

**PROJET C-MANTIC**

**Spécification Fonctionnelles**



**SUIVI DU DOCUMENT**

|  |  |  |  |
| --- | --- | --- | --- |
| Mises à jour | | | |
| Version | Date | Auteurs | Objet de la mise à jour |
| 1.0 | 30/05/2013 | Julien BRETON | Version initiale |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

|  |
| --- |
| **Liste de diffusion** |
| Mr. SANANES  Florian CONSTANT  Christophe BARRETO  Julien BRETON |

Sommaire

[1.1 Objet du document 4](#_Toc357844675)

[1.2 Enjeux et Objectifs 4](#_Toc357844676)

[1.3 Périmètre fonctionnel 4](#_Toc357844677)

[1.4 Cadre technique 4](#_Toc357844678)

[1.5 Description générale 5](#_Toc357844679)

[1.6 Planification du projet 6](#_Toc357844680)

# Objet du document

L’objet de ce document est de définir les spécifications fonctionnelles détaillées de l’application C-mantic.

Les spécifications fonctionnelles ont pour but de décrire précisément l’ensemble des fonctionnalités de l’application.

Toutes les fonctionnalités prévues lors de la phase de conception sont précisées dans ce document en indiquant l’implémentation de ces fonctionnalités dans l’application.

# Enjeux et Objectifs

Avec l'essor du Web, la production et la diffusion de documents se sont développées de manière exponentielle. La pratique du Copier/Coller se standardise sans pour autant faire référence aux auteurs.

Il existe même des sites payants (ex : http://www.academon.fr/) qui proposent leurs services de rédaction de documents. Ce sont des sites frauduleux qui fournissent des travaux toutes disciplines confondues. Le Copier/Coller devient un jeu d'enfants et les fraudeurs sont

nombreux.

Il est relativement facile de repérer si une production est plagiée ou pas. Il suffit de faire attention aux changements de tons du discours pour déceler si le texte a bien été écrit ou pas par l'auteur. Ensuite, pour repérer la source du texte plagié, on peut toujours utiliser les moteurs de recherche mais cela nécessite un certain temps que l'enseignant ne trouve pas forcément lorsqu'il doit corriger un grand nombre de copies.

Ainsi, l'objectif de l'application C-mantic est de pouvoir disposer d'un outil qui permette de détecter le plagiat à partir d’une source donnée sous forme de fichier.

# Périmètre fonctionnel

L’application C-mantic est principalement constituée d'une interface graphique à destination de l'utilisateur. A partir de cette interface, il sera possible d'effectuer les actions suivantes :

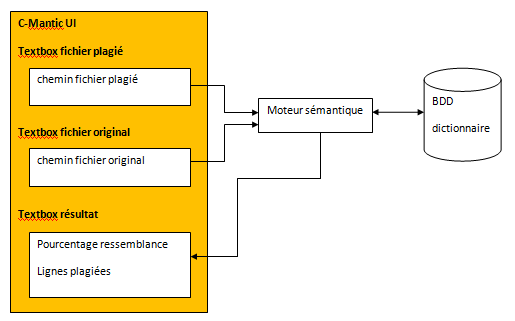
* Comparer deux textes : D'un côté le texte original et de l'autre celui dont on soupçonne d'avoir été plagié
* Lancer le traitement pour vérifier si le texte a effectivement été plagié ou non
* L'utilisateur pourra consulter en résultat :
* Le pourcentage de ressemblance entre les deux textes (plagié et original)
* Le texte complet qui a été vérifié avec la mise en évidence des phrases qui ont été plagiées à l'aide d'une coloration syntaxique particulière

# Cadre technique

L’application C-mantic reposera sur les outils suivants :

* Une interface graphique qui reposera sur la librairie graphique Qt
* Une BDD de type MySQL qui sera le « dictionnaire » de l’application
* Le moteur de sémantique codé en C++ qui accèdera à la base de données

# Description générale



Comme l'illustre la figure ci-dessus, l'application C-mantic est constituée de plusieurs modules:

* Le module d'interface graphique qui comprend les fonctions de saisie des noms de fichiers à comparer ainsi que l'affichage des résultats obtenus
* Le module de traitement constitué du moteur sémantique et de la BDD dictionnaire

Dans un premier temps, les deux fichiers sont pris en compte par le moteur sémantique à partir de leur chemin absolu. Après quoi, le moteur va parser les 2 fichiers en se débarrassant des mots trop communs telles que les prépositions (le, la, les, des, ...) par exemple. Ainsi, seuls les mots les plus importants seront conservés sous forme de liste.

Le texte sera décomposé sous forme de liste de listes:

* 1 ligne = 1 phrase
* 1 phrase = 1 liste de mots successifs

Les mots restants dans les listes seront ensuite comparés que ce soit en terme de temps, de synonymes et de longueur. Ainsi, dès qu'un mot sera jugé être un plagiat, il sera ajouté à une liste similaire à celle vue ci-dessus. Cette liste sera ensuite exploitée afin de calculer le taux de plagiat du fichier mais aussi pour faire ressortir les phrases du texte concernées par le plagiat sur l'interface graphique.

Au final, l'utilisateur pourra consulter le texte dans sa totalité depuis l'interface graphique avec les lignes plagiées mises en évidence à l'aide d'une coloration syntaxique particulière.

# Planification du projet

