

TP Conception logicielle – Scrum 2025



Parseur d'articles scientifiques en format texte (carnet du produit)

*** Sprint 3

Suite à une discussion au LIA, nous avons décidé que les fichiers parsés plein texte ne suffisent plus à nos besoins. Nous avons entendu parler que les fichiers structurés XML possèdent plusieurs avantages et qu'ils sont largement utilisés en TAL. Nous préférons donc comme sortie un fichier XML utf8 avec la structure suivante:

```
<article>
  <preamble> Le nom du fichier d'origine </preamble>
  <titre> Le titre du papier </titre>
  <auteur> La section auteurs et leur adresse </auteur>
  <abstract> Le résumé de l'article </abstract>
  <biblio> Les références bibliographiques du papier</biblio>
</article>
```

Cependant quelques chercheurs veulent garder la sortie texte aussi... Cette nouvelle sortie .xml complète la sortie texte .txt du sprint précédent.

Ainsi, au moment de lancer votre parseur, il faudra ajouter un argument (**-t** | **-x**) pour bien choisir le type de sortie : soit générer des fichiers dans une version texte (**-t**), soit au format XML (**-x**) :

Groupe (**Torres**) 1: **avant sprint 4**

Le système final, et les versions intermédiaires doivent être placés sur Github ; chaque version dans une branche différente. Le fichier README.md doit contenir une explication générale du système et la procédure nécessaire pour lancer les programmes. Ainsi il faut documenter tous les différents événements de la méthode Scrum en ajoutant sur ENT des images de votre travail (photos de réunions de planification de sprint, mêlée quotidienne, revue de sprint, rétrospective du sprint) ainsi que captures d'écran et ou des documents écrits à la main (dessins, idées, etc.).

Consignes

- Équipes : fixée en séance 1
- Méthodologie agile : SCRUM
- Plateforme : le projet doit fonctionner sous GNU/Linux en ligne de commande
- Aucune conversion en ligne via navigateur n'est acceptée
- Langage : fixé en séance 2
- Dépôts : sur ENT et GIT (lien envoyé par mail à votre prof)
- Soyez agiles !



Vous êtes en compétition ! Attention aux fuites des idées ou de logiciel !