



# RECHERCHE & INNOVATION

RAPPORT D'ACTIVITES

V2.0 / 2020

Rostand NYA DJIKI  
Directeur Recherche & Innovation  
T : 06 31 54 16 10

[rostand.nya-djiki@datavalue-consulting.com](mailto:rostand.nya-djiki@datavalue-consulting.com)

# SOMMAIRE

## Table des matières

<b>1. PREAMBULE</b> .....	<b>3</b>
<b>2. NIVEAU DE MATURITÉ</b> .....	<b>3</b>
<b>3. INDICATEURS DE RECHERCHE</b> .....	<b>4</b>
<b>4. PROGRAMMES DE R&amp;D</b> .....	<b>5</b>
4.1. PROJET “SYNAPSE” .....	5
4.2. PROJET “PREDIRH+” .....	6
4.3. PROJET “LORIA” .....	7
4.4. PROJET “CÉPHÉE” .....	8

## 1. PREAMBULE

La Division Recherche & Innovation mène depuis 2017 de nombreux projets de recherche, de développement et d'innovation s'inscrivant en partie, dans une démarche d'excellence opérationnelle des services du groupe DATAVALUE CONSULTING, ceux en matière de fabrication d'outils technologiques comme accélérateurs d'innovation. Forts de leurs ambitions, ils font de cette division, un vrai incubateur de technologies logicielles innovantes tout en s'inscrivant dans un vrai écosystème R&D alliant, dans une logique d'innovation ouverte, des laboratoires de recherche, des éditeurs de logiciels, des utilisateurs finaux et divers organismes de financement de l'innovation.

Au cœur de la stratégie de transformation digitale menée chez nos clients, notre laboratoire conçoit des solutions disruptives à partir des résultats de nos travaux de recherche et de notre veille technologique active. Notre R&D est ouverte à des partenariats externes afin d'être au plus proche d'organisations créatives et de pouvoir éprouver des programmes nationaux et internationaux.

"L'Innovation est dans l'ADN de DATAVALUE CONSULTING".

## 2. NIVEAU DE MATURITÉ

Nos projets de R&D et d'Innovation sont supervisés sur la base des critères de l'échelle TRL (Technology Readiness Level) dans le but de consolider leur niveau de maturité technologique respectif. Ci-contre une classification de nos programmes majeurs à ce jour :

Code Projet	Description / Date	Niveau de maturité (TRL)
<b>SYNAPSE</b>	NOUVEAU FORMALISME DE CODAGE ET D'ORCHESTRATION DE SCRIPTS D'ARCHITECTURE LAMBDA/KAPPA POUR LES SYSTEMES EMBARQUES (IoT) Date de démarrage : 09/2017 Date de fin : <b>En cours</b>	TRL4 - Les composants technologiques de base sont intégrés afin d'établir que toutes les parties fonctionnent ensemble. C'est une "basse fidélité" comparée au système final. Les exemples incluent l'intégration 'ad hoc' du matériel en laboratoire.
<b>PREDiRH+</b>	SURVEILLANCE PREDICTIVE DES RISQUES RH (SOCIO-PROFESSIONNELS) Date de démarrage : 09/2019 Date de fin : <b>En cours</b>	TRL2 – Le projet est à ses débuts. Une fois les principes de base observés, les applications pratiques peuvent être inventées. L'application est spéculative et il n'y a aucune preuve ou analyse détaillée pour étayer cette hypothèse. Les exemples sont toujours limités à des études papier.
<b>LORIA</b>	PLATEFORME D'ANALYSE ET DE SYNTHÈSE AUTOMATIQUE DE LA LITTÉRATURE SCIENTIFIQUE Date de démarrage : 02/2020 Date de fin : <b>En cours</b>	TRL2 – Le projet est à ses débuts. Une fois les principes de base observés, les applications pratiques peuvent être inventées. L'application est spéculative et il n'y a aucune preuve ou analyse détaillée pour étayer cette hypothèse. Les exemples sont toujours limités à des études papier.
<b>CEPHEE</b>	GENERATEUR AUTOMATIQUE DE TASK FORCE PROJET Date de démarrage : 09/2019 Date de fin : <b>En cours</b>	TRL5 – La fidélité de la technologie s'accroît significativement. Les composants technologiques basiques sont intégrés avec des éléments raisonnablement réalistes afin que la technologie soit testée dans un environnement simulé.

