



**HAL**  
open science

# Du concept à la mise en œuvre du machine learning dans les entreprises : L'expérience de Datapred

Henri Bourdeau, Corentin Petit, Christophe Midler

## ► To cite this version:

Henri Bourdeau, Corentin Petit, Christophe Midler. Du concept à la mise en œuvre du machine learning dans les entreprises : L'expérience de Datapred. Les cahiers du Master PIC, 3, 2019, Les POC.. et après ?. <hal-02873935>

**HAL Id: hal-02873935**

**<https://hal.science/hal-02873935v1>**

Submitted on 18 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Les cahiers du  
Master PIC

## DOSSIER 1 : LES POC... ET APRÈS?

Cahier du master PIC n° 3

# Du concept à la mise en œuvre du machine learning dans les entreprises : L'expérience de Datapred

Henri Bourdeau  
DataPred

Corentin Petit  
DataPred

Christophe Midler  
Centre de Recherche en Gestion I<sup>3</sup>

---



Copyright © 2019 Observatoire Projet Innovation  
Conception – École polytechnique

L'Observatoire Projet Innovation Conception encourage, sous réserve de citation (référence bibliographique et/ou URL correspondante), la reproduction et la communication de ses contenus, à des fins personnelles, dans le cadre de recherches ou à des fins pédagogiques. Toute utilisation commerciale (en version imprimée ou électronique) est toutefois interdite.

Les opinions, interprétations et conclusions exprimées sont celles de leurs auteurs, et n'engagent pas l'Observatoire Projet Innovation Conception et l'École polytechnique en tant qu'institution ni les individus ou les organisations consultés dans le cadre de cette étude.

Citation: Bourdeau H., Petit, C. et Midler C. (2019). « Du concept à la mise en œuvre du machine learning dans les entreprises: l'expérience de Datapred » Les cahiers du Master P.I.C. N° 3, L'Observatoire Projet Innovation Conception – École polytechnique, Paris, France, 18 P.

## **AVANT-PROPOS**

Ce cahier exploite les résultats du projet de Master PIC mené pendant 18 mois par Henri Bourdeau avec l'entreprise Datapred. Il est le fruit d'un travail collectif impliquant l'étudiant de Master, l'enseignant qui l'a encadré dans la recherche et le responsable qui l'a suivi dans l'entreprise.

Nous tenons à remercier tous les acteurs qui ont collaboré à cette recherche, sans lesquelles le master PIC, dans son identité d'articuler théorie et pratique, ne pourrait pas exister.

## RESUME

L'Intelligence Artificielle est l'objet d'un engouement exceptionnel dans l'industrie, mais, dans la pratique, peu de données concrètes permettent d'attester des conditions et des résultats de sa mise en œuvre. Le projet de Master PIC a permis d'expérimenter le déploiement concret d'une offre de machine learning dans une entreprise leader du domaine du traitement des données temporelles, Datapred.

Les enseignements tirés sont triples.

D'abord, l'importance de la phase de Proof Of Concept dans la mise en œuvre de ces nouvelles technologies: étape clé de crédibilité de l'efficacité concrète de ces technologies pour les clients, étape majeure d'apprentissage des problématiques métiers abordées pour les experts de l'analyse de données.

Ensuite, l'ampleur de la distance qui sépare la promesse d'un POC réussi de la vente d'un logiciel d'IA finalisé. Du fait des processus décisionnels complexes chez les grandes entreprises clientes comme de la coordination du travail des ingénieurs POC et des développeurs produit dans l'entreprise de conception logicielle.

Enfin, la nécessité, pour franchir cette distance, d'opérer des mutations de la stratégie et de l'organisation de l'entreprise d'IA, mutation portant sur sa conception logiciel, l'organisation de sa R&D et son business model.

**Mots clés:** Intelligence artificielle, apprentissage automatique, POC, conception logiciel, stratégie de croissance.

## SUMMARY

Artificial Intelligence is the subject of exceptional enthusiasm in the industry, but in practice, there is little concrete data to attest to the conditions and results of its implementation. The PIC Master's project made it possible to experiment with the concrete deployment of a machine learning offer in a leading company in the field of time data processing, Datapred.

The lessons learned are threefold.

First, the importance of the Proof Of Concept phase in the implementation of these new technologies: a key step in ensuring the credibility of the concrete effectiveness of these technologies for customers, a major step in learning the business issues addressed for data analysis experts.

Then, the magnitude of the distance between the promise of a successful POC and the sale of finalized AI software. Due to the complex decision-making processes at large corporate customers as well as the coordination of the work of POC engineers and product developers in the software design company.

Finally, the need, in order to cross this distance, to make changes in the strategy and organization of the AI company, changes relating to its software design, the organization of its R&D and its business model.

**Keywords:** Machine Learning, Proof Of Concept, Digital transition, Growth Strategy



## SOMMAIRE

<b>1. INTRODUCTION</b>	<b>7</b>
<b>2. LE MACHINE LEARNING : TECHNOLOGIES ET MARCHÉS</b>	<b>9</b>
2.1. Le machine learning	9
2.2. Le marché du machine learning	10
2.3. Le cas Datapred	10
<b>3. DU POC À L'OFFRE FINALISÉE : UNE TRAJECTOIRE AU CONCRET</b>	<b>13</b>
3.1. Phase 1 de prototypage: Valider la pertinence opérationnelle. Objectif rapidité et performance	13
3.2. L'après POC (1) le développement logiciel. Objectif: robustesse	15
3.3. L'après POC (2): Une relation à l'entreprise cliente à reconstruire.	16
<b>4. RÉCONCILIER LE POC ET LE DEVELOPPEMENT PRODUIT, L'EXPÉRIENCE DE DATAPRED</b>	<b>17</b>
<b>CONCLUSION</b>	<b>19</b>
<b>REFERENCES</b>	<b>19</b>



## 1. INTRODUCTION

Le déploiement du machine learning (ou apprentissage automatique) dans l'industrie témoigne d'un paysage contrasté. D'un côté, ces nouvelles technologies sont l'objet d'un engouement exceptionnel, à compter le nombre de discours, colloques, séminaires et formations auxquels il donne lieu. D'un autre, dans la pratique, les industriels ont souvent des difficultés à l'exploiter. Plus précisément, ils arrivent à construire des prototypes, mais il faut généralement un à deux ans pour les convertir en solution. Une fois cette conversion faite, ils constatent alors que la solution ne fonctionne pas à la hauteur de leurs attentes.

