

Compte-rendu d'activité sur la plateforme OSIRIM

Année 2017

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 2017. 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.

[Projet SemDis \(Franck Sajous, sajou@univ-tlse2.fr\)](mailto:sajous@univ-tlse2.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

CLLE-ERSS recourt à la plateforme pour l'analyse syntaxique automatique de divers corpus de grande taille.

Ces corpus analysés sont ensuite utilisés dans le cadre de travaux portant sur la sémantique distributionnelle, notamment en utilisant différentes versions de word2vec.

Les analyses syntaxiques produites sur Osirim sont aussi utilisées pour l'enrichissement de lexiques électroniques.

Résultats obtenus grâce aux services de la plateforme OSIRIM :

L'utilisation de la plateforme a permis de produire des versions de corpus français et anglais analysés syntaxiquement par l'analyseur Talismane (<http://redac.univ-tlse2.fr/applications/talismane.html>).

Ces analyses sont un préalable aux travaux portant sur l'analyse distributionnelle.

OSIRIM a également été utilisée dans un système de prédiction de l'accent en italien et son application à GLAFF-IT, un lexique morphophonologique de cette langue.

Publications :

- B. Calderone, M. Pascoli, F. Sajous and N. Hathout. (2017)
Hybrid Method for Stress Prediction Applied to GLAFF-IT, a Large-Scale Italian Lexicon.
In: Gracia J., Bond F., McCrae J., Buitelaar P., Chiarcos C., Hellmann S. (eds) Language, Data, and Knowledge. LDK 2017, Lecture Notes in Computer Science, vol 10318, pp. 26-41. Springer, Cham.
- Ho-Dac L.-M. and Laippala V. (2016).
Le corpus WikiDisc : ressource pour la caractérisation des discussions en ligne.
In Ciara Wigham & Gudrun Ledegen (Eds.) Médias sociaux et les corpus de communication médiée par les réseaux, collection 'Humanités numériques', l'Harmattan : Paris
- N. Hathout and F. Sajous. (2016).
Wiktionnaire's Wikicode GLAWIfied: a Workable French Machine-Readable Dictionary.
Proceedings of the tenth International Conference on Language Resources and Evaluation (LREC 2016), pp. 1369-1376, Portorož, Slovenia.
http://fsajous.free.fr/papers/LREC2016/Hathout_Sajous_LREC2016_GLAWI.pdf
- F. Sajous and N. Hathout. (2015).
GLAWI, a free XML-encoded Machine-Readable Dictionary built from the French Wiktionary.
Proceedings of eLex 2015 conference, pp. 405-426, Herstmonceux, England.
http://fsajous.free.fr/papers/ELEX2015/Sajous_Hathout_ELEX2015_GLAWI.pdf

Equipe MELODI (Tim Van De Cruys, tim.vandecruys@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

L'équipe MELODI utilise le cluster OSIRIM pour effectuer diverses tâches de traitement automatique du langage naturel (TAL). Une première tâche consiste à l'annotation syntaxique automatique (syntactic parsing) de corpora massifs. Le cluster a été utilisé pour annoter automatiquement jusqu'à 3 milliards de mots de textes en anglais et jusqu'à 2 milliards de mots de textes français. En raison du fait que le processus est facilement parallélisable, le cluster est parfaitement adapté à la tâche.

Une deuxième tâche consiste à la construction d'espaces de mots vectorielles. Dans un modèle espace-mot, les différents mots d'une langue sont encodés dans un modèle d'espace vectoriel selon les contextes dans lesquelles ces mots apparaissent. Un tel modèle permet de comparer (calculer la similarité entre) les différents mots, ce qui est utile pour les applications sémantiques. De nouveau, la construction est fortement parallélisable, ce qui le rend avantageux d'utiliser le cluster.

En dehors de ces deux tâches importantes, le cluster a été utilisé pour la construction de modèles tensoriels (ce qui permet la modélisation de phénomènes de langage aux co-occurrences multiples), et pour diverses applications, telles que la désambiguïsation du sens de mots et le calcul du sens de mots dans leur contexte.

Finalement, le cluster, et notamment le base de tweets français, a été utilisé pour effectuer de l'apprentissage non-supervisé.