### 3. INDICATEURS DE RECHERCHE

Nos travaux de R&D sont rattachées à des opérations de R&D ayant créé une technique prouvée originale et/ou meilleure que celles existantes. En effet, les techniques et procédés eux-mêmes sont décrits et pourraient être réutilisés. En général, nos opérations apparaissent lorsqu'une difficulté a été identifiée dans le cadre d'activités de l'entreprise et c'est sa résolution qui constitue l'opération de R&D.

Nos travaux font l'objet de publications scientifiques, d'articles médias grand public, ainsi que de brevet d'invention. Nous communiquons sur nos projets lors de conférences internationales et salons à thèmes spécifiques. Dans cet élan de dissémination, ci-dessous quelques-unes de nos parutions à venir :

- [1] R. Nya Djiki, F. Mely, "Generalized Reward Function based on crude Insights for Financial Portfolio Strategies". In proceeding for IJCAI-2020 AI in FinTech, January 2021, Japan.
- [2] R. Nya Djiki, A. Loukili, N. Joudar, "New semantic layer for automatic code generation based IaaC script languages". In proceeding for IEEE International Conference on Embedded Software and Systems, August 2021.
- [3] R. Nya Djiki, S. Ka, " Automatic generation of Project Task Force based on reconciliation of skills files". In proceeding for tne 35th Conference on Neural Information Processing Systems, IE NeurIPS-SpicyFL 2021.

Nous travaillons d'autre part, sur le dépôt d'enveloppes Soleau.

## 4. PROGRAMMES DE R&D

Cette section présente des résumés de nos projets majeurs identifiés cette année 2020.

### 4.1. PROJET “SYNAPSE”

#### NOUVEAU FORMALISME DE CODAGE ET D’ORCHESTRATION DE SCRIPTS D’ARCHITECTURE LAMBDA/KAPPA POUR LES SYSTEMES EMBARQUES (IoT)

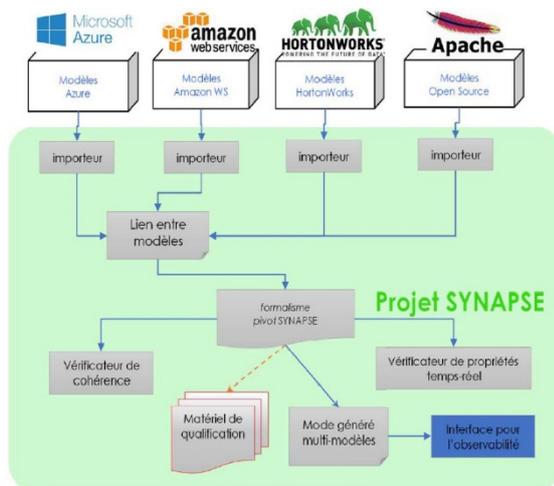


Figure 1 : Vue globale du projet SYNAPSE

Faciliter le développement logiciel des systèmes temps réel critiques dédiés au traitement de grands volumes de données en s’appuyant sur l’ingénierie orientée modèle, telle est l’ambition du projet SYNAPSE. Ce projet de recherche et d’innovation vise précisément à mettre au point une couche sémantique sophistiquée opérant sur un noyau d’intelligences artificielles hybrides développé sous licence libre, basé sur des standards d’interopérabilité ouverts, et capable de :

(i) Traiter de façon cohérente des sous-ensembles pertinents des langages de modélisation comportementaux et architecturaux conformes au modèle de calcul « Globalement Asynchrone, Localement Synchrone » (GALS) ; telles que les architectures Data temps réel en Cloud Azure, Amazon et Google.

(ii) Optimiser un modèle sémantique évolutif et formel assurant un bon équilibre entre expressivité et décidabilité pour la génération automatique de code pivot pour des systèmes embarqués.

