



Institut de Recherche  
en Informatique de Toulouse

# Utilisation d'Osirim

## Rapport annuel

### 2020

Les responsables des projets hébergés par la plateforme sont invités à fournir un bref bilan d'activité annuel. Ce document rassemble les éléments qui ont été retournés au titre de l'année 2020. Il n'est que partiel dans la mesure où tous les rapports attendus n'ont pas pu être collectés au moment de produire ce compte-rendu.

28/02/2021

## Table des matières

Laboratoire Interdisciplinaire Sciences Innovations Sociétés (LISIS).....	<b>2</b>
Laboratoire Cognition, Langues, Langage, Ergonomie (CLEE) .....	<b>3</b>
IRIT : Département Signaux et Images.....	
Equipe SAMOVA .....	<b>4</b>
Equipe MINDS .....	<b>10</b>
Equipe SC .....	<b>11</b>
IRIT : Département Gestion des Données .....	
Equipe SIG .....	<b>14</b>
Equipe IRIS .....	<b>20</b>
Equipe PYRAMIDE .....	<b>25</b>
IRIT : Département Intelligence Artificielle .....	
Equipe MELODI .....	<b>26</b>
Equipe ADRIA .....	<b>28</b>
IRIT : Département Calcul intensif, Simulation, Optimisation .....	
Equipe APO .....	<b>30</b>
Equipe REVA .....	<b>31</b>

# Laboratoire Interdisciplinaire Sciences Innovations Sociétés (LISIS)

Titre du projet : MUSK

Responsable : Nicolas Turenne

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Mes activités portent sur l'extraction terminologique et la distribution lexicale à travers l'analyse des médias sociaux. Le média qui focalise mon attention est Youtube.

L'idée consiste à étudier l'activisme: comme le zadisme ou le lancement d'alerte.

Pour cela il faut mettre en place une plateforme de traitement dont les modules principaux sont :

- le crawling
- le filtrage
- l'extraction lexicale
- la classification automatique

Dans ce projet on cherche à comprendre la structuration du domaine, les catégories mises en jeu, la portée politique.

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

L'espace de données et l'espace des caractéristiques peuvent être assez importants, ce qui nécessite une parallélisation des traitements. C'est ce que j'ai fait jusqu'à présent à travers une configuration technique R/Hadoop/MongoDB. La quantité de vidéos extraites peut avoisiner 500,000. Il en va de même pour l'espace de description terminologique (entités nommées, dépendances relationnelles, groupes nominaux, bigrammes). Avec de tels espaces de descriptions sur-dimensionnés, les approches de structuration nécessitent aussi une parallélisation. L'utilisation de la plateforme OSIRIM (slurm/hadoop) a permis de produire une chaîne de traitements capables de procéder à des traitements quantitatifs prometteurs.

# Laboratoire Cognition, Langues, Langage, Ergonomie (CLEE)

Titre du projet : SemDis

Responsable : Franck Sajous

- Activités scientifiques réalisées sur la plateforme OSIRIM :

CLLE recourt à la plateforme pour effectuer des calculs ayant trait à la sémantique distributionnelle. La plateforme est également utilisée en amont et en aval (construction et l'évaluation des différents modèles) pour constituer des ressources linguistiques : corpus textuels (web, Tweeter, etc.), divers lexiques et ressources lexicographiques (e.g. Wiktionary dans ses différentes versions de langue). Ces ressources peuvent être annotées sur la plateforme (e.g. analyse syntaxique automatiquement).

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

- corpus de Tweets utilisé pour la conception de modèles distributionnels (BERT, word2vec)
- lexique dérivationnel Glawinette, construit pour alimenter la base de données morphologique Démonette
- dictionnaire électronique ENGLAWI (version XML normalisée de Wiktionary)

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

- **Hathout, N., Sajous, F., Calderone, B. and Namer, F. (2020).** Glawinette: a linguistically motivated derivational description of French acquired from GLAWI. *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, pp. 3877-3885, Marseille, France.
- **Huyghe, R., & Wauquier, M. (2020).** What's in an agent? A distributional semantics approach to agents nouns in French. *Morphology*, 30, 185-218.
- **Miletic, F., Przewozny-Desriaux, A. & Tanguy, L. (2020).** Collecting Tweets to Investigate Regional Variation in Canadian English. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*. Marseille, France.
- **Sajous, F., Calderone, B. & Hathout, N. (2020).** ENGLAWI: From Human- to Machine-Readable Wiktionary. *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)*, pp. 3016-3026, Marseille, France.
- **Wauquier, M., Hathout, N., & Fabre, C. (2020).** Contributions of distributional semantics to the semantic study of French morphologically derived agent nouns. In J. Audring, N. Koutsoukos & C. Manouilidou (Eds), *Rules, patterns, schemas and analogy, MMM12 Online Proceedings*, 12, 111-121.
- **Wauquier, M., Hathout, N., & Fabre, C. (2020).** Semantic Discrimination of Nominalizations based on Distributional and Statistical Clues. *Journal of Word Formation/Zeitschrift für Wortbildung*, 4 (2), 100-119.

## IRIT : Equipe SAMOVA

### Titre du projet : OPRA : Objective amplification gain Prescription Rules based on Automatic speech recognition

Responsable : Julien Pinquier

#### - Activités scientifiques réalisées sur la plateforme OSIRIM

L'objectif du projet OPRA est d'optimiser le réglage de prothèses auditives pour la perte spécifique d'un patient donné, en mesurant le bénéfice apporté par le réglage sur la performance d'un moteur de reconnaissance automatique de la parole. En 2019, l'étude OPRA1 a montré une amélioration du confort auditif lors de l'évaluation sur les patients. Cependant, seules des variations de Gains d'Insertion (IG) à 65 dB étaient testées. Dorénavant des variations conjointes des IG à 65 dB, des IG à 85 dB, des seuils de compression (CT), des ratios de compression (CR) et des vitesses de compression sont testés. Comme nous étions limités par le temps de traitement imparti, diverses approches de recherche aléatoire ont été implémentées (dichotomique et génétique). Afin de réduire ce temps, les algorithmes s'exécutent en parallèle, exploitant au maximum les capacités d'OSIRIM (utilisation de 48 CPU conjointement)

Ce travail est une collaboration entre l'IRIT (équipe SAMoVA) et la société Archean Technologies.