On retrouve ici le constat général de la dynamique des attentes envers l'innovation technologique telle que la courbe de Gartner l'a décrite (Linden, A., & Fenn, J., 2003) ou les désillusions qui suivent généralement l'apogée des modes managériales (Midler 1986). Mais au-delà du constat général des effets sociaux amplificateurs sur l'adoption des innovations managériales, quelles sont aujourd'hui les difficultés spécifiques auxquelles se confronte l'adoption de ces nouvelles démarches? Quelles sont les dynamiques en cours pour les dépasser? Ce sont ces deux questions auxquelles ce cahier cherche à contribuer, par l'analyse de l'expérience d'une entreprise experte du domaine, Datapred.

Datapred est une start-up de conception de logiciel d'intelligence artificielle fondée en novembre 2014 spécialisée dans le traitement de données temporelles. Sa proposition de valeur ajoutée repose sur une nouvelle technologie, "l'agrégation séquentielle" et un savoir-faire sur le traitement des séries temporelles par le machine learning (branche de l'IA qui s'appuie sur des outils spécifiques). Celle-ci lui permet de combiner les résultats de plusieurs algorithmes pour obtenir des performances meilleures que chacun de ces algorithmes pris séparément.

Sa démarche commerciale comporte deux étapes. La première consiste en une phase de test (« proof of concept ») où une équipe de l'entreprise utilise ses méthodes sur un sujet réel spécifique posé par son client. Cette étape implique à la fois d'apprendre précisément la question de l'entreprise cliente, et d'ajuster sur mesure les outils de prévision développés antérieurement par Datapred pour trouver une réponse la plus adaptée possible. Si ce test est positif, c'est-à-dire si la prévision donnée par les outils Datapred est meilleure que les prévisions réalisées antérieurement par les spécialistes de l'entreprise, DataPred propose de vendre des licences de son logiciel d'apprentissage automatique pour que l'entreprise l'intègre comme outil de prévision systématique mis en œuvre par ses prévisionnistes internes. L'hypothèse de croissance de l'entreprise repose sur la cumulativité des apprentissages réalisés par les missions suc-

cessives: chaque test chez un client permet d'améliorer les outils génériques et de les rendre plus robustes à des problématiques d'utilisation diversifiées. L'efficacité des missions d'exploration initiales devrait donc être croissante, augmentant tendanciellement la rentabilité du modèle.

Le projet de Master, réalisé d'avril 2017 à août 2018 s'est opéré en deux phases. La première a permis d'expérimenter et d'analyser cette démarche. On verra que la question du passage du POC, même réussi, à la production, la vente et l'implémentation d'un logiciel « clé en main » chez le client a constitué la difficulté majeure de la démarche. La seconde phase a permis de participer à la définition et la mise en œuvre de transformations nécessaires pour franchir cet obstacle du POC à la vente d'un logiciel finalisé.

Ce cahier PIC est organisé en trois parties. La première présentera rapidement le champ du machine learning dans l'activité de l'entreprise étudiée, Datapred. La seconde présentera le cycle de mise en œuvre de la démarche Datapred telle qu'expérimentée en 2018 et l'analyse a posteriori qui en a été faite. La troisième partie tirera les enseignements du cas quant aux évolutions à mettre en œuvre face à ce diagnostic, évolutions portant à la fois sur l'organisation et les processus de développement des logiciels, la conception de ces logiciels elle-même et le business model de l'entreprise.

Les principaux enseignements de cette recherche, que l'on découvrira dans ce cahier, sont triples.

Tout d'abord, l'importance de la phase de POC dans la mise en œuvre de ces nouvelles technologies. C'est à la fois une étape clé de crédibilité de ces outils nouveaux aux yeux des clients. C'est en même temps une étape majeure d'apprentissage pour les experts de l'analyse de données qui confrontent leurs techniques aux spécificités d'un questionnement professionnel qu'ils découvrent. C'est enfin une source de financement de l'entreprise, même si l'essentiel du revenu est attendu de la vente des licences logicielles.

Ensuite, l'importance du saut qui sépare la promesse d'un POC réussi de la vente d'un logiciel finalisé. Un saut qui tient à la fois à la structuration de la problématique de mise en œuvre des prévisions chez les clients et à l'articulation des missions de test et de développement produit dans l'entreprise de conception logiciel.

Enfin, la nécessité, pour franchir ce saut, d'opérer des mutations de la stratégie et de l'organisation de l'entreprise de data analytics, mutation portant sur sa conception logiciel, l'organisation de sa R&D et son business model.



## 2. 2. LE MACHINE LEARNING : TECHNOLOGIES ET MARCHÉS

### 2.1. LE MACHINE LEARNING

L'apprentissage automatique ou machine learning est une branche de l'intelligence artificielle qui concerne la mise au point de méthodes permettant à des machines d'adapter elles-mêmes leur façon de fonctionner pour répondre à un besoin, et ce en observant des données.

En pratique le besoin prend la forme de l'un de ces deux types de problèmes :

- Les problèmes supervisés: on veut être capable de prévoir la valeur d'une variable à partir des observations passées. Ainsi la poste dispose de machines qui sont capables de lire elles-mêmes les codes postaux manuscrits sur les enveloppes. Les variables sont ici les chiffres. Carrefour cherchera, par exemple, à associer le chiffre d'affaire de ses magasins aux différents critères de satisfaction client pour savoir sur lesquels investir. La variable prédite ici est le chiffre d'affaires.
- Les problèmes non supervisés: on veut être capable de regrouper les observations sans connaître a priori quels sont les groupes. En appliquant une telle méthode à ses films, Netflix les groupe par similarités. Elle est ainsi en mesure de suggérer à ses clients des films qui leur plairont probablement: ce sont ceux qui font partie des mêmes groupes.

La différence entre ces deux problèmes est la présence ou non d'une variable cible parmi les données initiales: pour Carrefour on connaît le chiffre d'affaire des magasins pour les mois précédents. C'est la variable que l'on souhaite prédire. Pour Netflix, les groupes n'existent pas a priori.

La démarche est ensuite la suivante: on fournit à un algorithme les données sous la forme d'une série d'observations. L'algorithme cherche ensuite à comprendre la structure sous-jacente de ces données. C'est à dire qu'il essaye de déterminer des caractéristiques qui lui permettent de différencier les observations et de répondre au problème en les comparant les unes aux autres.

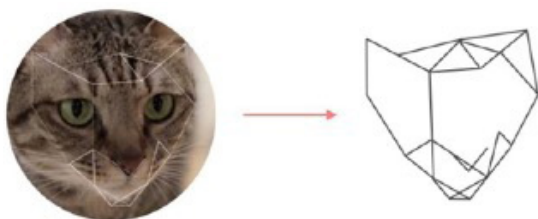


Figure 1 : Les algorithmes identifient la structure sous-jacente des données

Historiquement, pour faire du machine learning, on utilise une méthode dite "batch". C'est à dire que l'utilisation des algorithmes se fait en deux étapes: l'entraînement puis la prédiction.