(iii) Fournir un kit de qualification contenant les exigences, les plans de développement et de vérification, les cas de test, la traçabilité... pour permettre l’utilisation de la technologie SYNAPSE dans un contexte de certification compatible avec les normes DO-178B/C (aéronautique), ECSS-E-ST- 40C/Q-ST-80C (spatial) et ISO-26262 (automobile). La technologie SYNAPSE a vocation d’une part, à aider les DSI, les BE (Bureaux d’Etudes)... Dans la construction d’architectures Big Data « Lean » basée sur des modèles sémantiques sous contraintes élaborées ; et d’autre part, à contribuer aux initiatives d’interopérabilité et coopération telles que SESAR, INTERESTED et OPEES pour la consolidation de l’écosystème français et européen en matière d’outils.

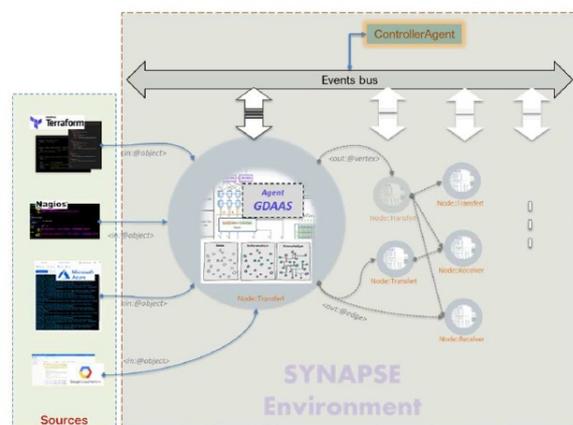


Figure 2 : Module d’Auto-mapping structurel

Nous atteignons un niveau TRL4 cette année, et poursuivons nos efforts d’expérimentations et de mise au point du concept sur la base de données réelles d’architectures Cloud hybrides.

*Technologies sources d’appui : Red Hat® OpenShift®, Kubernetes, Python, Ada, Python, C\C++, Nodejs.*

## 4.2. PROJET “PREDIRH+”

### SURVEILLANCE PREDICTIVE DES RISQUES RH

Face à la démocratisation du télétravail notamment accélérée en partie par la pandémie de COVID-19, de nombreuses organisations adoptent de plus en plus des mécanismes de maintien de la productivité principalement dirigés par les technologies numériques et du digital. Face cette mutation structurelle du monde du travail, comment aider les entreprises ou organisations à mieux renforcer la maîtrise de leurs risques RH?

La technologie PREDiRH+ ambitionne renforcer les performances de pilotage RH en exposant des services tactiques pour l’aide à la prise de décision stratégique. Ces derniers tirent leur originalité à travers un mécanisme sophistiqué de croisement de données multi-sources et d’évaluation d’indicateurs prédictifs capables de déceler et de rendre exploitables certains facteurs, certaines

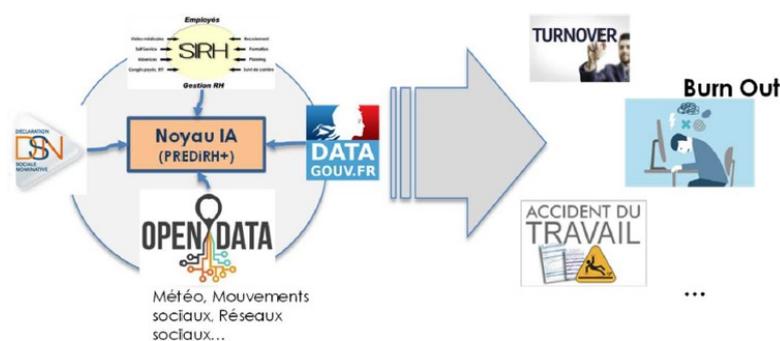


Figure 3 : Concept global du projet PREDiRH+

situations ou comportements déclenchant non perceptibles actuellement car non corrélés. Un des défis techniques et scientifiques de ce projet réside en la définition d’un modèle formel de données à partir de l’exploitation des données de DSN (Déclaration Sociale Nominative), de SIRH (Système d’Information RH), d’environnement de travail (individu, organisation, tâches, moment, lieu, équipements...), ainsi que de l’OpenData (météo, mouvements sociaux...).

