Créer des applications cognitives avec IBM Watson Engagement Explorer et Bluemix

Date : 10 avril 2015
Format : Conférence
Speakers : Philippe Comte, IBM

Watson comporte plusieurs technologies.
Les **frontières** entre l'humain et la technologie disparaissent peu à peu :

- Trans-humanisme chez Google
- Cœurs artificiels qui viennent compléter l'homme

Objectif de Watson : essaie d'améliorer le dialogue entre la machine et les humains.
En 1997, **DeepBlue** rivalisait avec les meilleurs joueurs d'échecs
Watson a récemment gagné face aux 2 meilleurs jours du **Jeopardy**. A nuancer tout de même : Watson n'a pas réponse à tout, il a été programmé pour ce jeu.

Depuis 2011, nous sommes rentrés dans l'ère de l'informatique cognitive : la machine prend des initiatives basés sur des statistiques, de l'entrainement … C'est le domaine du **Machine Learning**.
Watson permet de construire des robots assistants des hommes. L'équipe Watson d'IBM coopère avec la société Aldebaran (rachetée par SoftBank) qui construit des robots domestiques.
Les systèmes cognitifs ont la capacité d'apprendre et de donner des réponses à des questions non encore posées. Mais ils ne sont pas encore capables de reformuler une question. Aujourd'hui, le robot s'exprime en anglais. Adaptation en cours au Japonais et à l'Espagnol.

Technologies mises en œuvre pour :

1. Comprendre le langage naturel et la communication humaine
2. Générer et évaluer les hypothèses basées sur des références
3. S'adapter et apprendre

Watson utilise une base de connaissance dite « fermée », c'est à dire qu'il n'y a pas de contradictions à l'intérieur des documents. Les corpus sont vérifiés par un être humain avant d'être confiés au système.

IBM Watson fonctionne comme un expert. Par rapport à un simple moteur de recherche, le système expert cherche à comprendre la question, examine les réponses possibles, analyse les références puis fournit une seule réponse (ex : de quelle couleur est la voiture ?). Vis à vis du demandeur, le temps est raccourci entre la question et la réponse.

L'architecture de Watson est **massivement parallèle**. Elle utilise des algorithmes d'apprentissage et de raisonnement.
Contrairement à un humain, la machine a besoin d'évaluer de nombreuses références.

La plateforme est disponible sur le **Cloud** et est accessible via des **API**. Possibilité de l'installer dans un SI (ex : dans les hôpitaux pour des raisons de confidentialité des données).

Philippe travaille sur l'environnement de la retraite.
Démo disponible sur Chrome : http://devoxxthomasvox.eu-gb.mybluemix.net
Reconnaissance vocale disponible depuis Chrome.
Le résultat comporte la notion d'indice de confiance (sorte de ranking) et référence le document dans lequel les informations ont été retrouvées.

La solution repose sur une partie backend propulsée par Watson et un frontal Node.js déployé sur Bluemix.

La console Watson permet de gérer la base de connaissance, d'entraîner l'outil et de tester.

Base de connaissance
Le corpus est alimenté à partir de documents de format Word, PDF ou bien encore HTML. Plus le document est structuré, meilleur sera le taux de réussite. Philippe explique que, de part leur nature structurée, les documents HTML sont les meilleurs candidats.
L'analyse des documents se déroule en 16 étapes. A la fin, le corpus peut être exploité.

L'entrainement
L'utilisateur pose une première question. Il est alors possible de regrouper cette question à une autre de sens similaire. Si il n'en existe pas, à l'utilisateur de trouver une réponse dans les documents proposés par Watson. Si non trouvé, on peut saisir la réponse : déconseillé car trop spécifique.

Depuis http://hub.jazz.net on retrouve le code JS et le HTML généré.