

# MEHDAOUI Abdlemaine

Projet Technologique

October 2018

## 1 Introduction

Dans le cadre de l'unité d'enseignement : Projet technologique et afin d'avoir un aperçu sur un projet informatique, dès son début jusqu'à sa livraison, du coup, d'après ce projet nous avons eu l'occasion de réaliser un projet informatique qui répond à une problématique cruciale c'est celle de la compréhension du langage naturel via un forum de discussion pour enfants, ainsi qu'avec toute une démarche préalable avant de réaliser concrètement le projet telle que la recherche d'un module qui va répondre exactement au besoin d'analyser et de filtrer le langage naturel. Pour cela on va opter pour un cycle qui comprend premièrement le cahier des charges ou l'état de l'art c'est à dire les spécifications généraux du site (forum, blog...) deuxièmement une solution technique efficace pour le traitement du langage naturel pour protéger l'enfant de toutes sortes d'arcelement tout en les laissant communiquer entre eux.

## 2 Description général

### 2.1 Présentation du projet

L'objectif principal du projet est de fournir à des utilisateurs (enfants) la possibilité de communiquer entre eux à travers un forum de discussion en réseau. Le projet final doit principalement permettre aux utilisateurs de lire leurs nouveaux messages et d'en poster d'autres. En plus de ça un traitement et un filtre sur tous les messages reçus ou envoyés.

### 2.2 Présentation des acteurs

On peut identifier plusieurs acteurs au cours de notre projet

**Administrateur :** Il devra paramétrer et lancer le serveur. Ce dernier étant autonome, le travail de l'administrateur se restreint à la surveillance du bon fonctionnement du serveur. De plus, s'il a besoin d'aide pour la gestion du forum (pour la suppression des messages indésirables), il peut permettre à certains utilisateurs de supprimer ces messages. Ce sont, ce qu'on définira plus tard, des modérateurs. Il peut aussi, si un modérateur le lui demande, statuer sur la bonne utilisation d'un utilisateur du forum, et s'il le souhaite, le supprimer.

**Modérateur :** Il utilise le même client qu'un utilisateur standard, à la différence que le modérateur possède des droits étendus. Ainsi, il surveille le contenu des différents messages postés sur le forum et veille au respect des règles d'utilisation imposées par l'administrateur. Il peut, le cas échéant, supprimer les messages jugés indésirables sur le forum (ça peut être le parent dans ce cas).

**Utilisateur :** Il utilise le client. Il peut en utiliser les fonctionnalités de bases comme poster, lire, ou rechercher dans les archives. Selon l'identifiant utilisé pour se connecter à l'application, les droits d'administrations s'appliquent en conséquence. Par exemple, si l'identifiant d'un modérateur est reconnu, il aura accès à ses fonctionnalités spécifiques. Il faut préciser qu'aucun utilisateur non enregistré ne peut accéder aux messages du forum. Néanmoins, un formulaire dans l'application client permet de s'enregistrer de façon simple et rapide. Il aura alors des droits d'utilisations de base (utilisateur).

### 3 Contraintes générales

1. Contraintes de délais: Ce document doit être mené à bien pour début November au plus tard. Les programmes client et serveur ainsi que leur code respectif devront être rendus, quant à eux, au début Avril
2. L'ergonomie: le logiciel doit pouvoir être utilisés avec le maximum de confort, de sécurité et d'efficacité par le plus grand nombre d'utilisateurs. Utiliser les technologies les mieux adaptées pour chacun des modules : le choix de la technologie aura une incidence sur la rapidité, l'efficacité, ainsi que sur le résultat donné par le logiciel et donc sur la qualité. Dans la description générale, nous avons clairement identifié les objectifs de l'application. A présent nous allons nous appuyer sur les objectifs afin d'en exprimer les besoins. Les besoins nous permettront de clarifier les fonctionnalités de l'application

## 4 Expression des besoins

Dans cette partie, nous allons présenter les fonctionnalités attendues pour le projet. En prenant en compte les besoins des utilisateurs, nous allons étudier dans un premier temps les besoins fonctionnels (opérations sur les messages), puis nous mettrons en relief les besoins non fonctionnels (contraintes graphiques). Nous prendrons en considération les paramètres inhérents à ces besoins afin d'améliorer la qualité du projet.