#### - Résultats obtenus grâce aux services de la plateforme OSIRIM

Grâce à notre nouvelle approche sur la plateforme OSIRIM, le temps de calcul moyen est divisé par un facteur 5 au minimum par rapport à une exécution séquentielle sur un ordinateur standard. À l'issue de milliers de tests réalisés de manière automatique sur OSIRIM, notre nouvelle approche fournit théoriquement de meilleures améliorations que la précédente. Pour attester ces résultats, elle sera évaluée prochainement sur une trentaine de patients en déterminant pour chacun la meilleure des 2500 configurations possibles.

#### - Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

- Fontan, L., Le Coz, M., Azzopardi, C., Stone, M. A., & Füllgrabe, C. (2020). Improving hearing-aid gains based on automatic speech recognition. *The Journal of the Acoustical Society of America*, 148(3), EL227-EL233.
- Fontan, L. (mars 2020). Using automatic speech recognition to improve hearing-aid fitting. *Speech in noise conference (SPIN 2020)*, Toulouse, France.
- Fontan, L., Le Coz, M., Farinas, J., Segura, B., Stone, M. A., Füllgrabe, C. (2020). Using automatic speech recognition to improve hearing-aid fitting. *Speech in noise conference (SPIN 2020)*, Toulouse, France.

IRIT : Equipe SAMOVA

Titre du projet : LINTO

Responsable : Isabelle Ferrané

- Activités scientifiques réalisées sur la plateforme OSIRIM

Les travaux réalisés par l'équipe SAMoVA dans le cadre du projet LinTo portent sur l'analyse de données audio-visuelles correspondant à des réunions de travail entre plusieurs participants. Nous nous sommes intéressés d'abord à la détection du locuteur actif, à la signature audio des locuteurs, la signature visuelle des participants et la fusion des deux pour obtenir une signature audio-visuelle de chaque participant. Ceci dans le but d'extraire des informations sur la structure des réunions et sur l'interaction entre participants, en faisant abstractions des informations personnelles (pas de reconnaissance/identification des personnes).

- Résultats obtenus grâce aux services de la plateforme OSIRIM

Les capacités de la plateforme ont permis d'utiliser des architectures de Deep Learning pour les appliquer à des corpus de données existants (Corpus AMI) puis de "fine tuner" ces méthodes sur les données spécifique du projet. Trois types de modalités ont été exploitées : l'analyse acoustique de l'audio, la détection du buste et le flot optique.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

- Jorge Francisco Madrigal Diaz, Frédéric Lerasle, Lionel Pibre, Isabelle Ferrané. Audio-Video detection of the active speaker in meetings. IEEE International Conference on Pattern Recognition (ICPR 2020), IAPR : International Association of Pattern Recognition, Jan 2021, Milan (virtual), Italy. ([hal-03125600](#))

IRIT : Equipe SAMOVA

Titre du projet : SAMODI

Responsable : Isabelle Ferrané

- Activités scientifiques réalisées sur la plateforme OSIRIM

Les travaux réalisés par l'équipe SAMoVA dans le cadre du projet SAMODI ont concerné l'utilisation de Snorkel pour l'apprentissage faiblement supervisé. Appliqué à la base textuelle OPUS, la définition de règles a permis d'annoter des séquences de sous-titres en niveau de difficulté lexicale et syntaxique. Ceci a été fait dans le cadre d'un troisième stage co-encadré en 2020 par les équipes SAMOVA et MELODI.

- Résultats obtenus grâce aux services de la plateforme OSIRIM

Les capacités de la plateforme osirim ont permis l'utilisation des fonctionnalités de snorkel pour l'appliquer à un très gros volume de données.

IRIT : Equipe SAMOVA

Titre du projet : ALAIA

Responsable : Isabelle Ferrané

- Activités scientifiques réalisées sur la plateforme OSIRIM

Les travaux réalisés par l'équipe SAMoVA dans le cadre du Laboratoire Commun ALAIA portent sur la détection, la localisation et la caractérisation d'erreurs de prononciation dans un contexte d'apprentissage de langue étrangères. Cela nécessite l'apprentissage de modèles acoustique de la langue cible en vue du décodage phonétique des énoncés des apprenants. La deuxième phase consiste à utiliser ses résultats pour apprendre puis repérer des types d'erreurs de prononciations. L'effort réalisé en 2020 a surtout porté sur l'annotation de corpus.

- Résultats obtenus grâce aux services de la plateforme OSIRIM

Les capacités de la plateforme n'ont pas encore été exploitées car il était nécessaire de préparer les données au préalable. Le travail sur l'apprentissage et le développement de système de détection et de localisation d'erreur va débuter en 2021.



IRIT : Equipe SAMOVA

Titre du projet : Lightly-supervised and Unsupervised Discovery of Audio Units using Deep Learning (LUDAU), ANR-18-CE23-0005-01

Responsable : Thomas Pellegrini

- Activités scientifiques réalisées sur la plateforme OSIRIM

Mise en oeuvre d'un nouvel algorithme d'optimisation de seuils pour la classification multi-label (SGLThresh)

Dans le cadre de la thèse de Léo Cances :

- Entraînement et inférence de réseaux de neurones profonds sur des tâches de détection d'événements sonores.
- Recherche des meilleurs hyper-paramètres
- Tests de nouvelles méthodes en condition d'apprentissage légèrement supervisé et semi-supervisé.

Dans le cadre du stage M2 de Etienne Labbé (continuation en thèse) :

- Mise en oeuvre de nouveaux algorithmes de deep learning semi-supervisé : Mixmatch et Fixmatch

- Résultats obtenus grâce aux services de la plateforme OSIRIM

- Participation au challenge DCASE 2020 à la tâche 6 "légendage automatique audio", *audio captioning*, 6ème place sur 11 équipes internationales,
- Amélioration des performances dans la situation d'apprentissage légèrement/semi supervisé

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

- Pellegrini, T., Zimmer, R., & Masquelier, T. (2020). Low-activity supervised convolutional spiking neural networks applied to speech commands recognition. arXiv preprint arXiv:2011.06846.
- Pellegrini, T. (2020). IRIT-UPS DCASE 2020 audio captioning system. Technical report. [http://dcase.community/documents/challenge2020/technical\\_reports/DCASE2020\\_Pellegrini\\_131\\_t6.pdf](http://dcase.community/documents/challenge2020/technical_reports/DCASE2020_Pellegrini_131_t6.pdf)

IRIT : Equipe SAMOVA

Titre du projet : thèses CIFRE Lucile Gelin

Responsables : Julien Pinquier et Thomas Pellegrini

- Activités scientifiques réalisées sur la plateforme OSIRIM

Lucile Gelin travaille sur des systèmes de reconnaissance automatique de la parole d'enfants.