PREDiRH+ opère une stratégie de contrôle anticipatif de certains risques majeurs socio-professionnels, en se basant d’une part, sur un réseau d’interconnexions d’agents prédictifs ; et d’autre part, sur une technique de fusion d’indicateurs formels. Le projet tente de définir une couche graphique générique et interactive à destination de RRH (Responsables des Ressources Humaines) ou Responsables d’équipes ou groupes de travail. Cette couche exposerait des tableaux de bords de visualisation de ces indicateurs prédictifs concourant à expliquer chaque thématique des risques RH (Burn Out; Accident du travail; Turn Over; Arrêts maladie...).



Au stade de TRL2 cette année, nous posons les bases d’expérimentations sous hypothèses dans le but de renforcer la pertinence du projet auprès d’acteurs des métiers des RH. Les exemples sont actuellement limités à des études papier.

*Technologies sources d’appui : Nifi, Python, Vue.js, Angular.js.*

### 4.3. PROJET "LORIA"

#### PLATEFORME D'ANALYSE ET DE SYNTHÈSE AUTOMATIQUE DE LA LITTÉRATURE SCIENTIFIQUE

Les crises posées par la pandémie de COVID-19 ont été de divers ordres, entre autres sociétales, économiques et sanitaires. C'est sur ce dernier volet en particulier que nous avons souhaité contribuer à l'effort de résilience humanitaire, en formalisant un projet de recherche dans le domaine l'E-santé.

Exposer des informations intelligibles et fondamentales à partir d'un brassage efficient de la littérature scientifique issue d'une thématique de santé, dans le but de faciliter, d'accélérer et de permettre une

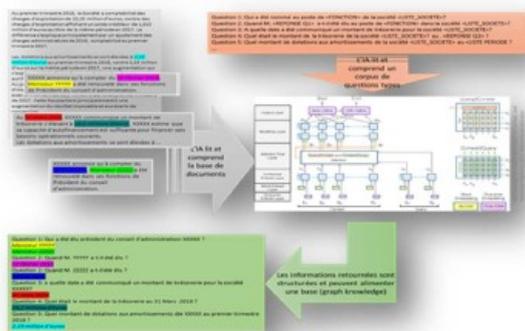


Figure 4 : Concept analytique LORIA

prise de décision pour la recherche médicale ; telle est l'ambition du projet LORIA. Il vise plus précisément à générer en automatique des résumés d'articles scientifiques issues d'un tri sophistiqué dans des bases bibliographiques conventionnelles, à partir d'un intitulé de thème de recherche prédéfini. De façon moins conceptuelle, il s'apparente à proposer un moteur de recherche bibliographique augmentée d'une stratégie de production automatisée de résumés d'articles scientifiques.

La technologie LORIA ambitionne s'appuyer sur un réseau novateur d'agents artificiels hybrides,

distribués et explicables, combinant des modèles connexionnistes et symboliques, pour fournir des patterns sémantiques enrichies et capitalisées à travers une base de données orientée graphe. Cette dernière alimenterait un générateur de texte qui se charge de produire des verbatims moyennant un modèle sémantique évolutif et formel assurant un bon équilibre entre expressivité et décidabilité.

Sur un sujet donné, "SARS-CoV-2" par exemple, la technologie LORIA analyse un réseau de co-publications et de citations type bioRxiv.org ; pubMed.ncbi.nlm.nih.gov... pour fournir d'une part des index bibliographiques récents et pertinents ; et d'autre part, un résumé de la compréhension automatique opérée pour chacun de ces articles.

LORIA ambitionne être utilisée par un public très large, allant des chercheurs et chercheuses aux entreprises, et sur des thématiques variées. En TRL 2 cette année, le projet LORIA se structure via des concepts et cas d'applications formulés.