L'entraînement est la phase pendant laquelle l'algorithme cherche à déterminer cette fameuse structure sous-jacente. Poursuivons sur l'exemple classique de la reconnaissance d'un chat. Pendant la phase d'entraînement, elle va donner à l'algorithme des milliers d'images en indexant celle qui représentent un chat. Le deuxième jeu d'images permettra à l'algorithme de faire la différence entre les images de chat et celle qui n'en représentent pas. L'algorithme travaille ensuite avec ces milliers d'images. Une fois qu'il a terminé, on peut le relier à la caméra et le laisser décider de la nature des animaux photographiés.

Cette approche batch est extrêmement efficace pour ce genre de situations où la structure des données ne change pas au cours du temps: un chat reste un chat. En revanche, lorsque celle-ci évolue au cours du temps, elle est très limitante: il faut ré-entraîner l'algorithme sur de nouveaux exemples contenant les valeurs nouvelles des variables que l'on veut prédire. Et il faut ensuite recommencer pour chaque nouvelle valeur de la variable évolutive. Avec la popularisation du "Big Data" et des objets connectés, les entreprises amassent des quantités impressionnantes de données qui nécessitent des temps de traitement considérables. La solution proposée par les grandes plateformes de machine learning telles que l'Amazon cloud service ou IBM analytics est le "batch haute fréquence". Les algorithmes sont ré-entraînés suffisamment fréquemment pour compenser les changements de structure. La location de serveurs de calcul capables de digérer le volume mensuel de données en un temps suffisamment restreint pour être utilisable représente un coût significatif. C'est pour répondre à ce problème qu'une nouvelle approche a été créée, le machine learning séquentiel.

En machine learning séquentiel les algorithmes évoluent pas à pas. Au lieu de séparer une phase d'apprentissage et de prédiction, ces deux étapes sont faites conjointement. On peut à tout moment demander à l'algorithme de faire l'un comme l'autre: soit on lui donne une nouvelle observation d'entraînement et il met à jour son analyse de la structure de donnée, soit on lui demande de prédire et il le fait en se basant sur ses connaissances actuelles. En cas d'erreur on peut lui donner ensuite la bonne réponse à titre de donnée d'entraînement.

Cette nouvelle approche permet d'obtenir des algorithmes qui se mettent à jour effectivement en temps réel tout en remplissant leur rôle de prédiction. Cet avantage des algorithmes séquentiels vis-à-vis de leur équivalent batch apparaît particulièrement net lorsqu'ils s'appliquent à des grandeurs temporelles, qui évoluent au cours du temps comme par exemple un nombre de ventes par jour. Datapred est à ce jour leader dans l'application du machine learning séquentiel à des cas industriels.

## 2.2. Le marché du machine learning

Nous n'insistons pas ici sur l'ampleur et la variété des clients intéressés par les applications potentielles des technologies du machine learning. Comme précisé en introduction, l'explosion du volume des données de toutes natures liées à l'internet comme aujourd'hui à l'internet des objets déclenche depuis quelques années des espérances nouvelles dans la possibilité de transformer ces données pléthoriques en connaissances utilisables dans des domaines variés : du marketing à la logistique, la finance, la maintenance ou la gestion de production.

Sur le plan de l'offre, trois types d'acteurs investissent ce nouvel eldorado.

Les sociétés Tech (grandes ou petites) spécialisées en informatique et analyse de données. C'est le cas de Datapred bien sûr, mais aussi de Google, IBM, AutoML, Datarobot...

Les cabinets de conseil qui ont créé des branches dédiées à l'apprentissage automatique comme McKinsey, Bain, le BCG par exemple. Ces entreprises proposent comme Datapred une étude approfondie du contexte opérationnel de leur client, la création d'une solution sur-mesure et l'intégration de cette solution à un niveau d'exigence industrielle dans la structure informatique du client.

Enfin, les équipes internes de développement de solutions d'apprentissage automatique des grandes entreprises clientes, notamment tous les départements de numérisation, "*digitalisation*", etc. cherchant à monter en compétences sur ce domaine jugé stratégique (par exemple G&E et Thales).

Il faut noter que les relations entre ces trois types d'acteurs évoluent vers des situations qui ne prennent pas la forme d'une concurrence frontale. On retrouve sur la technologie du machine learning la complémentarité entre sociétés de service en organisation, équipes internes spécialisées et entreprise pure player.

## 2.3. Le cas Datapred

Dans ce contexte du marché de l'IA, Datapred est une startup financée en fonds propres fondée en octobre 2014 qui repose sur une méthode unique pour traiter les données séquentiellement et une technologie propriétaires: KPP, une variante d'un algorithme très populaire et polyvalent. Elle est accélérée par les programmes d'accélération de la SNCF et AIRBUS et compte désormais 11 membres. Lauréate de la fondation RAISE, Datapred mène actuellement des projets auprès de 16 clients parmi les grandes entreprises françaises Autode secteurs variés, de l'automobile au retail en passant par les industries agroalimentaires.

Sa proposition de valeur repose sur trois piliers :

- Une bibliothèque informatique générique conçue par l'équipe de R&D interne de Datapred, fruit de l'expérience accumulée sur des projets menés pour de grands industriels. Par rapport aux autres compétiteurs, Datapred se différencie notamment par la prise en compte de l'aspect temporel des données et sa capacité à intégrer d'éventuelles prédictions pré-existantes chez le client. Ces deux facteurs lui ont permis jusqu'ici d'obtenir des performances sensiblement meilleures sur les sujets clients pour lesquels il y a eu concurrence directe (dans des logiques d'appels d'offres).
- Une offre de service data science pour construire une solution à l'aide de la bibliothèque, répondant à la problématique spécifique de chaque client. Cette compétence s'exprime lors de la phase initiale de la relation avec les clients, qui donne lieu à l'élaboration d'un Proof of Concept (POC) visant à tester la pertinence de la technologie de Datapred à fournir une réponse meilleure que les prévisions existantes avant son intervention. Cette phase permet d'étudier conjointement avec le client son problème et la meilleure façon pour Datapred d'y répondre. Les équipes de Datapred mettent en place cette solution et la testent sur les données historiques du client.
- Si les résultats du POC sont satisfaisants pour le client, Datapred effectue un passage en production en installant la solution sur un serveur chez le client. Le logiciel Datapred est conçu pour s'intégrer dans le système du client de façon transparente: les échanges d'informations sont faits aux formats des outils déjà en place et les opérations de prédiction sont faites de façon entièrement automatisées. Le logiciel est vendu sous la forme d'une licence annuelle. Cette licence inclut la mise à jour du logiciel au fur et à mesure de son perfectionnement par les équipes de recherche et développement de Datapred et les éventuelles modifications que le client souhaiterait apporter à son mode de fonctionnement.