### 4.1 Besoins fonctionnels

Maintenant, nous allons analyser les opérations dont bénéficie l'utilisateur de base. Ces opérations se faisant sur des articles (archivés ou non), il nous faut préciser ce qu'est un article. Un article est un texte répertorié par un numéro d'identification unique permettant son archivage. En effet, la première partie d'un article est constituée de son entête, elle-même composée de l'identité de l'auteur, de la date de création et du sujet du message. Il est mis en ligne (sur le serveur) par un client.

**Système d'enregistrement:** Chaque nouvel utilisateur peut, par le biais du client, s'enregistrer auprès du serveur. Ainsi, si un utilisateur souhaite participer au forum, lire des messages ou bien même en poster, il peut, au démarrage du client, s'enregistrer tel un nouvel utilisateur(ici vaut mieux que l'enregistrement soit avec le mail d'un des deux parents). Il aura alors des droits de base. Il faut tout de même préciser que chacun des champs du formulaire aura une longueur maximum à respecter. Celle-ci sera de l'ordre d'une centaine de caractères, et sera précisée au sein de l'interface. Si l'utilisateur tente d'entrer une chaîne de caractères trop longue, il verra que sa chaîne sera tronquée par l'interface.

**Système de reconnaissance du login:** Selon le login entré par un utilisateur désirant se connecter au serveur, le client est capable d'afficher une interface appropriée à l'utilisation que l'utilisateur va faire du client. Ainsi, si un administrateur est reconnu, il aura accès à des fonctionnalités disponibles sur son interface graphique que des utilisateurs de base n'auront pas. En fait, le fonctionnement de ce système est le suivant :

Envoi du login et du mot de passe à travers le réseau jusqu'au serveur

1. Si le login est reconnu comme étant enregistré sur le serveur et que le mot de passe est bon
  - Le serveur renvoie une réponse au client. Selon cette réponse, le client affiche l'interface graphique adéquate.
  - Toutes les fonctionnalités nécessaires sont mises en fonction. Elles sont prêtes à être utilisées par l'utilisateur
  - L'utilisateur peut alors commencer à se servir du forum.

bon

2. Si le login n'est pas reconnu, ou que le mot de passe n'est pas bon
  - Un message est affiché à l'écran et explique à l'utilisateur la nature de l'erreur.
  - Puis il est renvoyé à la première fenêtre lui demandant s'il veut se connecter ou bien s'enregistrer.

**Envoi d'un nouveau message** Chaque utilisateur peut, s'il le désire, poster un nouveau message dans le forum. Ainsi, un bouton est prévu dans l'interface graphique. Ce bouton ouvre une fenêtre avec un champ permettant la saisie du texte à mettre en ligne. Le client s'occupe de la mise en forme du message (avec l'écriture de l'en-tête au nom de l'utilisateur courant notamment). Le message est, une fois le bouton d'envoi pressé, envoyé par le réseau au serveur. Le message 10 est alors traité par ce dernier et est prêt à être lu par tous les utilisateurs du forum. Il sera d'ailleurs marqué comme non-lu .

**Suppression d'un message** Chaque utilisateur se voit offert la possibilité de supprimer les messages qu'il a posté lui même. Ainsi, il ne peut pas supprimer les messages postés par d'autres utilisateurs. Seuls les modérateurs et l'administrateur du forum peuvent supprimer les messages de leur choix, qu'ils soient les leurs où ceux d'un autre utilisateur. Cela implique une gestion des droits entre chacun des types d'acteurs du forum.

**Gestion des droits** L'administrateur possède par le biais de son interface personnalisée la possibilité d'établir des droits spécifiques pour chaque utilisateur. Ainsi, il peut permettre à un utilisateur de supprimer les messages de son choix (il devient modérateur). Néanmoins, un modérateur ne peut modifier les droits d'un autre utilisateur.

**Système de recherche des messages** Un utilisateur peut rechercher un message qui a été archivé. Il peut ainsi soit effectuer une recherche par le numéro de message (unique pour chaque message), soit effectuer une recherche en tapant un mot. Tous les messages comportant ce mot sont alors listés à l'écran

**Moteur de recherche avancé** Un utilisateur peut rechercher s'il le désire un message qui a posté sur le forum. Il peut ainsi soit effectuer une recherche par le numéro de message (comme précisé dans la fonction Système de recherche de message), soit effectuer une recherche en tapant un mot. C'est cette fonction que nous allons détailler ici. En effet, cette fonction va permettre à l'utilisateur de pouvoir rechercher un sujet suivant un mot qui sera soit dans le titre du sujet soit dans le corps selon le choix de l'utilisateur. Tous les messages comportant ce mot sont alors listés à l'écran.

