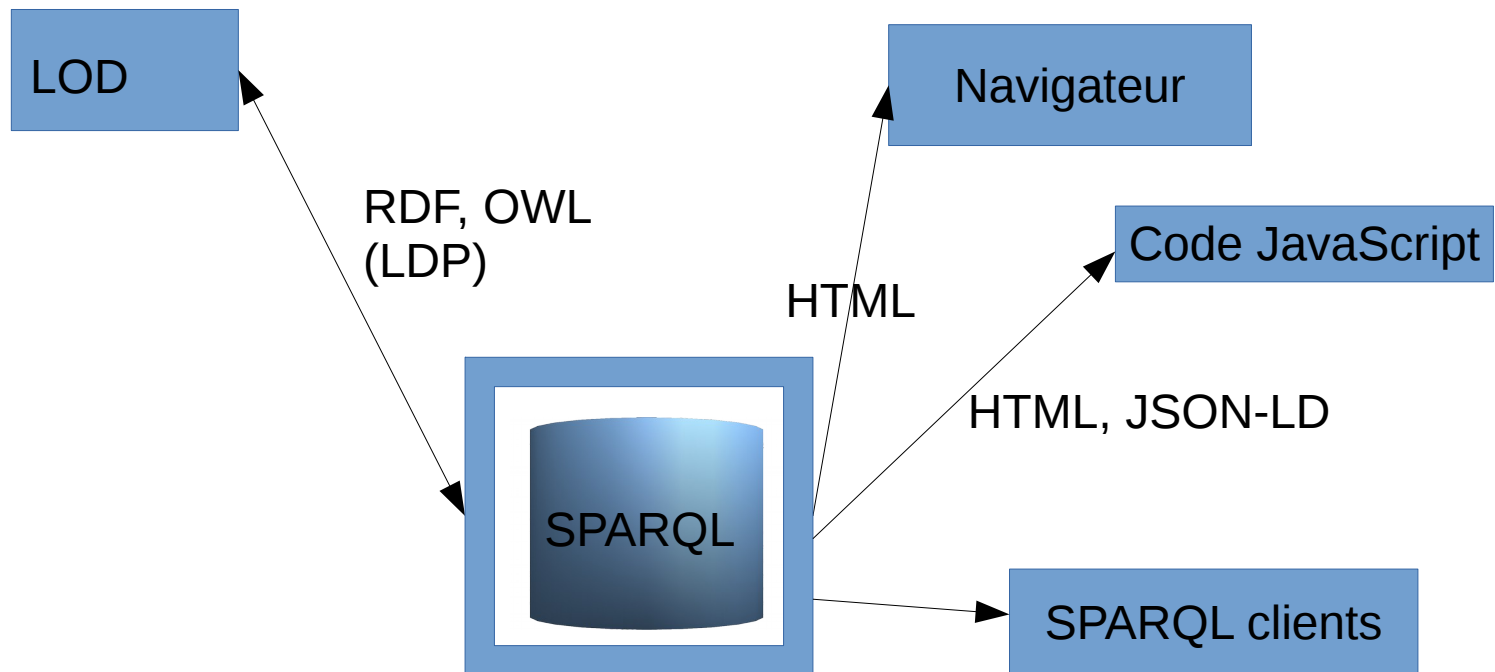
Le Web Sémantique dans SF

Standards du W3C

- pour la navigation de proche en proche : RDF
- pour la modélisation : OWL
- pour la base de données : SPARQL
- Pour le REST ;) LDP

⇒ tout se retrouve dans la base SPARQL de SF

SF : big picture



Semantic_forms

- Le coeur: un générateur de formulaire robuste, générique, en temps réel
 - les ontologies guident la saisie
 - Spécification de formulaire = liste de propriétés RDF
- visualiser et éditer RDF de manière collaborative
- solution Open Source simple (13 000 lignes)
 - distribution zip, seule dépendance Java 8

Formulaires : automatique ou spécification

- Inférence via `rdfs:domain`, `rdfs:range`, `rdf:type`, `rdfs:subclassOf`
- Spécification de formulaire
 - Fixer l'ordre des champs
 - Détails d'un champs: lookup dbPedia

```
forms:personForm
```

```
  a form:specification ;
```

```
  form:classDomain foaf:Person ;
```

```
  form:showProperties (
```

```
    foaf:givenName foaf:familyName
```

```
    foaf:homepage
```

```
    foaf:mbox          foaf:phone          foaf:img
```

```
    foaf:topic_interest          foaf:knows
```

```
    foaf:currentProject          cco:expertise
```

```
  ) .
```

Formulaires flexibles

- Aspects wiki:
 - création d'URI,
 - contenu HTML
- Multi valeurs
- Multi type,
- typage dynamique

Fonctionnalités

- Recherche plein texte dans la base SPARQL avec Lucene
- Complétion par rapport à dbPedia pour compétences, etc
- Historique des actions par utilisateur
- tracé de graphes avec RDFViewer
- console d'administration SPARQL (à la PHPMyAdmin)
- divers algorithmes: dédoublonnage, restructuration d'ontologie

Sous le capot

La pile technologique

- Play! Framework,
- application Web générique avec Play! framework
- librairie de génération de formulaires (API Java et Scala)
- cache SPARQL,
 - utilise un graphe nommé par source de source de données, plus un par utilisateur
- langage Scala et librairie Banana-RDF.
 - Grâce à Banana-RDF, semantic_forms peut utiliser d'autres bases SPARQL, en embarqué (via API) comme BlazeGraph ou Sesame, ou n'importe quelle base via HTTP.
- Base SPARQL : Jena TDB (en embarqué) + Lucene

Architectural

- Le cache sémantique : agrégation de documents RDF hébergés ailleurs, lors de la navigation
 - le cache, via LDP, peut être vu comme un composant d'infrastructure qui fédère des données de différentes provenances
 - Implémentation partielle LDP : expose les documents RDF via HTTP GET
- Session utilisateur via Play! framework
- Service web renvoyant un formulaire nu
 - créer des applications 100 % HTML + JavaScript
 - `/form?displayuri=http://jmvanel.free.fr/jmv.rdf#me`
 - `/form?formuri=http://deductions-software.com/ontologies/forms#personForm`

Sémantique > SQL, noSQL, OO

- Par rapport à SQL
 - Cardinalité flexible, requêtes fédérées
 - propriétés exceptionnelles possibles
- Par rapport à no-SQL (orienté document)
 - Pas besoin de lire-écrire un gros JSON pour changer un champ
 - Standard du W3C solides et nombreuses implémentations
- Par rapport à l'Orienté Objet
 - Les propriétés ne sont pas enfermées dans les classes
 - Inférences
 - UML mélange implémentation et modélisation des données; le standard UML a toujours été instable

En chantier

- Améliorations ergonomie
- Calcul de proximité sur le graphe SPARQL :
moteur de recommandation générique
exploitant les données dbPedia
- Complétion en Français dans tous les champs
de saisie via Lucene
- Assistant de sémantisation pour CSV, XML,
SQL, JSON

Conclusion

- Plutôt que courir après les données SQL, texte, etc,
 - il faut créer du sémantique;
- rôle central de dbPedia / Wikidata
 - référentiel de connaissances
 - recommandation via parcours de graphe
- Place au SI sémantique : CRM, CMS, ERP, ...