

Compte rendu de réunion et Ordres du jour Projet tutoré 2015

Roland Bary, Charles Follet

Encadrants : Marie-Noelle Bessagnet, Annig Lacayrelle, Albert
Royer, Christian Sallaberry

Contents

1	Jeudi 5 Février	2
2	Lundi 9 Février 2015	2
	2.1 Paramétrage de GATE	2
	2.2 Corrections de bugs	3
3	Mardi 03 Mars 2015	4

1 Jeudi 5 Février

■ Discussion :

- Rencontre entre l'équipe d'encadrants et les groupes associés aux projets.
- Présentation brève des cahiers de charges relatifs aux différents projets et de l'environnement d'annotation thématique GATE.



♣ Tâches à réaliser pour le 09/02/2015 :

- Installer l'environnement GATE et effectuer quelques tests pour se familiariser avec.
- Installer le logiciel TreeTagger et effectuer quelques tests pour se familiariser avec.
- Lire la documentation de GATE jusqu'à la partie 6.
- Lire des ressources documentaires de chaque projet.



2 Lundi 9 Février 2015

2.1 Paramétrage de GATE

Convention de nommage :

GATEdir : Chemin d'installation de GATE.

TREETAGGERdir : Chemin d'installation de TreeTagger.

ANNIE

- Lancer le CREOLE plugin manager en cliquant sur le logo CREOLE (puzzle).
- Cocher la case Load Now et Load Always du plugin ANNIE.
- Cliquer sur Apply All.
- Fermer le plugin manager.

Unicode Tokenizer

Dans GATE :

- Clic droit sur Processing Ressource dans l'arborescence de gauche.
- New GATE Unicode Tokenizer avec encoding : utf8, rulesURL :
GATEdir/plugins/Lang_French/tokeniser/FrenchTokeniser.rules

Inclusion TreeTagger

- Éditer

```
GATEdir/plugin/Tagger_Framework/resources/TreeTagger/treetagger-french-gate
```

- Ajouter/modifier les lignes suivantes

```
BIN=TREETAGGERdir/bin  
LIB=TREETAGGERdir/lib  
CMD=TREETAGGERdir/cmd
```

Récupérer une application

- Clic droit sur Applications puis Restore application from file...
- sélectionner

```
GATEdir/plugins/Lang_French/french+tagger.gapp
```

- Dans French_NE : clic gauche sur Treetager-FR-NoTokenization
- modifier encoding: utf-8 et taggerBinary :

```
GATEdir/plugins/Tagger_Framework/resources/TreeTagger/tree-tagger-french-gate
```

2.2 Corrections de bugs

GAWK

installer gawk :

```
sudo apt-get install gawk
```

Renommage

Copie de french-utf8.par en french.par dans le dossier

```
TREETAGGERdir/lib
```

LuckyException

Dans

```
GATEdir/plugin/Lang_French/tokeniser/postprocess.jap
```

3 Mardi 03 Mars 2015

■ Discussion :

Il a été au cours de cette réunion, de présenter le travail que nous avons effectué durant les vacances du mois de février :

- Définition des éléments d'information ciblés, par une ébauche de construction de règles d'annotations. Les éléments ciblés étant les suivants :
 - Spatiale: villes, départements, villes de pays frontaliers
 - Temporelle: période d'émission des monnaies
 - Thématique: nature de la monnaie, légende du droit et légende du revers, le type de la monnaie, le nom de l'atelier, personnages associés aux monnaies
- Construction d'un premier Gazetter à partir des nom de villes et départements de France récupérés sur le site web de L'INSEE
- Tests d'annotation sur un fichier d'extension .txt (Tiré de la ressource documentaire) avec le dit gazetter constitué.

Aussi plusieurs remarques ont été mises en évidence, à savoir :

- L'amélioration de l'ébauche de construction des règles d'annotation qui a été présentée, par l'ajout des thématiques 'collection', 'Trésors' et 'trouvailles' et
- Le traitement en priorité cas généraux d'annotation et noter les cas spécifiques rencontrés
- Formaliser aux mieux l'information, à l'aide de notions bien connues (i.e Expression Régulières, automates, ...)
- Constitution de Gazetter à partir d'informations directement extraites de la ressource documentaire



♣ **Tâches à réaliser pour le :** A l'issue de cette réunion il nous a été recommandé de réaliser les tâches suivantes pour la prochaine rencontre:

- Scan et Océrisation des pages constituant les chapitres I à III, afin d'extraire des informations nécessaires à la constructions de Gazzeters
- Améliorer l'ébauche de construction des règles d'annotation, par des illustrations, et des notes de problèmes éventuellement rencontrés
- Être capable de retrouver les noms des ateliers, les périodes, les collections, trésors, trouvailles...



♣ **Ordre du jour du :**

- Présentation
-
-