**Système de sauvegarde du serveur** Afin d'assurer une certaine stabilité dans le temps le serveur fera des sauvegardes. Cela permettra ainsi de sauvegarder la base de données concernant les utilisateurs à savoir les login, mot

de passe et droit d'utilisateur ainsi que la base de donnée des sujet posté. On sauvegardera donc dans un fichier a par la liste des utilisateurs après qu'un nouvel utilisateur se soit enregistrer. Il en est de même pour la sauvegarde des sujet, après chaque nouveau sujet de discussion la liste des sujet sera sauvegarder

## 4.2 Besoins non-fonctionnels

Dans cette partie, nous allons préciser l'environnement dans lequel le projet s'inscrit et les langages utilisées

**Environnement** Notre application client-serveur s'inscrit fondamentalement dans un réseau internet. Il fonctionnera sur toute les plateformes sans exception.

**Outils de développement** Comme stipulé dans le sujet de projet on a choisi à travailler avec:

- Langages de programmation: PHP, JavaScript, Python
- Bibliothèques pour l'application web: DOM, Bootstrap, JQuery,PDO
- Bibliothèques logicielle NLU: Natural language toolkit(NLTK)

## 5 Étas de l'art et solution retuenus

### 5.1 Extraction de l'information:

L'Extraction d'Information ou EI (en anglais, Information Extraction ou IE) désigne une technologie récente qui vise à extraire et à structurer automatiquement un ensemble d'informations précises apparaissant dans un ou plusieurs documents textuels écrits en langue naturelle. Ces informations sont destinées à créer ou alimenter un entrepôt de données (appelé aussi banque de données) [Piazenza 1997]. La tâche d'extraction est réalisée grâce au remplissage de formulaires prédéfinis (template).

### 5.2 Utilité dans le projet

Laisser un écolier de primaire chatter seul sur le web pose quelques problèmes moraux et de sécurité. Sans parler de dépendance ludique dont il pourrait être victime, l'enfant risque de se faire pirater son compte, de livrer des informations confidentielles, de subir un harcèlement, d'entendre des propos violents ou choquants,du coup la toute première solution technique qu'on peut voir est d'utilisé une API dédié et qui va etre on open source (contrairement à l'api google)

### 5.3 IBM Cloud API est une éventuelle solution

Avec cette API diverses caractéristiques du contenu du texte à l'échelle. Fournissez du texte, du code HTML brut ou une URL publique et IBM Watson Natural Language Understanding vous donnera les résultats des fonctionnalités demandées. Le service nettoie par défaut le contenu HTML avant l'analyse, de sorte que les résultats peuvent ignorer la plupart des publicités et autres contenus indésirables.

Vous pouvez créer des modèles personnalisés avec Watson Knowledge Studio pour détecter des entités et des relations personnalisées dans Natural Language Understanding.

Natural Language Understanding inclut un ensemble de fonctionnalités d'analyse de texte que vous pouvez utiliser pour extraire la signification de données non structurées.

### 5.4 Fonctionnement général de l'API

Pour faire de l'extraction de l'information et avec l'utilisation de cette API on peut procéder par différents façons autrement dit la manière dont l'information voulue est repérée pour cela on peut trouver plusieurs catégories:

**Concepts:** Renvoie les concepts de haut niveau dans le contenu. Par exemple, un document de recherche sur l'apprentissage en profondeur pourrait renvoyer le concept "d'intelligence artificielle", bien que le terme ne soit pas mentionné.

**Émotion:** Détecte la colère, le dégoût, la peur, la joie ou la tristesse véhiculés par le contenu ou par le contexte entourant les phrases cibles spécifiées dans le paramètre cibles. Vous pouvez analyser des émotions pour des entités détectées avec `entities.emotion` et pour des mots-clés avec `keywords.emotion`.

**Entités:** Identifie les personnes, les villes, les organisations et les autres entités du contenu. Voir Types et sous-types d'entités.

**Mots clés:** Renvoie les mots clés importants dans le contenu.

**Métadonnées:** Renvoie les informations du document, y compris le nom de l'auteur, le titre, les flux RSS / ATOM, l'image de la page principale et la date de publication. Prend en charge les types d'entrée URL et HTML uniquement.

## 6 Conclusion:

La chose qui nous a laissés penchés un petit peu sur cette API c'est la multitude des catégories de la façon d'extraction de l'information ainsi la possibilité de développer ça en python est qui est un outil à utiliser stipulé aussi dans la présentation du projet.