La capacité de croissance de la proposition de valeur de Datapred repose à la fois sur la performance de sa technologie de prévision, mais aussi la généralité de la technologie (qui va la rendre applicable à des classes de traitement de données variées) et la facilité d'adaptation à des contextes et problématiques variées (pour limiter la réécriture du code pour chaque cas d'usage). L'effort de l'entreprise vise à développer cette généralité et cette adaptabilité de la bibliothèque informatique, de manière à augmenter, de projet en projet, le rendement des phases de POC et de production des logiciels dédiés clients. C'est précisément dans le cadre de cette stratégie que va s'inscrire le projet de Master PIC.

Au moment où il s'engage, Datapred était constituée de 6 membres. Ils portaient chacun bien sûr plusieurs casquettes, mais la répartition globale des ressources était approximativement la suivante :

- Deux membres prennent en charge la relation avec les clients, la conduite des projets et la construction des prototypes

- Deux membres cristallisent l'expertise mathématique en soutien sur les projets et effectuant de la recherche sur de nouveaux modèles
- Deux membres concentrent la compétence de développement informatique et effectuent les mises en production en apportant un soutien technique à la partie projet.

Début 2017, deux nouveaux développeurs sont recrutés pour accroître ses capacités et le projet de Master PIC est lancé pour accompagner la transition depuis la logique sur-mesure vers un logiciel générique clé en main pour des contextes d'utilisation variés. Pour parvenir à cet objectif, le projet PIC va s'organiser en trois phases. D'abord, participation à une mission de réalisation d'un POC chez un client réel, en l'occurrence la prédiction de demande de voitures pour Autoun constructeur automobile, que nous appellerons Auto. Ensuite, analyse de l'étape de passage du POC à la vente du logiciel finalisé. Enfin, proposition et mise en œuvre d'évolutions d'organisation et de processus pour améliorer ce passage dans le cadre du développement de l'entreprise.



Figure 2: Présentation de l'évolution de Datapred



### 3. DU POC À L'OFFRE FINALISÉE : UNE TRAJECTOIRE AU CONCRET

La conduite d'un projet de construction d'une solution d'apprentissage automatique dans l'industrie se fait classiquement par une étude de faisabilité puis par un passage en production. Ces deux phases nécessitent des compétences et des modes de travail très différents l'un de l'autre.

#### 3.1. Phase 1 de prototypage : Valider la pertinence opérationnelle. Objectif rapidité et performance

L'interaction avec l'entreprise cliente démarre par une phase initiale de workshops (non facturés) avec le client pour trouver la bonne problématique et définir les objectifs du projet prototypé. La question de l'interlocuteur client est à la fois importante et complexe dans les grandes entreprises comme Auto. Ici, le sujet choisi a été l'amélioration des prévisions de ventes de véhicules à plusieurs mois, de manière à optimiser la chaîne de production. Le projet a été piloté par la division Supply Chain, mais on comprend facilement qu'il concerne aussi fortement le département commerce, qui, jusque-là est l'unité qui fait les prévisions de ventes. Les enjeux de ces deux unités ne sont pas les mêmes : la supply chain a besoin des prévisions les plus fiables possibles pour éviter la surproduction et les surstocks couteux. Le commerce, d'un autre côté veut avoir toujours du stock pour vendre. Cette phase d'initialisation a duré ici deux mois.

Dans cette phase de sélection des prestataires sur un sujet majeur comme celui-là, Auto lance plusieurs POC en parallèle avec d'autres start up, prestataires grand cabinet de conseil et en interne dans son département Digital en cours de montée en puissance.

Une fois Datapred sélectionné par le client Supply Chain et le cas d'application identifié, s'engage une phase de Proof of Concept de 3 mois qui sera menée avec conjointement la Supply chain et le commerce. Une équipe spécialisée dans la construction de modèles, l'équipe POC par la suite, identifie quels vont être les entrées et sorties du modèle. Ce qu'on lui demande ici, c'est d'avoir une très bonne compréhension de la situation opérationnelle de l'entreprise. Il faut que la sortie qu'elle va décider de choisir pour son modèle ait un sens interprétable et utile pour les ou opérationnels qui vont interagir ensuite avec la solution.

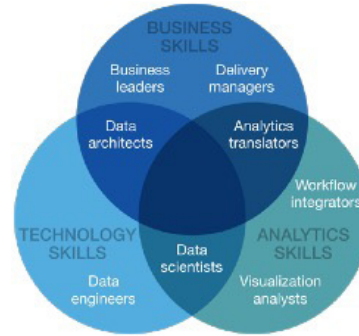


Figure n° 3 : Connaissance opérationnelle de l'entreprise, compréhension des structures de données et capacités de modélisation :  
Les trois compétences nécessaires pour faire de l'apprentissage automatique  
Source : McKinsey Global Institute (2016)

Cette étape correspond à analyser le fonctionnement de toute la chaîne logistique du constructeur pour identifier le point de planification future. Ce point existait évidemment déjà chez le client : il disposait d'une équipe d'experts qui effectuaient des prédictions pour un nombre réduit de modèles. Un algorithme s'occupait ensuite d'affiner ces prédictions pour prendre en compte l'intégralité de l'offre du constructeur. Le sujet était de venir intégrer les prédictions, soit au niveau simplifié, soit au niveau affiné, et de les améliorer pour les réinjecter ensuite dans le système existant.

Sans une bonne connaissance du fonctionnement de la logistique du client, il serait tout simplement impossible de connaître l'existence même de ces prédictions pré-existantes. Pour la prédiction de volume de véhicules, cette tâche-là est plus compliquée qu'il n'y paraît. Un véhicule est, d'un point de vue fonctionnel, la combinaison d'une centaine d'options : du type de moteur jusqu'à la présence de toit ouvrant ou de GPS intégré. Cela représente des milliers de configurations de véhicules différents qui doivent toutes être prises en compte : La présence ou non d'un toit ouvrant a des conséquences sur les pièces utilisées pour le cadre du véhicule. On ne peut pas les traiter indépendamment. Mais au-delà même de cette combinatoire, il existait 9 colonnes différentes dans la base de données correspondant à des niveaux d'avancement différents des véhicules dans le processus de fabrication et de vente. Il aura fallu deux semaines de discussion avec les experts du client pour déterminer finalement celle qui était la plus pertinente.

Une fois que la sortie est déterminée, l'équipe POC va étudier le contenu des structures de données du client pour identifier quelles données seront utiles pour le bon fonctionnement de la solution d'apprentissage automatique.