- Résultats obtenus grâce aux services de la plateforme OSIRIM

Elle entraîne des modèles acoustiques à l'aide de Kaldi et de PyTorch pour les modèles de type end-to-end

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

- Gelin, L., Daniel, M., Pellegrini, T., Pinquier, J. (2020) Reconnaissance de phones fondée sur du Transfer Learning pour des enfants apprenants lecteurs en environnement de classe, Proc. JEP-TALN-RECITAL 2020

## Equipe MINDS

Titre du projet : Co-encadrement de la thèse de Karen Sanchez de l'Université de Bucaramanga, en Colombie.

Responsables du projet : Denis Kouamé et Adrian Basarab

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Karen a utilisé Osirim pour entrainer des réseaux de neurones convolutifs pour la classification d'images tomographiques pulmonaires, dans le but de détecter des pneumonies. Ce projet a été réalisé en collaboration avec le CHU de Toulouse, qui a fourni les données. L'idée a été de fusionner deux bases de données existantes afin d'améliorer les résultats de classification sur une faible base d'apprentissage.

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

Osirim a permis les nombreux tests d'apprentissage.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- Subspace-Based Domain Adaptation using Similarity Constraints for Pneumonia Diagnosis within a Small Chest X-Ray Image Dataset . Sanchez, K.; Hinojosa, C.; Arguello, H.; Freiss, S.; Sans, N.; Kouamé, D.; Meyrignac, O.; and Basarab, A. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Nice, 13/04/2021-16/04/2021*, 2021..

## Equipe SC

Titre du projet : "Compression Onboard Of Satellite Images using neural NETworks" (COOSINET)

Responsable : Marie Chabert

### - Activités scientifiques réalisées sur la plateforme OSIRIM :

Dans ce projet, nous utilisons la plateforme OSIRIM pour entraîner et tester des architectures neuronales destinées à la compression d'images. Nous avons essentiellement testé des architectures à base d'autoencodeurs, éventuellement variationnels faisant intervenir trois couches convolutionnelles associées à des non-linéarités de type GDN (Generalized Divisive Normalization). L'objectif des simulations effectuées sur la plateforme est tout d'abord d'adapter ces algorithmes de la littérature à des images d'observation de la Terre fournies par le CNES dont le format diffère de celui des images naturelles pour lesquelles ces algorithmes ont été conçus. Dans un deuxième temps, nous avons opéré des modifications sur ces architectures pour pouvoir les adapter aux contraintes bord : ressources réduites en termes de calcul et de mémoire, Nous avons effectué un réglage fin de chaque paramètre de l'architecture d'origine: nombre de filtres, taille des filtres, fonctions d'activation et surtout le modèle de distribution requis pour la partie variationnelle de l'autoencodeur.

### - Résultats obtenus grâce aux services de la plateforme OSIRIM :

La plateforme OSIRIM a permis de montrer que les architectures neuronales proposées dans la littérature récente ont des performances supérieures, en termes de débit et de distorsion, aux compresseurs utilisés actuellement à bord des satellites d'observation de la Terre. Des expériences réalisées sur la plateforme ont également permis de montrer que leur complexité pouvait être réduite en vue d'une implantation bord, notamment en réduisant le nombre de filtres constituant les couches convolutionnelles. La plateforme nous a permis de tester différentes combinaisons de paramètres d'autoencodeur afin de proposer des modifications conduisant à une complexité réduite avec un impact minimal sur les performances résultantes. Récemment, la plateforme nous a également permis d'entraîner de nouveaux modèles pour effectuer le débruitage conjointement à la compression. Les résultats préliminaires indiquent que les modèles entraînés pour cette nouvelle tâche montrent des performances supérieures à celle des architectures actuellement embarquées.

### - Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- Article paru dans les actes de la conférence OBPDC 2020 (International Workshop on board payload data compression) en septembre 2020: « *Simplified entropy model for reduced-complexity end-to-end variational auto-encoder with application to on-board satellite image compression* », V. Alves de Oliveira, T. Oberlin, M. Chabert, C. Poulliat, M. Bruno, C. Latry, M. Carlavan, S. Henrot, F. Falzon, and R. Camarero, 7th International

Workshop on On-Board Payload Data Compression, OBPDC2020, CNES-ESA, 21-23 September 2020, Online Event.

- Article de revue paru dans le numéro spécial du MDPI Remote Sensing, intitulé « Remote Sensing Data Compression » en janvier 2021: Alves de Oliveira, V.; Chabert, M.; Oberlin, T.; Poulliat, C.; Bruno, M.; Latry, C.; Carlavan, M.; Henrot, S.; Falzon, F.; Camarero, R. Reduced-Complexity End-to-End Variational Autoencoder for on Board Satellite Image Compression. *Remote Sens.* **2021**, *13*, 447. <https://doi.org/10.3390/rs13030447>

Equipe SC

Titre du projet : DeepRad - Multiple target extraction, identification and tracking for automotive radar

Responsable : Thomas Oberlin

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Mise au point et test de modèles d'apprentissage profond pour analyser et traiter les données radar. Utilisation du JupyterLab de la plateforme (en CPU la plupart du temps) pour prototyper les modèles, analyser rapidement la base de données et y faire quelques ajustements. Les scripts étaient développés en local et une fois testés les simulations étaient lancées dans des containers singularity, le plus souvent le soir.

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

Une comparaison des modèles de l'état de l'art pour des données radar

IRIT : Equipe SIG

Titre du projet : Thèse CIFRE Clément Lejeune

Responsable : Olivier Teste, Josiane Mothe

- Activités scientifiques réalisées sur la plateforme OSIRIM :

La détection d'anomalies dans les time-series multivariées est un enjeu dans de nombreux domaines dont l'aéronautique dans lequel se situe cette thèse. Nous nous intéressons à la détection des phénomènes anormaux, communément appelés valeurs aberrantes ou anomalies.