Publications :

- Nicholas Asher, Tim Van de Cruys, Antoine Bride, Márta Abrusán. 2016. Integrating Type Theory and Distributional Semantics: A Case Study on Adjective–Noun Compositions, *Computational Linguistics*, 42:4:703–725.
- Damien Sileo, Camille Pradel, Philippe Muller, Tim Van de Cruys. 2017. Système d'ensemble pour la classification de tweets. DEFT2017. https://deft.limsi.fr/2017/actes_DEFT_2017.pdf
- Damien Sileo, Camille Pradel, Philippe Muller, Tim Van de Cruys. 2017. Synapse at CAP 2017 NER challenge: Fasttext CRF. CAP2017. <https://arxiv.org/pdf/1709.04820.pdf>

Equipe SIG, Learning to Rank, Tweets (Josiane Mothe, josiane.mothe@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

Techniques d'apprentissage automatique pour l'apprentissage d'ordonnement (learning to rank).

Application sur l'apprentissage des configurations de systèmes de recherche d'information.
Application sur l'apprentissage d'ordonnement des documents dans les moteurs de recherche d'information

Résultats obtenus grâce aux services de la plateforme OSIRIM :

Des collaborations : avec l'université de Montréal, l'INSA de Lyon

Publications :

- Anthony Bigot, Sebastien Déjean, Josiane Mothe.
Learning to Choose the Best System Configuration in Information Retrieval: the case of repeated queries.
Dans / In : Journal of Universal Computer Science, Consortium J.UCS, Numéro spécial / Special issue : Information Retrieval and Recommendation, Vol. 21 N. 13, p. 1726–1745, décembre / december 2015.
- Liana Ermakova, Josiane Mothe.
La structure thème-rhème pour l'ordonnement de documents en recherche d'information.
Dans / In : Document numérique, Hermès, 2017 (à paraître).
- Léa Laporte, Sebastien Déjean, Josiane Mothe.
Sélection de variables en apprentissage d'ordonnement : évaluation des SVM pondérés.
Dans / In : Document numérique, Hermès, Numéro spécial / Special issue : Evaluation en Recherche d'Information, Vol. 18/1 – 2015, p. 99–121, 2015.
- Romain Deveaud, Josiane Mothe, Jian-Yun Nie.
Learning to Rank System Configurations (short paper).
Dans / In : Conference on Information and Knowledge Management (CIKM 2016), Indianapolis, United States, 24/10/16–28/10/16, ACM, p. 2001–2004, octobre / october 2016.
- Josiane Mothe, Mahdi Washha.
Predicting the Best System Parameter Configuration: the (Per Parameter Learning) PPL method. (regular paper).
Dans / In : International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2017), Marseille, 06/09/17–08/09/17, Elsevier, p. 1–10, 2016.

Equipe SIG, Projet Depression (Josiane Mothe, josiane.mothe@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

Détection des personnes susceptibles d'être dépressives dans un réseau social.

Résultats obtenus grâce aux services de la plateforme OSIRIM :

Un modèle de détection

Publications :

- Idriss Abdou Malam, Mohamed Arziki, Mohammed Nezar Bellazrak, Farah Benamara, Assafa El Kaidi, Bouchra Es-Saghir, Zhaolong He, Mouad Housni, Véronique Moriceau, Josiane Mothe, Faneva Ramiandrisoa.
IRIT at e-Risk (regular paper).
Dans / In : International Conference of the CLEF Association, CLEF 2017 Labs Working Notes, Dublin, Ireland, September, 11/09/17-14/09/17,
Linda Cappellato, Nicola Ferro, Lorraine Goeuriot, Thomas Mandl (Eds.), 201

Projet PRINCESS (Yoann Pitarch, pitarch@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

Développement d'un système novateur pour effectuer une tâche de RI adhoc. Le système est basé sur un principe de compétitions entre documents. De nombreuses expérimentations ont été réalisées sur la plateforme pour valider la pertinence et l'efficacité du modèle.

Résultats obtenus grâce aux services de la plateforme OSIRIM :

Des résultats encourageants qui améliorent l'état de l'art.

Publications :

- G. Hubert, Y. Pitarch, K. Pinel-Sauvagnat, R. Tournier, et L. Laporte
TournRank: When Information Retrieval Becomes Document Competition.
A paraître dans Information Processing and Management (IPM) fin 2017 ou début 2018

Equipe IRIS, Thèse de Gia-Hung Nguyen : apprentissage de représentation pour la recherche d'information (Lynda Tamine Lechani, lynda.lechani@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

L'objectif des travaux de thèse de Gia est d'exploiter, adapter et développer des modèles d'apprentissage profond basé sur les réseaux de neurones pour 1) améliorer les représentations des textes en intégrant les connaissances issues de ressources sémantiques externes, exemples : WordNet, UMLS; 2) améliorer l'efficacité des modèles d'ordonnement de documents en réponse à une requête utilisateur en se basant sur des architecture de type 'end to end'.