Auto utilise des prédictions humaines dans son système actuel. Ces prédictions reposent sur des méthodes et de l'expertise interne et sont extrêmement efficaces pour déterminer les tendances de marché mais il est humainement impossible de traiter les milliers de configurations différentes. Elles ne sont donc uniquement présentes qu'au niveau des objets clés.

De plus, il y a déjà eu des tentatives par le passé d'introduire des algorithmes de machine learning chez Auto. Ceux-ci n'étaient pas performants et parce qu'ils avaient pour vocation de remplacer les équipes d'Auto, ils ont résulté en un faible taux d'adoption à la démarche.

Enfin, le but de la démarche est de prédire la demande libre du marché. En temps normal, en prédiction, on observe l'évolution de la valeur à prédire au cours du temps ainsi que celle d'indicateurs qui pourraient y être liées. Ici, on n'a pas accès à la valeur de la demande libre. C'est par essence quelque chose qui n'est pas mesurable.

Nous disposons cependant d'une estimation, la quantité de véhicules vendus par le passé. Elle est cependant biaisée. Les véhicules vendus dépendent de ceux construits qui sont eux-mêmes liés aux prédictions faites par les équipes d'Auto. Lorsque le marché ne se comporte pas comme prévu et qu'Auto a trop produit d'un type de véhicule, elle peut quand même le pousser sur le marché en le subventionnant. Ainsi dans le second semestre 2016, les prévisions avaient surestimé le volume de diesel qui serait demandé. Le surplus de production a cependant été vendu en "forçant" le marché. Cette capacité d'Auto à forcer le marché confère aux prédictions qu'elle utilise une capacité autoréalisatrice. La crainte chez Auto est donc que des algorithmes se basant uniquement sur le passé ne reproduisent ce genre d'erreur en continuant à prédire une demande en diesel dans notre exemple. On appellera cela biais de modélisation. Datapred a développé de nouveaux procédés pour répondre à ce biais en mobilisant dans le processus de calcul les expertises métiers.

Enfin, l'équipe POC peut travailler sur la création du modèle d'apprentissage automatique à proprement parler : choisir un modèle, ses paramètres, un optimiseur et un critère de performance. On retrouve ici la mobilisation de compétences en modélisation.

C'est ici que l'équipe POC tire parti de la technologie d'agrégation séquentielle propre à l'entreprise : on lui donne un ensemble d'algorithmes de prédiction.

Dans le cas d'Auto, cela inclut les prédictions de leurs experts en objets clés, la conversion en combinaisons d'objets faite par leur algorithme et les prédictions de différents algorithmes ajoutés par Datapred se basant sur l'historique des ventes, des fabrications et des stocks. Chacun de ces algorithmes est spécialisé dans un comportement de données : certains sont bons sur des comportements saisonniers, d'autres pour anticiper des fluctuations à court terme, d'autres à long terme. L'agrégation séquentielle détermine ensuite en temps réel la combinaison de ces algorithmes qui représente au mieux la réalité de la donnée. Cette approche permet d'obtenir facilement un algorithme qui soit toujours adapté à la donnée qu'il traite même si celle-ci peut présenter des structures très différentes : un produit commercial passe typiquement par des phases de lancement, de vie mature et de fin de vie. Ce sont des concepts que l'on comprend humainement, mais réaliser un unique algorithme qui puisse traiter toutes ces configurations simultanément est très délicat. L'agrégation simule un tel algorithme en "mélangeant" des modèles bons dans les différents cas en adaptant les proportions au cours du temps.

Une fois ce travail terminé, l'équipe entraîne le modèle et mesure ses performances. Le processus de validation de la solution se fait en deux parties : a) Les performances sont-elles satisfaisantes ? et b) Le résultat de l'algorithme est-il exploitable opérationnellement parlant ? Si ces deux critères sont satisfaisants, l'étude de faisabilité de la solution prend fin et on passe à la production.

Sur l'exemple de prédiction de demande du marché pour les véhicules, ce n'est qu'à ce moment que l'on commence à entraîner et sélectionner des modèles, à retraiter les entrées manuellement pour obtenir de meilleurs résultats.

En pratique, l'étude de faisabilité ne se résume pas à la création et la validation d'un seul modèle. Il faut souvent itérer à plusieurs reprises sur l'ensemble du procédé - analyse opérationnelle, analyse des données, analyse des modèles - pour affiner la valeur de sortie du modèle, afin qu'elle soit au mieux exploitable par les opérations. À chaque fois qu'une valeur de sortie est définie, il faut souvent également plusieurs tentatives avant d'obtenir un modèle avec une performance suffisante.

Dans le cas de la prévision du programme de vente d'Auto, ce travail a finalement produit une réduction de l'erreur de prévision de 30 % par rapport aux processus de prévision existants précédemment. C'est à dire que la prévision Datapred réduit de 30 % le nombre de véhicules mal anticipés et qui doivent être ensuite discountés. Cela représente 10 % du nombre total de véhicules produits dans la zone d'études. L'économie potentiellement réalisée est de l'ordre de la centaine de millions d'euros.

### 3.2. L'après POC (1) Le développement logiciel. Objectif : robustesse

À l'issue de cette phase POC, le département Supply Chain valide la solution. S'engage alors la phase de transition du POC à la vente et la livraison du logiciel intégré aux systèmes du client. Une phase qui comporte un versant de développement logiciel chez DataPred et un processus de négociation et décision d'implémentation chez le client Auto. Une phase difficile pour Datapred car elle va mobiliser des ressources significatives sans pour autant pouvoir donner lieu à facturation par les clients.

Une fois l'intérêt opérationnel du POC validé, il s'agit d'en faire un produit logiciel qui puisse s'intégrer à l'opérationnel de l'entreprise. Celui-ci va fortement interagir avec la structure de données existantes puisqu'il va devoir y puiser ses variables d'entrées, celles de sorties, s'entraîner, effectuer des prédictions sur des nouvelles variables d'entrées et enregistrer le résultat dans la structure de données.

Les solutions d'apprentissage automatique sont très sensibles à la qualité des données qui vont y entrer. Un défaut même temporaire dans la logique de la structure des données peut avoir des conséquences désastreuses à la fois à court terme sur la performance de l'entreprise et à long terme sur sa capacité à tirer profit des possibilités liées à l'apprentissage automatique. Or, il existe souvent des multitudes de systèmes qui interagissent avec la base de données et qui sont source de défauts potentiels : depuis des capteurs qui créent initialement cette donnée, viennent ensuite les canaux d'acheminement qui peuvent la corrompre. Ensuite interviennent dans le processus de traitement des données l'ensemble des applications développées ou intégrées par l'entreprise qui transforment les données ou en crée de nouvelles.