Notre approche repose sur la théorie des données fonctionnelles à partir desquelles il est possible d'observer le comportement des time-series au travers de la forme de leurs courbes. Nous développons plusieurs indicateurs géométriques afin de pouvoir détecter les anomalies présentes dans les time-series

- Résultats obtenus grâce aux services de la plateforme OSIRIM

Comparaisons entre nos algorithmes iFor(Vmap) iFor(Curvmap) iFor(Lenmap) OCSVM(Vmap) OCSVM(Curvmap) OCSVM(Lenmap) et deux compétiteurs

- [1] FUNTA (Kuhnt, S., Rehage, A., 2016. An angle-based multivariate functional pseudo-depth for shape outlier detection. *Journal of Multivariate Analysis* 146, 325–340)
- [2] Dir.out (Dai, W., Genton, M.G., 2019. Directional outlyingness for multivariate functional data. *Computational Statistics and Data Analysis* 131, 50–65)

Utilisation de deux jeux de données réelles de référence

- [1] ECG dataset (Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E., 2000. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation* 101, e215–e220)
- [2] PenDig dataset (Dua, D., Graff, C., 2017. UCI machine learning repository. URL:<http://archive.ics.uci.edu/ml>.)

Et un jeu de données synthétiques construit en fonction de

- [1] 5 modèles de génération en suivant l'approche développée par Dai, W., Genton, M.G., 2019. Directional outlyingness for multivariate functional data. *Computational Statistics and Data Analysis* 131, 50–65

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- *C. Lejeune, J. Mothe, A. Soubki, O. Teste, Shape-based outlier detection in multivariate functional data, International Journal of Knowledge-Based Systems, Elsevier Science Publisher, 2020*  
<https://doi.org/10.1016/j.knosys.2020.105960>

IRIT : Equipe SIG

Titre du projet : TRANSPARENS

Responsable : Imen Megdiche

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Nous étudions la classification d'hémorragies cérébrales par des algorithmes centralisés et fédérés. Le dataset exploité est disponible sur ce lien (<https://www.kaggle.com/c/rsna-intracranial-hemorrhage-detection>). Ce dataset contient 752,803 scanner dont 14% présentant des hémorragies. Il a été stocké et traité par des algorithmes deep learning selon plusieurs architectures (CNN, Resnet).

- Résultats obtenus grâce aux services de la plateforme OSIRIM

Les travaux en mode centralisé atteignent une accuracy de 80% pour un temps de traitement ~ 67 heures . Ces résultats approchent les résultats de certains travaux de la littérature.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- O. El Rifai, M. Biotteau, X. Deboissezon, I. Megdiche, F. Ravat, O. Teste, Blockchain-based Federated Learning in Medecine, International Conference on Artificial Intelligence in Medicine (AIME'20), Minneapolis, USA, 2020  
[https://doi.org/10.1007/978-3-030-59137-3\\_20](https://doi.org/10.1007/978-3-030-59137-3_20)
- *O. El Rifai, M. Biotteau, X. Deboissezon, I. Megdiche, F. Ravat, O. Teste, Blockchain-Based Personal Health Records for Patients' Empowerment, 14th International Conference on Research Challenges in Information Science (RCIS'20), Limassol, Cyprus, 2020*  
[https://doi.org/10.1007/978-3-030-50316-1\\_27](https://doi.org/10.1007/978-3-030-50316-1_27)



## IRIT : Equipe SIG

### Titre du projet : Convention Recherche SGE

### Responsable : André Péninou, Olivier Teste

#### - Activités scientifiques réalisées sur la plateforme OSIRIM :

Le but de cette étude est de mettre au point une méthode pour automatiser et consolider le processus de détection d'anomalies dans les séries temporelles issues de réseaux de capteurs.

Pour ceci, nous avons développé un algorithme, appelé « Composition based Decision Tree » (CDT), qui permet de produire automatiquement des règles de détection d'anomalies dans un ensemble de séries temporelles univariées. L'algorithme détermine par un processus bayésien un ensemble de motifs de détection de points remarquables dans les séries temporelles. Un arbre de décision est ensuite automatiquement construit en considérant les nœuds comme des compositions de motifs. L'arbre est converti en un ensemble de règles de décision, compréhensibles par les experts.

#### - Résultats obtenus grâce aux services de la plateforme OSIRIM

Comparaisons entre l'algorithme CDT et deux algorithmes compétiteurs de l'état de l'art

- [1] PART (Daud Nor Ridzuan and Corne David Wolfe. 2009. Human readable rule induction in medical data mining. In Proceedings of the European Computing Conference (vol.1 ed.), Vol. 27 LNEE. 787–798)
- [2] JRip (Ian H. Witten and Eibe Frank. 2005. Data Mining: Practical Machine Learning Tools and Techniques (2nd ed.). Morgan Kaufmann)

Utilisation de deux jeux de données réelles

- [1] Les datasets du SGE
- [2] The Yahoo datasets Nikolay Laptev and Saeed Amizadeh. 2015. A labeled anomaly detection dataset S5 Yahoo Research, v1. <https://webscope.sandbox.yahoo.com/catalog.php?datatype=s&did=70>

Un dépôt GitHub de l'approche est disponible : <https://github.com/IBK-TLS/CDT>.

#### - Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- I. Ben Kraiem, F. Ghazzi, A. Péninou, G. Roman-Jimenez, O. Teste, Human-Interpretable Rules for Anomaly Detection in Time-series, EDBT/ICDT Joint Conference - Nicosia, Cyprus, 23-26 Mars 2021.

IRIT : Equipe SIG

Titre du projet : Deep learning pour des images multi-échelles

Responsable : Josiane Mothe

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Recherche sur les méthodes d'apprentissage profond pour la détection dans des images de différentes résolutions :

- Des images satellites d'observation de la Terre en lien avec le projet FabSpace en particulier pour la détection de changements plus particulièrement dans des zones forestières,
- Des images médicales (WSI) pour l'aide à la détection de cancers

- Résultats obtenus grâce aux services de la plateforme OSIRIM

- Collections de données annotées d'observation de la Terre combinée avec des textes descriptifs
- Evaluation de méthodes d'apprentissage profond sur des collections de référence d'images médicales
- Compréhension des approches d'apprentissage profond : mise au point de méthodes pour l'étude de l'impact de la distribution des jeux de test

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- Ismat Ara Reshma, Sylvain Cussat-Blanc, Radu Tudor Ionescu, Hervé Luga, Josiane Mothe. Natural vs Balanced Distribution in Deep Learning on Whole Slide Images for Cancer Detection. *The 36th ACM/SIGAPP Symposium on Applied Computing (SAC '21)*, Mar 2021, Virtual Event, South Korea. [10.1145/3412841.3441884](https://doi.org/10.1145/3412841.3441884). [hal-03057209](https://hal.archives-ouvertes.fr/hal-03057209).