Dans ce cadre, la plateforme OSIRIM est exploitée pour : 1) indexer des collections de documents volumineuses (eg. Gov2, PubMed) et des ressources (WordNet,, UMLS); 2) exploiter les librairies (eg. Indri, Lucene, Tensor Flow, Torch) offertes pour implémenter les modèles neuronaux; 3) lancer les programmes d'apprentissage, test et les analyses des sorties associées

Résultats obtenus grâce aux services de la plateforme OSIRIM :

- Modèles d'ordonnement end-to-end avec des appariements combinant les mots et concepts
- Techniques d'apprentissage joint des concepts, documents et mots
- Analyses spatio-temporelles d'événements détectés sur Twitter (en cours)

Publications :

- Gia Hung Nguyen, Laure Soulier, Lynda Tamine, Nathalie Souf. DSRIM: A Deep Neural Information Retrieval Model Enhanced by a Knowledge Resource Driven Representation of Documents (regular paper).
Dans : International Conference on the Theory of Information Retrieval (ICTIR 2017), Amsterdam, 01/10/17-04/10/17, ACM, (en ligne), 2017
- Gia Hung Nguyen, Lynda Tamine, Laure Soulier, Nathalie Souf.
Learning Concept-Driven Document Embeddings for Medical Information Search (regular paper).
Dans : Conference on Artificial Intelligence in Medicine (AIME 2017), Vienna, Austria, 21/06/17-24/06/17, Springer, p. 160-170, juin 2017.

Equipe IRIS, Thèse de Paul Mousset dans le cadre du projet CIFRE IRIT-ATOS (2016-2019) : Valorisation de media par annotation événementielle générées à partir de flux d'informations hétérogènes (Lynda Tamine Lechani, lynda.lechani@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

L'objectif des travaux de thèse de Paul Mousset est de développer des modèles de résumés spatio-temporels d'événements incluant : des techniques d'annotation spatiale de flux d'informations dans les réseaux sociaux; des techniques d'agrégation spatio-temporelle de l'information

Dans ce cadre, la plateforme OSIRIM est exploitée pour : 1) stocker des collections de tweets et de POI dans une base NoSQL, 2) utiliser les capacités de calcul d'OSIRIM pour différentes analyses (évolution de mesures comme l'entropie de l'information en fonction des distances spatiales entre utilisateurs ayant émis des tweets).

Résultats obtenus grâce aux services de la plateforme OSIRIM :

- Modèles d'ordonnement end-to-end avec des appariements combinant les mots et concepts
- Techniques d'apprentissage joint des concepts, documents et mots
- Analyses spatio-temporelles d'événements détectés sur Twitter (en cours)

Projet MUSK (Nicolas Turenne, nturenne@u-pem.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

Les activités se sont poursuivies dans la continuité des projets en cours :

- pestobserver (analyse des bsv)
- youcraw (crawling de youtube)
- jdmdh (crawling de arxiv)
- twitter (crawling de twitter)

s'est rajouté un projet en IST (information scientifique et technique) au doux nom d' ISTEEX qui vise à récupérer des données de la plateforme istex et indexer le contenu par une nouvelle métrique.

Projet LISTIC (Guillaume CABANAC, guillaume.cabanac@irit.fr)

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.

Publications :

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. L'analyse qualitative des données est en cours et la soumission d'articles dans des revues du champ des sciences de l'information et de la communication est à notre programme en 2018.