À l'intégration d'une nouvelle couche il faut donc prêter une attention particulière aux interdépendances ainsi créées qui influencent les efforts de maintenance et de correction de bugs. Ce point n'est qu'un constat structurel préalable à la création de la solution à proprement parler. Il doit être pris en compte par la nouvelle équipe développement au moment de la création de la solution opérationnelle.

L'équipe développement (DEV) fait donc face à un problème de développement informatique radicalement différent de celui de l'équipe POC : il lui faut créer un logiciel qui soit rigoureux, performant en ressources de calcul, robuste vis-à-vis de bugs. Cela demande un processus de développement lent et rigoureux, avec l'équivalent informatique d'un contrôle qualité et d'une traçabilité des données automatique en interne.

Le problème est que le code produit par l'équipe POC ne répond pas à ces standards : elle s'intéressait avant tout à valider le résultat final au plus vite. Une grande partie de son travail consiste à transformer les données issues de la structure pour en faire des grandeurs plus facilement exploitables par le modèle (le *preprocessing* en anglais). Cette partie est faite à partir d'un code sur-mesure qui se construit par ajout de couches successives. L'approche priorisant l'obtention de résultats rapides se traduit par une méthode de travail concentrée sur l'action qui empêche la planification et la structuration du code pour le rendre compréhensible, ce qui n'est pas nécessaire puisque chaque partie n'est consultée-crée qu'une fois, au moment où le traitement particulier nécessaire à la progression de l'action est fait.

La conséquence directe est que l'équipe DEV doit presque toujours "déconstruire" morceau par morceau le résultat de l'étude de faisabilité pour inférer la finalité et le fonctionnement interne de chacun des composants. L'équipe DEV peut ensuite "reconstruire" en partant de zéro la solution industrielle suivant les standards appropriés. Les deux appellations de "déconstruction" et "reconstruction" proviennent directement du discours tenu par la plupart des clients de Datapred. Ce travail de déconstruction-reconstruction a deux conséquences directes qui ont un impact néfaste sur la probabilité de réussite du projet dans son ensemble :

- Le temps nécessaire à l'opération est important. Les clients de Datapred parlent d'expérience de durées de l'ordre de 6 à 24 mois. Ce délai entre la fin de l'étude de faisabilité et le passage en production effectif vient augmenter la probabilité que le modèle mis en place pendant l'étude de faisabilité soit déjà devenu obsolète lors de la production. L'équipe POC n'utilisant généralement pas de méthodologie appropriée à la prise en compte de la temporalité pour les raisons évoquées ci-dessus, l'entreprise observe alors à ce moment de gros écarts de performances entre les promesses de l'étude de faisabilité et la réalité de l'exploitation. Cet écart a pour conséquence directe de décrédibiliser le travail de l'équipe POC et risque de faire passer la solution sous son seuil de rentabilité.

- L'équipe DEV a entièrement reconstruit la solution suivant ses standards, car la solution initiale lui était incompréhensible. La situation est maintenant inversée : le code de la solution, rédigé suivant les standards du développement informatique industriel, est peu compréhensible pour les membres de l'équipe POC. Ses membres sont incapables de le modifier sans détruire tout le processus de contrôle de qualité et de traçabilité interne de la solution. Le problème est que ce sont les membres de l'équipe POC qui possèdent les compétences nécessaires à la maintenance du modèle intégré à la solution. Donc corriger la solution, pour intégrer de nouveaux éléments ou corriger une obsolescence, se fait en pratique en refaisant une étude de faisabilité à partir de zéro.

On pourrait imaginer qu'avec le temps les deux équipes apprennent à travailler ensemble. C'est en revanche peu probable : Schön (1993) met bien en évidence avec son jeu du silence qu'il faut mettre en place une démarche sociale pour permettre à des personnes de formations différentes de concevoir ensemble. Or ici les équipes POC et DEV ne se fréquentent en pratique presque pas. Si la plupart des entreprises disposent d'employés susceptibles de remplir le rôle de l'équipe développement dans leurs services IT, il n'est pas rare qu'elles doivent sourcer l'équipe POC auprès de services de conseil techniques. De plus, les deux équipes ne travaillent pas simultanément sur les mêmes projets et l'équipe développement est rarement dédiée spécifiquement au transfert des prototypes d'apprentissage automatique en production. On verra dans la partie suivante les évolutions organisationnelles et de design informatique qui ont été mises en œuvre pour améliorer cette situation chez Datapred.

### 3.3. L'après POC (2) : Une relation à l'entreprise cliente à reconstruire.

Comme vu précédemment, le passage du POC au logiciel finalisé implique un changement profond dans l'identité de l'équipe au sein l'entreprise d'IA. Mais il se traduit aussi par un changement majeur concernant les interlocuteurs du client et le mode de relation qui les relie.

La phase de test a essentiellement concerné les professionnels de la prévision du commerce et de la supply chain dans l'entreprise cliente. Elle s'est conclue de manière positive par la validation par la hiérarchie de la Supply Chain de la solution Datapred. Mais cette étape, loin d'être la fin d'un processus décisionnel d'adoption va être au contraire le début d'un processus de négociation interne complexe où la start up n'a plus de prise.

D'abord, le pricing élevé du produit fini ralenti considérablement les négociations, lorsque le POC avait pu avancer rapidement car les budgets en causes n'étaient pas élevés. Les montants des licences obligeaient à une démarche d'achat traditionnel, alors que l'entreprise n'était pas référencée dans le panel fournisseur de l'entreprise.

Ensuite, l'évaluation du travail de développement nécessaire à l'intégration du logiciel Datapred dans les systèmes Auto va lui aussi prendre du temps.

6 mois de négociation seront encore nécessaires pour obtenir l'accord de la Direction commerciale.

Finalement, la structure Auto Digital va décider de récupérer le projet, puis, l'ayant récupéré, va abandonner la collaboration avec Datapred et repartir à zéro avec un nouveau POC.

## 4. RÉCONCILIER LE POC ET LE DÉVELOPPEMENT PRODUIT, L'EXPÉRIENCE DE DATAPRED

Les publications sur les méthodologies de créativité mettent à juste titre l'accent sur l'importance des POC comme dispositifs efficaces voire indispensables à la fois pour l'apprentissage et la démonstration de l'intérêt d'innovations de rupture, comme l'adoption de l'apprentissage automatique. Mais peu s'intéressent au cheminement de l'après POC, là où se retrouvent d'importantes difficultés mis entre parenthèses dans la phase d'expérimentation.