## IRIT : Equipe SIG

### Titre du projet : Recherche d'information, réseaux sociaux et résumés automatiques

### Responsable : Josiane Mothe

#### - Activités scientifiques réalisées sur la plateforme OSIRIM :

Moteur de recherche multi-cœur : sélection de la configuration adaptée en fonction des contextes de la requête de l'utilisateur et prédiction de la difficulté des requêtes ; fouille de réseaux sociaux et résumés automatiques

#### - Résultats obtenus grâce aux services de la plateforme OSIRIM

- Evaluation d'une méthode de sélection de noyaux de moteurs de recherche d'information,
- Evaluation de modèles de fouille de réseaux sociaux
- Evaluation de modèle de détection de signaux faib

#### - Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- Faneva Ramiandrisoa, Josiane Mothe. Aggression Identification in Social Media: a Transfer Learning Based Approach Second Workshop on Trolling, Aggression and Cyberbullying, European Language Resources Association (ELRA), May 2020, Marseille, France. pp.26-31. hal-02635019v1
- Faneva Ramiandrisoa, Josiane Mothe. IRIT at TRAC 2020 Second Workshop on Trolling, Aggression and Cyberbullying, European Language Resources Association (ELRA), May 2020, Marseille, France. pp.49-54. hal-02635076v1
- Faneva Ramiandrisoa, Josiane Mothe. Early Detection of Depression and Anorexia from Social Media: A Machine Learning Approach Circle 2020, Iván Cantador; Max Chevalier; Massimo Melucci; Josiane Mothe, Jul 2020, Samatan, France. hal-02877723v1
- Josiane Mothe, Pratik Parikh, Faneva Ramiandrisoa. IRIT-PREVISION AT HASOC 2020 : Fine-tuning BERT for Hate Speech and Offensive Content Identification Hate Speech and Offensive Content Identification in Indo-European Languages (HASOC@FIRE 2020), Dec 2020, Hyderabad (virtual), India. hal-03040547v1
- Sebastien Déjean, Radu Tudor Ionescu, Josiane Mothe, Md Zia Ullah. Forward and Backward Feature Selection for Query Performance Prediction

35th Annual ACM Symposium on Applied Computing (SAC 2020), Mar 2020, Brno, Czech Republic. pp.690-697, [10.1145/3341105.3373904](https://doi.org/10.1145/3341105.3373904)hal-02942304v1

## Equipe IRIS

### Titre du projet : LISTIC

### Responsable : Guillaume CABANAC

#### - Activités scientifiques réalisées sur la plateforme OSIRIM :

LisTIC (« Liens socionumériques et Technologies (mobiles) de l'Information et de la Communication ») est un projet de recherche interdisciplinaire coordonné par des chercheurs en sociologie du LISST de l'université de Toulouse 2 en partenariat avec des chercheurs en sciences de l'information et de la communication et en informatique affiliés au LERASS et à l'IRIT de l'université de Toulouse 3. Ce projet est financé de 2016 à 2020 par l'Agence Nationale de la Recherche (projet ANR-16-CE26-0014-01) et a obtenu le label du GDRI Web Science du CNRS. Les différentes enquêtes réalisées dans le cadre de ce projet, à l'aide des méthodes des humanités numériques, ont été mises en place à l'aide de la plateforme OSIRIM.

Le projet LisTIC explore les usages contemporains des téléphones mobiles et des applications développées par les principaux médias sociaux dans le but de mieux cerner leurs effets sur la composition et la morphologie des réseaux relationnels, sur les sociabilités tissées au travail et sur les formes médiatisées de la participation politique, que ce soit en France durant les élections présidentielles ou dans le cadre des mouvements sociaux actuels au Brésil.

Pour cartographier ces pratiques, le projet LisTIC propose de développer les méthodologies des humanités numériques en exploitant les ressources offertes par les smartphones afin d'en documenter les usages, notamment les formes nomades de la participation aux applications mobiles des réseaux socionumériques (mSNS). L'objectif est de faire converger des protocoles de recherche susceptibles de nous livrer un aperçu exhaustif de l'appropriation contemporaine de ces technologies et de la manière dont elles infléchissent, voire renouvellent, nos structures sociales et relationnelles. Pour relever cet ambitieux défi sociologique, des méthodologies innovantes sont développées en lien étroit avec les méthodes classiques des SHS et leurs acquis théoriques pour favoriser au maximum la cumulativité des connaissances.

L'hypothèse de ce projet revient à percevoir les smartphones et les déclinaisons nomades des réseaux socionumériques comme le creuset technologique où se forge la propagation d'une forme sociale qui tend à être dominante aujourd'hui, celle d'un « individualisme en réseaux », soit une manière d'être et de tisser des liens sociaux de plus en plus travaillée de manière individualisée sous l'amplification des sollicitations sociales et relationnelles émanant des TIC. Cette hypothèse sociologique est explorée dans le cadre de quatre recherches qui ont été conçues pour appréhender trois modes d'appropriation des smartphones :

- mSNS & Réseaux : la manière dont les usages nomades des réseaux socionumériques viennent s'articuler avec les pratiques téléphoniques et leurs effets sur la morphologie des réseaux de relations personnels
- E-participation en France et au Brésil : les formes nomades de la e-participation aux mouvements sociaux (un cas en France, un cas au Brésil) et la manière dont elles favorisent une hybridation des arènes publiques à l'interface entre les territoires urbains et numériques
- mSNS & Travail : l'inscription des usages des réseaux socionumériques dans l'écologie des activités professionnelles et la manière dont la disponibilité permanente des affinités

relationnelles personnelles vient concurrencer les sociabilités professionnelles d'autant plus vécues comme « contraintes ».

