

Interfaces des services Mostrare

Matthieu Keith & Fabien Torre

Équipe INRIA Futurs Mostrare
<http://mostrare.futurs.inria.fr/>

5 février 2008

Table des matières

- 1 Interfaces des services Formating
- 2 Interface StructuredInformationExtraction (service Mostrare)

Formating

Formating est un service générique regroupant les services dont la finalité est de produire un document dans le format pivot WebContent à partir d'un document.

ConfigurableService (méthodes)

Formating implémente l'interface *ConfigurableService* pour fixer les ressources à utiliser par le service.

Méthodes offertes

- configure(userId : URI, config : Annotation)
- resetConfiguration(userId : URI)

Chaque service doit préciser la syntaxe de sa configuration (sous forme de RDFS par exemple).

Formating (méthodes)

Méthodes offertes

- `processRawDocument(doc : RawDocument, userId : URI) : Document`
- `processURI(docuri : String, userId : URI) : Document`
- `processWCdocument(doc : Document, userId : URI) : Document`
- `configure(userId : URI, config : Annotation)`
- `resetConfiguration(userId : URI)`

Chaque service doit préciser la ou les version(s) de *process* qu'il implémente.

ConfigurableService (configure)

configure(userId : URI, config : Annotation)

- *description* : définit la configuration du service pour les utilisations ultérieures de l'utilisateur.
- *entrées* : une URI identifiant l'utilisateur et une annotation décrivant les ressources à utiliser pour le traitement
- *sorties* : aucune.
- *pré-condition* : l'annotation respecte le format spécifique du service et contient des URI désignant effectivement des ressources utilisables par le service. L'utilisateur est enregistré par le service.
- *post-condition* : le service est prêt à utiliser les ressources.

ConfigurableService (resetConfiguration)

resetConfiguration(userId : URI)

- *description* : restitue la configuration du service à sa valeur par défaut, éventuellement à l'état "non configuré".
- *entrées* : une URI désignant l'utilisateur.
- *sorties* : aucune.
- *pré-condition* : l'URI désigne effectivement un utilisateur enregistré par le service.
- *post-condition* : la configuration est ré-initialisée pour l'utilisateur.

Formating (processRawDocument)

```
processRawDocument(doc : RawDocument, userId : URI) :  
Document
```

- *description* : produit un document dans le format pivot WebContent à partir d'un document du web et des ressources définies dans la configuration antérieure.
- *entrées* : un document dans un format à préciser et l'URI d'un utilisateur.
- *sorties* : un document au format pivot WebContent reprenant des éléments du document d'origine.
- *pré-condition* : document dans le format attendu, ressources configurées préalablement (méthode *configure*) et utilisables par le service.
- *post-condition* : un document au format pivot WebContent est produit.

Formating (processURI)

processURI(docuri : URI, userld : URI) : Document

- *description* : produit un document dans le format pivot WebContent à partir d'un document du web pointé par son URI.
- *entrées* : les URIs d'un document et de l'utilisateur.
- *sorties* : un document au format pivot WebContent.
- *pré-condition* : URI valides et utilisables.
- *post-condition* : un document au format pivot WebContent est produit.

Formating (processWCdocument)

`processWCdocument(doc : Document, userId : URI) : Document`

- *description* : produit un document dans le format pivot WebContent à partir d'un document déjà dans le format pivot.
- *entrées* : un document dans le format pivot WebContent et faisant référence à sa source, URI de l'utilisateur.
- *sorties* : un document au format pivot WebContent. du document d'origine.
- *pré-condition* : présence la méta-donnée `dc:source` dans le document, URI valides et utilisables.
- *post-condition* : un document au format pivot WebContent est produit.

StructuredInformationExtraction (méthodes)

Méthodes offertes

- `configure(userId : URI, config : Annotation)`
- `resetConfiguration(userId : URI)`
- `processRawDocument(doc : RawDocument, userId : URI) : Document`
- `processURI` : non fournie
- `processWCdocument` : non fournie

StructuredInformationExtraction (processRawDocument)

`processRawDocument(doc : RawDocument, userId : URI) : Document`

- *description* : réalise l'extraction sur le document fourni.
- *entrées* : un document au format XML et l'identifiant d'un utilisateur enregistré par le service.
- *sorties* : un document au format pivot WebContent.
- *pré-condition* : le document fourni doit respecter les règles XML, et les URIs définies dans les annotations doivent désigner des wrappers existants sur le serveur.
- *post-condition* : un nouveau document WebContent est produit, il contient une table dont chaque ligne correspond à un tuple de la relation, tuple dont les composantes ont été trouvées dans le document d'origine.