



Opinions Libres

le blog d'Olivier Ezratty

Modèle d'étude de cas de l'IA

L'IA est un sujet relativement nouveau dans l'actualité des entreprises. C'est même "le" sujet le plus à la mode, après les vagues du cloud, de la mobilité et de la transformation digitale multiforme. Il génère son lot d'IA washing, un phénomène aux contours flous où de nombreux éditeurs de logiciels et startups habillent d'IA leurs solutions. Très souvent, cette appellation correspond à un usage d'une ou plusieurs techniques d'IA qui ne sont malheureusement pas toujours précisées. Les briques technologiques de ces solutions sont soit externes soit internes à la société. Il n'y a pas de mal à ce qu'elles soient externes car on ne reprochera pas à une startup de ne pas réinventer la roue.

Vu des clients, il est critique d'accéder à des études de cas de ces fournisseurs, histoire d'évaluer l'intérêt de lancer tel ou tel projet d'IA dans son entreprise. J'avais évoqué la question du "**benchmark de l'IA**" dans un post en mai 2017.

Nous avons grandement besoin de formalisme pour décrire les études de cas. Cela permet par exemple d'éviter les déclarations enflammées relayées par les médias et qui ne sont étayées par rien du tout. Comme l'exemple ci-dessous où l'on apprend qu'au Japon, une IA a été battue de justesse par des créatifs humains dans une agence de communication (**source**). Mais pas moyen de mettre la main sur les méthodes employées et sur les réponses des créatifs et de l'IA ! Ni, bien entendu, sur les techniques employées ! On ne voit que des clips vidéos produits par une IA et par des créatifs humains.

AI-CD β

robot utilisé comme directeur de création chez McCann Japan

... mais histoire très mal documentée !

Human Beats AI CD in McCann Japan's Creative Battle



By Erik Oster on Sep. 1, 2016 - 10:27 AM [Comment](#)

After introducing its A.I. CD early this spring, McCann Japan decided to pit the AI-CD β robotic creative director against its human counterpart, namely creative director **Mitsuru Kuramoto**, in a creative battle. Both were given the task of creating a spot for Mondelez Japan brand Clorets Mint Tab, communicating the brand message of "instant, long-lasting refreshment that lasts for 10 minutes" and then turning to a nationwide poll to declare the winner.



Voici donc une proposition de modèle de documentation d'étude de cas de projet intégrant de l'IA. C'est un modèle extensif qui sera probablement rarement complètement rempli. Peu d'entreprises ont envie de documenter leurs projets avec ce niveau de détails. Mais ces études de

cas peuvent être réalisées par certains éditeurs pour des projets présentés “behind closed doors”. J’ai notamment pu le constater dans une présentation de Justine Baron de **Recast.ai** (solution de création de chatbots) lors du séminaire Intelligence Artificielle organisé par Frenchweb lors du Cristal Festival de Courchevel les 14 et 15 décembre 2017. Lors de ce séminaire d’un jour et demi, je faisais un tour d’horizon assez large des techniques et usages de l’IA.

Qu’est-ce qui est spécifique à l’IA dans ce modèle ? C’est ce qui est en bleu *ci-dessous*.

Société cliente

- Secteur d’activité.
- Taille de l’entreprise. Bien préciser la taille de l’entité couverte par la solution. “Total” ou “Orange” n’est pas assez précis. On est souvent trompé par les études de cas qui ne précisent pas leur portée dans une très grande entreprise. Très souvent, les projets n’en concernent qu’une toute petite entité.
- Lieu, ce qui intéressant dans le cas de déploiements internationaux.

Solution

- Description métier du besoin et de la solution. **Comment faisait-on avant ? Quelles techniques classiques étaient utilisées ? Quels étaient les surcoûts engendrés par l’existant ?**
- Description technique de la solution. **Quelles techniques d’IA intègre-t-elle : de l’IA symbolique (système expert, moteur de règle, logique floue), du machine learning, des réseaux de neurones simples, du deep learning, des réseaux convolutionnels, des réseaux récurrents ou à mémoire, des techniques de traitement du langage.**
- Copies d’écrans de la solution, vue de l’utilisateur. L’interface utilisateur d’une solution logicielle est aussi importante que sa fonction !
- **Schémas fonctionnels, un diagramme des flux des données avec leurs sources étant indiqué.**

Données

- Nature, volume, origine et coût des données d’entraînement puis de production. Quels capteurs les ont générées (logs Internet, objets connectés, ...). **Quelles données sont d’origine interne et externe à l’entreprise ? Quelles données exploitées relèvent de l’open data. Quelles sont leurs conditions d’obtention commerciales ou en open data ?**
- Fréquence de la mise à jour opérationnelle des données. **Comment le modèle est-il réentraîné avec l’arrivée de nouvelles données ?**
- Taux d’erreur mesuré de la solution si applicable. **Ce taux est mesuré après l’entraînement du système d’IA si celui-ci utilise du machine learning ou du deep learning.**
- Anonymisation des données exploitées le cas échéant. **Est-ce que les données qui alimentent le machine learning ou le deep learning sont bien anonymisées. Normalement, c’est toujours**

le cas.

- Vidéo avec témoignage et démonstration, le cas échéant.

Fournisseurs

- Technologies. Au sens : logiciels de base (TensorFlow), d'infrastructure (Spark, Hadoop), logiciels divers et autres.
- Prestataires de services. En indiquant leur apport dans le projet.
- Ressources en cloud si pertinent. Et notamment, si des processeurs spécialisés (GPU ou neuromorphiques) sont utilisés, notamment pour l'entraînement d'un modèle de deep learning.

Dates

- Début du projet.
- Date des premiers tests opérationnels. Ce que l'on appelle un "PoC", pour proof of concept.
- Date de la mise en production. Et portée de la mise en production en nombre d'utilisateurs.

Economie

- Coût du projet. Ressources humaines consommées en interne et en externe pour créer la solution. Types de compétences : développeurs, data-scientists, etc.
- Durée d'entraînement des modèles, dans le cas de solutions à base de deep learning.
- Nombre d'utilisateurs de la solution. Aujourd'hui et demain.
- Retour sur investissement. C'est la partie la plus difficile à mesurer sur de nombreux projets. Il faut pouvoir y intégrer l'ensemble des coûts relatifs au projet, y compris la formation des utilisateurs.
- Validation du projet au regard de la RGPD, la Règlementation Générale de la Protection des Données européenne qui entre en vigueur le 25 mai 2018.

J'ai aussi proposé le modèle à l'équipe du Hub#FranceAI dans le cadre de ses activités de formation des entreprises.

Est-ce complet ? Que manque-t-il à un tel modèle ?

Un volontaire pour créer une première étude de cas entièrement documentée ? Je la publierai volontiers dans ces colonnes !

Cet article a été publié le 18 décembre 2017 et édité en PDF le 5 septembre 2020.
(cc) Olivier Ezratty - "Opinions Libres" - <https://www.oezratty.net>