Dans ce cadre, OSIRIM héberge depuis novembre 2016 une machine virtuelle sur laquelle les services suivants sont déployés par les partenaires IRIT :

- site web du projet : <https://listic.irit.fr>
- moissonneur Twitter DMI-TCAT (<https://github.com/digitalmethodsinitiative/dmi-tcat/wiki>)
- moissonneur Facebook : développement interne basé sur Netvizz (<https://apps.facebook.com/netvizz>)
- plugin d'enquêtes de sociologie (développement interne) basé pour LimeSurvey (<https://listic.irit.fr/limesurvey>)
- gestionnaire de code SVN

#### - Résultats obtenus grâce aux services de la plateforme OSIRIM :

La première phase du projet a consisté à moissonner des données depuis novembre 2016, liées à l'élection présidentielle française de 2017. Afin de questionner l'usage des médias sociaux par des activistes en politique, ces données sont collectées à partir de :

- Twitter : 42 millions de tweets (72 Go) relatifs,
- Facebook : 450 000 posts (4 Go) liés à 81 groupes et 214 pages.

L'analyse des données moissonnées a permis d'informer la constitution de l'échantillon des personnes enquêtées par entretien individuel. Elles permettent de confronter les discours des différents partis politiques (par analyse lexicométrique), d'identifier les individus influenceurs (par analyse résiliaire), les moments de rupture dans la campagne (par analyse de séries temporelles), les pratiques des activistes (par analyses de pratiques égocentrées), etc.

Nos résultats de recherche sont publiés dans les actes des conférences :

- ACM Hypertext & Social Media (HT 2018),
- AAAI International Conference on Weblogs and Social Media (ICWSM 2018),
- International Conference on the Theory of Information Retrieval (ICTIR 2017)

et ont fait l'objet de plusieurs communications, notamment au Datapol organisé par Sciences Po Paris : <http://www.medialab.sciences-po.fr/projets/datapol/> .

#### - Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

----- revues internationales -----

- J. Figeac, P. Ratinaud, N. Smyrniotis, G. Cabanac, O. Fraissier, T. Salord et F. Seffusatti. « Mobile phones in the spread of unreliable information on Twitter. Evidence from the 2017 French presidential campaign ». *Mobile Media and Communication* (2021). to appear. doi:10.1177/2050157920972157. SAGE.
- J. Figeac, N. Smyrniotis, T. Salord, G. Cabanac, O. Fraissier, P. Ratinaud et F. Seffusatti. « Information-sharing practices on Facebook during the 2017 French presidential campaign: An “unreliable information bubble” within the extreme right ». *Communications* 45.s1 (2020), p. 648-670. doi:10.1515/commun-2019-0193. de

Gruyter.

----- revues nationales -----

- F. Bousquet, J. Figeac, G. Cabanac et C. Noûs. « Pratiques de médiations informationnelles sur Facebook: l'appropriation politique de la presse quotidienne régionale lors de la campagne présidentielle de 2017 ». *Mots: Les langages du politique* 123 (2020), p. 81-102. doi:10.4000/mots.26797. ENS Éditions.
- J. Figeac, P. Ratinaud, N. Smyrnaiois, G. Cabanac, O. FRaisier-Vannier, T. Salord et F. Seffusatti. « Les téléphones mobiles, un outil de désinformation ? La circulation des informations peu fiables dans Twitter lors de la campagne présidentielle française de 2017 ». *tic&société* 14.1-2 (2020), p. 371-400. doi:10.4000/ticetsociete.5413.
- J. Figeac, T. Salord, G. Cabanac, O. Fraisier, P. Ratinaud, F. Seffusatti et N. Smyrnaiois. « Facebook favorise-t-il la désinformation et la polarisation idéologique des opinions ? » *Questions de communication* 36 (2019). Presses Universitaires de Nancy – Éditions universitaires de Lorraine, p. 167-187. doi : 10.4000/questionsdecommunication.21149.

## Equipe IRIS

### Titre du projet / activité de recherche :

- Activités de thèse de Sourty Rapahël dans le cadre du projet CIFRE IRIT-RENAULT (2019-2022) : Modèles d'apprentissage de représentations de textes augmentées par les graphes de connaissances.
- Activités de thèse de Nicolas Bizzoréro dans le cadre du projet ANR CoST (2019-2022) : Modèles Seq2Seq pour la prédiction de requête suivante en recherche d'information.
- Activités de stage (M2) de Jesus Levon Melgarejo dans le cadre du projet ANR CoST : Etude du phénomène d'"oubli catastrophique" des modèles neuronaux de recherche d'information..

### Responsable : Lynda Tamine Lechani

#### - Activités scientifiques réalisées sur la plateforme OSIRIM :

##### 1- Thèse de Sourty Raphaël (2019-2022)

L'objectif des travaux de thèse de Sourty Raphaël est de développer des modèles d'apprentissage de représentations de textes qui intègrent les connaissances issues de ressources externes (ex. WordNET, Free Base, etc.). Les modèles utilisés sont basés sur les réseaux de neurones et mécanismes de distillation des connaissances (knowledge distillation) entre modèles.

Dans ce cadre, la plateforme OSIRIM est exploitée pour : 1) entraîner des modèles de représentations d'entités issus de graphes de connaissances ; 2) réaliser le mécanisme de distillation entre modèles et évaluer son apport dans des tâches de prédiction de lien.

##### 2- Thèse de Nicolas Bizzorrero (2019-2022)

L'objectif des travaux de thèse de Nicolas Bizorréro est de définir des modèles neuronaux d'accès à l'information adaptées à des tâches complexes de recherche d'information (multi-sujets, multi-étapes, etc.). Précisément, une première activité est de proposer des modèles de prédiction de la requête suivante selon l'historique de recherche basé sur la notion de session. Dans ce cadre, la plateforme OSIRIM est exploitée pour entraîner des modèles de l'état de l'art pour la prédiction de requête, basés sur l'apprentissage profond.

##### 3- Stage de Jesus Levon Melgarejo (avril-septembre 2020)

L'objectif de ce stage était de mener une étude empirique du phénomène d'"oubli catastrophique" dans le cas des modèles neuronaux de recherche d'information. Dans ce cadre, la plateforme OSIRIM est exploitée pour entraîner des modèles neuronaux de l'état de l'art pour la recherche d'information.