Il faut noter par ailleurs que, sur le plan empirique non plus, cette question n'apparaît pas au-devant de la scène dans la phase initiale de « hype » de la courbe de Gardner. Le démarrage de la croissance de l'entreprise Datapred s'est faite sur ce qu'on pourrait appeler un marché de « pur POC », la question de la transformation

en produits finalisés étant relayée à une seconde phase de la croissance du marché. La conjoncture s'est retournée dans l'année 2017. Les problèmes qui ont été décrits avec Auto n'ont pas été de ce point de vue isolés. Ils ont confronté l'entreprise à l'obligation de trouver des réponses efficaces et rapides à ce problème de « vallée de la mort » entre le POC et la sortie du produit logiciel fini. Ces réponses ont constitué un pivot dans la trajectoire de l'entreprise, impliquant une évolution de la stratégie commerciale, du processus de conception logiciel et de l'organisation interne. A ce jour, après le « trou d'air » de 2017, elle retrouve une dynamique de croissance tout à fait satisfaisante.

Repartons de la schématisation ci-dessous du cycle de croissance de Datapred.

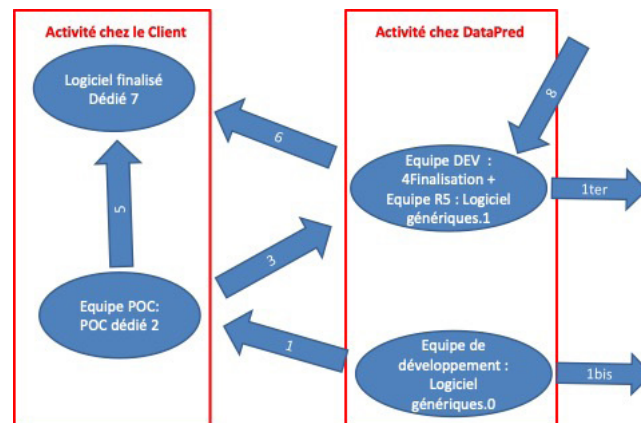


Figure 4 : les étapes du cycle de croissance de Datapred

Il peut se décrire selon la séquence suivante :

- Etape 0 : constitution en interne de l'actif de compétence sur des logiciels génériques d'apprentissage automatique
- Etape 1, 1bis, 1ter... : obtention d'un contrat de POC avec un client
- Etape 2 : réalisation du POC actualisant les logiciels génériques dans le contexte et la problématique du client
- Etape 3 : retour en interne « déconstruction- » du POC
- Etape 4 : Reconstitution du logiciel finalisé pour le client
- Etape 5 : Nouvelle version des logiciels génériques
- Etape 6 et 7 : vente et implémentation du produit fini chez le client
- Etape 8 : Intégration de progrès techniques dans les logiciels génériques.
- Etape 1ter : réalisation d'un POC chez un autre client à partir de la bibliothèque enrichie

Le moteur de ce cycle de croissance tient :

- A la capacité à élargir la compétence d'application en multipliant et diversifiant les POC
- A la capacité à augmenter l'efficacité de réalisation des POC et de la transformation POC/produit en rendant le logiciel générique le plus facilement adaptable
- A augmenter le rendement de l'action commerciale en accumulant les références de succès.
- On vérifie que les POC jouent un rôle clé dans ce cycle de croissance. Mais ils ne sont pas les seuls.

On a vu que Datapred ne visait que les grandes entreprises, ce qui est logique vu que le data analytics est essentiellement un outil de gestion de la complexité se nourrissant d'une masse de données et nécessitant des moyens de calcul significatif. Autant de facteurs que l'on retrouve évidemment dans les grands groupes.

Mais ce choix présente aussi des aspects négatifs qu'on a vus dans l'exemple décrit et que l'on rencontre de manière générale dans les relations entre startup et grands groupes : complexité et incertitude décisionnelle, concurrence interne vs externe, ambiguïté dans la relation aux prestataires entre partenariats d'innovation et mise en concurrence... Plusieurs projets PIC ont analysé ces problèmes par le passé.

Plus spécifiquement, nous avons vu que les problèmes de la relation client venaient de ce que l'on pourrait appeler la « distance » entre le POC et le produit fini : distance technique et temporelle liée au processus de déconstruction/construction de la solution, distance économique liée à l'écart entre le coût modique du POC pour l'entreprise et le coût important de la licence finale, distance organisationnelle, entre les acteurs impliqués dans le POC et les responsables qui décident ou non de généraliser la solution finalisée.

Cela a amené l'entreprise à revoir sa stratégie d'exploration marché comme ses modalités d'engagement des relations commerciales avec ses clients.

Concernant la stratégie Marketing, la stratégie de genericité totale des domaines d'application est certes cohérente avec l'esprit de l'IA et ambitieuse, mais s'est traduite par des difficultés de mise en œuvre dans la pratique. Il est plus facile et rapide de capitaliser (techniquement et en termes de réputation) sur des opérations réussies dans un même domaine que sur des POC issus de contextes très différents comme la gestion des stocks de véhicule, le risque d'impayés sur les crédits à la consommation ou le cours des matières premières. Cela a amené Datapred à recentrer son exploration commerciale sur des « verticales » où l'entreprise avait conduit avec succès l'ensemble du cycle.

Concernant la relation commerciale, il s'agit d'être plus exigeant sur les points d'engagement de la relation :

- Anticiper la négociation pour la vente finale avant l'engagement du POC, en définissant plus précisément au départ les conditions de succès avec le client, en présentant le pricing au départ et en négociant la phase d'installation si le POC est un succès
- En vendant une licence « explore » dès le départ peu chère pour être référencé aux achats et intéresser le département de numérisation au succès du projet.
- Demander un sponsor haut placé dès le départ du POC et une présentation des résultats au COMEX de la grande entreprise.

Concernant la conception logicielle et son organisation :

- La bibliothèque générique est l'actif clé de l'entreprise. La poursuite de l'apprentissage systématique sur ce corpus à partir des POC réussis comme des autres sources technologiques est évidemment central et ne doit pas être relégué en seconde priorité par rapport à une finalisation « adhoc » de solutions clients trop spécialisées. De ce point de vue, le recentrage marketing sur des « verticales métiers » peut libérer du temps occupé auparavant au développement de la solution finale spécialisées et le consacrer à la rationalisation de la bibliothèque générique.
- Sur le plan de l'organisation, la participation de développeurs à des POC, expérimentés dans le cadre du projet de Master, permet de diminuer la distance entre les deux cultures professionnelles, et donc raccourcir globalement la boucle déconstruction/reconstruction de la solution.

## CONCLUSION

Ce cahier se focalise sur le rôle des POC dans le déploiement de l'IA. D'un côté, il a montré l'importance d'autant plus forte de ces « démonstrateurs » (Ben Mahmoud Jouini et al. 2014) que ces initiatives reposent sur des fascinations et que bien peu savent concrètement de quoi il s'agit ni ce qui peut vraiment en sortir. Le POC est donc une étape clé d'apprentissage relativement peu coûteuse pour les clients. C'est en même temps une étape majeure d'apprentissage pour les experts de l'analyse de données qui confrontent leurs techniques génériques aux spécificités d'un questionnement professionnel qu'ils découvrent. C'est enfin une source de financement de l'entreprise, même si dans le modèle économique choisi par l'entreprise Datapred, l'essentiel du revenu est attendu de la vente des licences logicielles.

