

La révolution de l'assurance par la donnée : défis scientifiques de l'extraction à la gestion de connaissances

Marcin Detyniecki* ** ***

*GIE AXA, Data Innovation Lab, Suresnes, France
Marcin.DETYNIECKI@axa.com

**Polish Academy of Sciences, IBS, Warsaw, Poland

***Sorbonne Universités, CNRS, UPMC Univ Paris 06, UMR 7606,
LIP6, Paris, France

Biography

Marcin Detyniecki is senior R&D officer at the Data Innovation Lab recently created by AXA. Additionally, he is professor at the Polish Academy of Science (IBS PAN) and associate researcher at the computer science laboratory LIP6 of the University Pierre and Marie Curie (UPMC). Currently, his research focuses on the emerging challenge popularly named big data. In the past, he has worked on the usage of new media, with challenges ranging from multimedia information retrieval to image understanding. Several of the developed applications have not only been deployed in the market, but they have also been singled out in international competitions such as TrecVid, ImageClef, MediaEval. This applicative success is the results of a dialogue with more theoretical works on topics such as new challenges in approximate reasoning, information aggregation and fusion, and machine learning from a computational intelligence perspective. Marcin Detyniecki studied mathematics, physics and computer science at the University Pierre and Marie Curie (UPMC) in Paris. In 2000 he obtained his Ph.D. in Artificial Intelligence from the same university. Between 2001 and 2014, he was a research scientist of the French National Center for Scientific Research (CNRS). He has been researcher at the University of California at Berkeley and at Carnegie Mellon University (CMU). He has been visiting researcher at the University of Florence and at British Telecom Research labs. Today he is member of the research and academic council of UPMC University, member of the executive board of laboratory SMART, elected member of the LIP6 laboratory council, and member of the editorial board of the International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (IJFUKS). He also funded and animated until 2014 the UPMC – Sorbonne Universités Computer Science Colloquium. Dr. Detyniecki has over 90 publications in journals and conference proceedings, including 6 keynotes.

Résumé

La quantité de données dans notre monde a explosé et l'analyse de grands ensembles de données – aussi connu dans l'industrie sous le nom « *Big Data* » – deviendra un atout majeur de compétitivité, principalement dû à une croissance de productivité et surtout à grâce à plus d'innovation. La croissance exponentielle de données est alimentée par la facilité de la captation et par la multiplication de canaux numériques d'acquisition. On pense non seulement à tous les processus qui sont informatisés aujourd'hui, mais aussi aux médias sociaux et aux objets connectés.

L'assurance vie une révolution tout particulière. L'assureur, traditionnellement gestionnaire du risque en s'appuyant sur une longue expérience, qu'on traduirait aujourd'hui par une captation systématique de données, est après la révolution numérique partiellement exclus de canaux digitaux.

Ceci est en même temps une menace et une opportunité. Il s'agit d'un défi puisque l'industrie doit réaliser une forte mutation pour se positionner là où la donnée se trouve aujourd'hui, *i.e.* dans le digital. Il s'agit d'une opportunité puisque ces nouvelles données permettront de mieux appréhender les risques, et plus particulièrement, permettront d'estimer au plus près les risques à la source, plutôt que passer par de variables intermédiaires, comme peut l'être l'âge pour le risque d'accident en conduite. L'opportunité est d'autant plus grande qu'en accédant aux données au plus près des utilisateurs il est possible de faire de la prévention évitant ainsi des accidents coûteux pour l'assureur, mais surtout désastreux pour les victimes.

Une fois la révolution engagée, ceci implique, un certain nombre de transformations dans les processus d'extraction et gestion de connaissances. Les défis scientifiques sont nombreux, allant de la captation non-intrusive de la donnée, à la visualisation et gestion de connaissances extraites, en passant par de l'apprentissage artificiel pour pouvoir servir à de millions d'utilisateurs simultanément. Dans cette présentation nous allons couvrir rapidement chacune de ces thématiques avec une attention particulière aux défis scientifiques sous-jacents.

Nous allons illustrer notre propos par un exemple phare de cette révolution : la famille d'offres d'assurance dite « *pay as you drive* » où généralement on obtient une décote ou réduction en fonction de sa façon de conduire. Nous allons ce que ceci implique en termes d'extraction et de gestion de connaissances.

Pour conclure, il est important de mentionner que cette révolution implique d'autres challenges cruciaux qui dépassent ce qui est abordé ici. En particulier, pour ne mentionner que deux grands axes : la protection de la vie privée, aussi bien du point de vue technique que juridique ; et la transformation de métiers accompagné d'une pénurie de talents déjà entamé.