#### - Résultats obtenus grâce aux services de la plateforme OSIRIM :

- Modèle de transfert de connaissance entre ressources (bases de connaissances) basé sur le mécanisme de distillation



- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

- Raphael Sourty, Jose G Moreno, François-Paul Servantes, Lynda Tamine, Knowledge Base Embedding By Cooperative Knowledge Distillation, COLING 2020
- Jesus Levon Melgarejo, Laure Soulier, , Karen Pinel-Sauvagnat, Lynda Tamine: Studying Catastrophic Forgetting in neural ranking models, European Conference on Information Retrieval ECIR (Papier long) 2021

Equipe PYRAMIDE

Titre du projet : Stage M2 CSA de Ibrahim Mukhtar

Responsables du projet : Shaoyi Yin

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Ibrahim Mukhtar a comparé la consommation de ressources de Hive et Spark pour un même workload (un benchmark de type TPC-H). Une estimation de coûts monétaires a été faite. Le but final du travail est de développer un nouveau benchmark qui prend en compte les SLA et permet de comparer d'une part la rentabilité des SGBDs multi-tenants (côté fournisseur) et d'autre part la dépense financière d'un tenant en utilisant des différents services DBaaS

IRIT : Equipe MELODI

Titre du projet : SLANT

Responsable : Philippe Muller

- Activités scientifiques réalisées sur la plateforme OSIRIM

Collecte de données dans le cadre du projet ANR SLANT: articles de journaux en anglais et en français, dans le but de faire des analyses de l'expression linguistique du biais selon la source et son orientation politique.

- Résultats obtenus grâce aux services de la plateforme OSIRIM

- Collecte encore en cours, mais recueil pour l'instant d'environ 1M d'articles répartis sur quelques milliers de sources différentes
- Quelques analyses préliminaires

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

Non applicable: le projet a débuté à l'été 2020.

IRIT : Equipe MELODI

Titre du projet : thèse CIFRE de Jordane Dorne avec Thales Alenia Space Toulouse

Responsable : Nathalie Aussenac-Gilles

- Activités scientifiques réalisées sur la plateforme OSIRIM

Dans le cadre de ma thèse « Représentation sémantique de données geo-spatiales pour l'analyse de changements » j'utilise la plateforme OSIRIM.

La plateforme OSIRIM me permet de travailler sur le corpus de tweets hébergé sur celle-ci. Ce corpus me permet de pouvoir rechercher l'ensemble des tweets en rapport avec un lieu donné et deux dates.

J'utilise principalement la machine 48CPU via jupyter Notebook car il s'agit de la recherche de texte dans des fichiers JSON ce qui nécessite majoritairement des ressources « read » sur le disque et non des calculs.

J'ai écrit un script Python que j'exécute dans le notebook qui prend en paramètre le nom de l'entité à rechercher dans les tweets et les deux dates qui servent de bornes.

Une fois les tweets trouvés, je les regroupe dans un seul fichier JSON que je télécharge via le service SFTP sur mon PC.

- Résultats obtenus grâce aux services de la plateforme OSIRIM

Le résultat obtenu est un fichier JSON composé uniquement des tweets contenant l'entité recherchée entre les deux dates. Ce fichier me permet ensuite, d'extraire les mots-clés les plus fréquents afin de qualifier un événement.

A ce jour, je n'ai travaillé que sur les incendies et les résultats sont probants.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme

- <https://doi.org/10.1145/3423336.3429347>.

## Equipe ADRIA

Titre du projet : Argumentation abstraite: le passage à l'échelle

Responsable : MC Lagasque-Schiex

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Réalisation d'expérimentations dans le cadre de la thèse de Mickaël Lafages. Le but de ces expérimentations est de comparer la vitesse de résolution de problèmes d'argumentation par différents solveurs. Ces solveurs sont pour la plus part multithreadés et les problèmes étudiés longs à résoudre. La plateforme OSIRIM est donc essentielle pour l'avancement de ce travail.

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

Grace aux expérimentations réalisées sur OSIRIM, on a pu mettre en évidence l'aptitude des algorithmes développés par Mickaël Lafages à résoudre très rapidement certains problèmes d'argumentation. Pour certaines catégories, ces algorithmes donnent en effet de meilleurs résultats que tous les solveurs jusqu'ici testés.

Les expérimentations ont également permis l'amélioration de ces algorithmes, notamment en tirant un meilleur parti de la parallélisation.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- [Sylvie Doutre](#), [Mickaël Lafages](#), [Marie-Christine Lagasque-Schiex](#). *A Distributed and Clustering-Based Algorithm for the Enumeration Problem in Abstract Argumentation (regular paper)*. Dans : *International Conference on Principles and Practice of Multi-Agent Systems (PRIMA 2019), Torino, Italy, 28/10/2019-31/10/2019*, Matteo Baldoni, Mehdi Dastani, Beishui Liao, Yuko Sakurai, Rym Zalila Wenkstern (Eds.), [Springer](#), LNAI 11873, p. 87-105, 2019.

## Equipe ADRIA

Titre du projet : Raisonnement Explicite et Apprentissage sur imagerie Médicale (REAM)

Responsable\_ : Florence Dupin de Saint-Cyr Bannay

- Activités scientifiques réalisées sur la plateforme OSIRIM :

OSIRIM a été utilisée dans le cadre de la thèse de Mona MAYOUF pour faire tourner les algorithmes d'apprentissage profond pour la classification automatique des images histologiques du cancer du sein en exploitant la base de données publique BreakHis qui a été chargée dans le répertoire du projet (la base comprend environ 8000 images de 460x700 pixels, et après les augmentations réalisées elle occupe 5 Go). Un premier travail avec un large éventail de protocoles de tests a été mené dans le but de sélectionner la meilleure manière de préparer les données. Ces expérimentations sont très gourmandes en GPU, l'environnement offert par OSIRIM nous est indispensable. De plus, la possibilité d'utiliser plusieurs noyaux sur OSIRIM, ainsi que d'installer des bibliothèques supplémentaires, nous a facilité plusieurs étapes techniques et nous a épargné beaucoup de tracas d'installations.