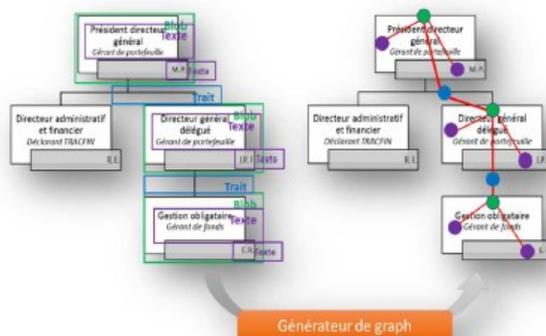


Figure 5 : Modélisation orientée Graphe

*Technologies sources d'appui : SpaCy, JanusGraph.*

## 4.4. PROJET “CÉPHÉE” GENERATEUR AUTOMATIQUE DE TASK FORCE PROJET

La transformation digitale des entreprises connaît un essor de plus en plus important. Son émergence a entériné la prise de conscience par les organisations, d’un mouvement nécessaire pour s’adapter aux nouvelles contraintes de l’environnement concurrentiel et à la modification des comportements et des attentes des consommateurs. Face à la recrudescence de ces projets digitaux, le projet CEPHEE vise à élaborer un dispositif serviciel technologique capable d’opérer une analyse profonde des interrelations sémantiques entre des entités nommées d’une base de cartographies structurelles et des contraintes de spécifications fonctionnelles, dans le but de générer des scénarios optimaux. Ce projet vise plus précisément à proposer un ordonnanceur d’analyse sémantique sophistiquée capable de générer à la volée, des séquences de dispositifs projet les plus adéquats à la réalisation de missions décrites dans un cahier des charges, sous contraintes d’indices prédéfinis de qualification.

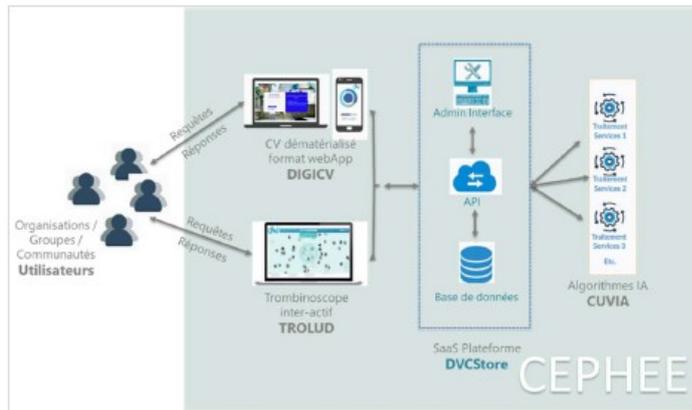


Figure 6 : Vue globale du projet CEPHEE

Le projet CEPHEE est structuré en 4 principaux sous-projets :

- (i) DIGICV : Dispositif de dématérialisation des dossiers de compétences sous la forme d’un modèle multi-dimensionnel ou ontologie propice à l’élaboration d’un graphe sémantique de connaissance. Ce volet prépare la cartographie des compétences disponibles dans un système de management de profils de compétences d’une organisation. Il se charge de la mise au point d’un modèle sémantique évolutif assurant un bon équilibre entre expressivité décidabilité pour le matching.
- (ii) TROLUD : Dispositif Interface homme-machine ludique pour la visualisation interactive de cartes auto-adaptatives de graphes générées à partir de modèles de liens établis.
- (iii) CUVIA : Noyau IA dédiés aux traitements analytiques des modèles de données, à la génération des graphes de connaissance et à la génération temps-réel de labels de groupes.
- (iv) DVCSTORE : Plateforme SaaS d’exposition des services CEPHEE pour souscription à abonnements, pour gestion des organisations clientes et l’administration générale.



Figure 7 : Aperçu TROLUD

*Technologies sources d’appui : Node.js, Express, Python, Webgl.*



**DataValue**<sup>®</sup>  
CONSULTING

[www.datavalue-consulting.com](http://www.datavalue-consulting.com)

**Contactez-nous** : 01 81 89 41 70 - [contact@datavalue-consulting.com](mailto:contact@datavalue-consulting.com)

5 Boulevard du Général de Gaulle - 92120 Montrouge