Mais les POC ne sont pas des prototypes représentatifs du logiciel final. Ce sont des démonstrateurs (Ben Mahmoud Jouini, et al. 2014) qu'il va falloir déconstruire et reconstruire pour en faire un produit informatique intégrable dans les systèmes logiciels des clients. En ce sens, le POC, même réussi, est autant le début d'une conception que la fin d'une autre. Nous avons défini le concept de « distance » entre le démonstrateur et le produit fini pour qualifier plus précisément le « reste à faire » pour franchir l'écart entre le POC et le produit fini. Un concept plus large que celui de « représentativité » qui relie étroitement le prototype au produit fini et qui est ici trop étroit pour rendre compte des différentes dimensions qu'il reste à franchir: dimension technique bien sûr, mais aussi dimension organisationnelle pour partager les compétences entre explorateurs et développeurs dans l'entreprise comme pour aligner les décideurs dans la grande entreprise cliente.

Ce projet de Master PIC a aussi permis d'analyser comment une stratégie de croissance pouvait se construire dans un domaine, celui des mathématiques et de l'informatique appliquées à des problèmes concrets, où se pose en permanence la question de l'articulation entre l'universel et le spécifique des applications. La survie et la croissance d'une startup d'une dizaine de personnes dans un tel contexte tient largement à la capacité à allouer de manière pertinente les ressources rares que constituent les experts entre la compréhension applicative et celle du cœur technologique de la firme.

Enfin, l'analyse de la dynamique de cette entreprise sur les 2 ans et demi qu'ont duré les deux projets de Master successifs chez Datapred montre combien, dans ces domaines technologiques soumis aux vagues de la mode, la plasticité stratégique de la start-up est la condition centrale d'une croissance continue et robuste. L'enchaînement des POC, marchés à part entière dans la phase de Hype (Linden & Fenn 2003), amènerait à l'échec s'ils ne s'inscrivaient pas dans une stratégie d'apprentissage sur un actif technologique spécifique, et si cette stratégie ne savait pas évoluer d'une ambition d'universalité d'application, probablement nécessaire au départ pour « Pitcher » devant des publics de VC, à une stratégie de « Bowling Alley » (Moore, 1995) plus à même de soutenir la trésorerie de l'entreprise dans sa phase actuelle de croissance.

## RÉFÉRENCES

- Schön Donald A. (1993). "Learning to design and designing to learn". Nordisk Arkitekturforskning, 6(1).
- McKinsey Global Institute, (2016) Digital Europe: Pushing the frontier and capturing the benefits, June.
- Linden, A., & Fenn, J. (2003). *Understanding Gartner's hype cycles*. Strategic Analysis Report N° R-20-1971. Gartner, Inc.
- Moore, Geoffrey (1995). *Inside the Tornado: Strategies for Developing, Leveraging, and Surviving Hypergrowth Markets*, HarperBusiness, New York.
- Midler, C. (1986). "Logique de la mode managériale", *Gérer et Comprendre.*, Les Annales des Mines, 1986, pp.74-85.
- Ben Mahmoud-Jouini S., Midler C. (2014) "Creative artefacts How stimulators, demonstrators and prototypes contribute to the creative processes?", *Academy of Management Annual Meeting Proceedings*.



Créé en 2002 à l'École polytechnique, développé ensuite en partenariat avec les prestigieuses Grandes Écoles Françaises que sont Les Mines Paristech, HEC, TelecomParisTech et l'Ensta, le master Projet Innovation Conception analyse et participe à la transformation du management de l'innovation des entreprises à partir d'un dispositif original, associant des entreprises, des enseignants chercheurs reconnus sur le domaine et des étudiants issus de formations supérieures d'ingénieurs et de gestion. Un cursus en alternance permet aux étudiants de s'impliquer pendant au moins un an dans un projet d'innovation des entreprises partenaires, en bénéficiant de l'expertise des enseignants chercheurs associés au Centre de Recherche en Gestion de l'École polytechnique.

Les sujets sont divers et reflètent la variété des problématiques contemporaines de transformation du management de l'innovation : implication dans des projets de rupture, mise en place de cellules visant à déployer des méthodologies nouvelles comme le design thinking, l'open innovation, l'analyse de l'expérience clients ou les transformations numériques, conception et mise en œuvre de concours d'innovation en interne ou en externe, management de communauté d'innovateurs au sein de l'entreprise, le redéploiement international des processus d'innovation, etc... Plus de 300 partenariats ont permis d'explorer ces dynamiques du management de l'innovation dans des contextes d'organisation et selon des points de vue variés : de la grande entreprise du CAC 40 à la jeune start up, de l'innovation produit high tech à l'innovation de service, de l'organisme publique à l'association à mission sociale en passant par l'incubateur ou le fond de capital risque.

L'exigence pédagogique d'un master de recherche, l'implication des entreprises partenaires et l'appui d'experts académiques reconnus permettent, d'une part, une analyse fine et rigoureuse de ces cas, et, d'autre part, une mise en relation étroite de ces pratiques avec les enseignements les plus actuels des sciences de gestion concernés.

Les cahiers du Master PIC ont vocation à diffuser les enseignements de ce dispositif à des publics tant académiques que professionnels dans les entreprises. Ils paraîtront de manière régulière sur le site de L'Observatoire du Master PIC.

—

Percie du Sert, F.G., Midler C., 2019. « D'un projet exploratoire à l'organisation de l'ambidextrie: Le cas ZDS », Les cahiers du Master P.I.C., n° 1, L'Observatoire Projet Innovation Conception – École polytechnique, Paris, France.

Moniot, A., Massé, D., 2019. « La face cachée des programmes d'incubation pour startups: entre vecteurs d'alignement des acteurs et outils de vassalisation ». Les cahiers du Master P.I.C. n° 2, L'Observatoire Projet Innovation Conception – École polytechnique, Paris, France.

Bourdeau, H. Midler C. Petit, C., 2019. « Du concept à la mise en œuvre du machine learning dans les entreprises: L'expérience de Datapred », les cahiers du Master P.I.C., n° 3, L'Observatoire Projet Innovation Conception – École polytechnique, Paris, France.

Blanchard, B. da Motta Cerqueira T. Maniak, R., Midler C. « De l'évaluation au pilotage de la création de valeur des innovations », les cahiers du Master P.I.C., n° 4, L'Observatoire Projet Innovation Conception – École polytechnique, Paris, France.