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

L'étude de l'impact des différentes combinaisons de préparation des données sur la qualité d'apprentissage, nous a permis de sélectionner une combinaison performante pour la classification automatique des images histopathologiques du cancer du sein, ce travail en cours d'expansion est en cours de rédaction. Nous avons également pu mettre en évidence une nouvelle technique d'alimentation du réseau de neurones qui optimise l'apprentissage en exploitant des connaissances de haut-niveau.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- Nous avons soumis en février 2021 au journal "IEEE TRANSACTIONS ON MEDICAL IMAGING" un article décrivant nos résultats sous le titre "Incremental Deep Learning on BreakHis DataSet" auteurs Mona Mayouf, Florence Dupin de Saint-Cyr et Tim van de Cruys..

Equipe APO

Titre du Projet : DAML

Responsable : Serge Gratton

- Activités scientifiques réalisées sur la plateforme OSIRIM :

OSIRIM a été utilisé pour entraîner des réseaux de neurones profonds dans le cadre de l'assimilation de données. Ces réseaux de neurones sont implémentés en Pytorch et, pour obtenir les meilleurs résultats possibles, nous sommes allés chercher les limites des NVIDIA GTX 1080 TI.

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

Grâce aux réseaux de neurones, nous avons montré que l'apprentissage machine génère des méthodes d'assimilation de données plus efficaces que les meilleurs algorithmes conçus manuellement. Plus précisément, cela témoigne qu'il est possible d'apprendre avec très peu d'hypothèses comment estimer l'évolution des distributions de probabilités de l'état d'un système dynamique chaotique. Les applications de ce travail sont, entre autres, la prévision météorologique.

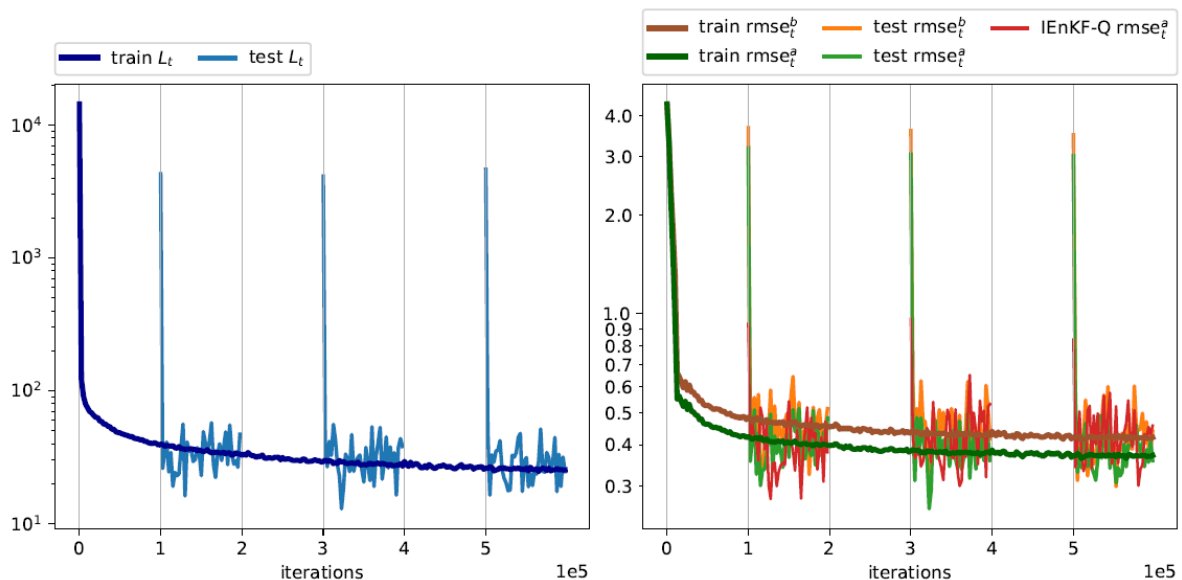


Figure 1: Le graphique de droite montre l'évolution des performances d'assimilation du réseau au cours de l'apprentissage. En vert clair les performances en test, en vert foncé les performances en entraînement et en rouge les performances de l'état de l'art.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- Un papier intitulé `_DAN - An optimal Data Assimilation framework based on machine learning Recurrent Networks_` est actuellement en révision dans SIMODS (SIAM journal on mathematics of data science)

## Equipe REVA

Titre du Projet : Estimation de posture 3D à partir de données imprécises et incomplètes

Responsable : Thibault Blanc-Beyne

- Activités scientifiques réalisées sur la plateforme OSIRIM :

Les travaux réalisés ont pour but de construire un système d'estimation de la posture 3D d'opérateurs humains à partir d'images de profondeur, dans un contexte industriel. Ce système, divisé en trois parties, est basé sur l'utilisation de plusieurs modèles de réseaux de neurones. Tout d'abord, nous entraînons un algorithme d'apprentissage à l'aide d'une segmentation automatique dont les meilleurs échantillons sont automatiquement sélectionnés au cours de l'entraînement, nous permettant d'extraire l'opérateur de l'image. En parallèle, nous construisons un réseau de neurones léger et optimal pour l'estimation de la posture 3D de l'opérateur sur des images de profondeur générées numériquement. Ce modèle n'étant pas utilisable sur les données acquises dans les centres de tri, nous rendons les images de profondeur de synthèse plus réalistes à l'aide d'un réseau de neurones génératif. Ces nouvelles images sont utilisées pour réentraîner l'algorithme d'estimation de posture, qui permet finalement d'obtenir des résultats convaincants sur les images réelles.

- Résultats obtenus grâce aux services de la plateforme OSIRIM :

L'utilisation de la plateforme nous a permis de tester différents modèles de réseaux de neurones sur nos données, et d'élaborer les hyperparamètres optimaux.

- Références bibliographiques des articles et publications présentant des résultats ou des contenus dérivant de l'utilisation de la plateforme :

- Thibault Blanc-Beyne, Estimation of 3D pose from inaccurate and incomplete data: Application to human operators' activity analysis in a sorting center, thèse de l'Université de Toulouse, novembre 2020.
- Thibault Blanc-Beyne, Axel Carlier, Sandrine Mouysset, Vincent Charvillat, Unsupervised Human Pose Estimation on Depth Images, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Ghent, Belgique, septembre 2020
- Thibault Blanc-Beyne, Axel Carlier, Vincent Charvillat, Iterative dataset filtering for weakly supervised segmentation of depth images, ICIP, Taipei, Taïwan, septembre 2019.