Projet C2SI /Parolothèque (Julie MACLAIR, julie.mauclair@irit.fr)

Données brutes se trouvant sur la plateforme OSIRIM :

- Fichiers sons provenant de l'enregistrement de patients de l'oncopole ayant des cancers des voies ORL (CHU) (plusieurs exercices différents ont été enregistrés en fonction des besoins des partenaires)
- Fichiers logs contenant la transcription de l'affichage écran produit au patient et qui doit être lu par eux ainsi que les temps relatifs d'affichage

Méta-données se trouvant sur la plateforme OSIRIM :

- Fichiers tableurs indiquant notamment la sévérité de l'atteinte de la voix des patients et d'autres critères médicaux et personnels apportés par le personnel soignant (CHU)
- Fichiers sons convertis en 16k mono et 48k mono pour les tests perceptifs (IRIT)
- Transcription 1 provenant de tests de perception explicités plus loin (LPL)
- Segmentation des fichiers bruts pour n'effectuer les tests perceptifs que sur des mots et pas une série de mots (LPL)
- Transcription 2 contenant la transcription donnée par un système automatique de reconnaissance de la parole sur une partie des fichiers sons (IRIT)
- Transcription 3 contenant la transcription donnée par un système automatique de reconnaissance de la parole sur une partie des fichiers sons (LIA)
- Annotations de certains fichiers sons sur la sévérité de l'atteinte (LIA)
- Fichiers tableurs et ppt contenant les résultats d'analyses acoustiques (IRIT)

Activités scientifiques réalisées sur la plateforme OSIRIM :

- Tests de perceptions où des auditeurs transcrivent ce que les patients ont prononcé => Transcription 1 (LPL)
- Tests perceptifs où des auditeurs jugent de l'intentionnalité du patient (Octogone-Lordat)
- Plusieurs tests de reconnaissance automatique et analyse acoustique (IRIT, LIA)

Travaux de Thèse de Mahdi Washha (Florence SEDES, florence.sedes@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

We use the OSIRIM platform to carry out experiments to validate mahdi washha's thesis work about information quality in online social networks.

Utilisation Methods:

1-Processing large amount of Twitter information to detect and filter out spam and scam tweets and accounts :Java

2- Running a long term crawler to collect Twitter accounts to conduct our experiments on : Java

3- Conducting the experiments through using threads and parallel processing.

Résultats obtenus grâce aux services de la plateforme OSIRIM :

We have validated different hypothesis, mathematically formulated, related to spam and scan detection using Twitter data of more than three million tweets.

Publications :

- Mahdi Washha, Aziz Qaroush, Manel Mezghani, and Florence Sèdes. Information quality in social networks: A collaborative method for detecting spam tweets in trending topics. In *Advances in Artificial Intelligence: From Theory to Practice - 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2017, Arras, France, June 27-30, 2017, Proceedings, Part II*, pages 211–223, 2017. doi: 10.1007/978-3-319-60045-1_24. URL https://doi.org/10.1007/978-3-319-60045-1_24
- Mahdi Washha, Dania Shilleh, Yara Ghawadrah, Reem Jazi, and Florence Sèdes. Information Quality in Online Social Networks: A Fast Unsupervised Social Spam Detection Method for Trending Topics. In *ICEIS 2017 - Proceedings of the 19th International Conference on Enterprise Information Systems, Volume 2*, Porto, Portugal, 26-29 April, 2017, pages 663–675. SciTePress, 2017. ISBN 978-989-758-248-6
- Mahdi Washha, Aziz Qaroush, Manel Mezghani, and Florence Sèdes. Information Quality in Social Networks: Predicting Spammmy Naming Patterns for Retrieving Twitter Spam Accounts. In *ICEIS 2017 - Proceedings of the 19th International Conference on Enterprise Information Systems, Volume 2*, Porto, Portugal, 26-29 April, 2017, pages 610–622. SciTePress, 2017. ISBN 978-989-758-248-6
- MahdiWashha, Aziz Qaroush, and Florence Sèdes. Leveraging time for spammers detection on twitter. In *Proceedings of the 8th International Conference on Management of Digital EcoSystems*, pages 109–116. ACM, 2016

- Mahdi Washha, Manel Mezghani, and Florence Sèdes. La qualité de l'information dans les réseaux sociaux en ligne : une approche non supervisée et rapide de détection de spam. In Actes du XXXVème Congrès INFORSID, Toulouse, France, May 30 - June 2, 2017, pages 247–262, 2017. URL http://inforsid.fr/actes/2017/INFORSID_2017_paper_15.pdf
- Mahdi Washha, Manel Mezghani, and Florence Sèdes. Une méthode collaborative pour identifier les spams: contribution à la qualité de l'information dans les réseaux sociaux. In Conférence en Recherche d'Informations et Applications - CORIA 2017, 14th French Information Retrieval Conference. Marseille, France, March 29-31, 2017. Proceedings, Marseille, France, March 29-31, 2017., pages 139–152, 2017. doi: 10.24348/coria.2017.41. URL <https://doi.org/10.24348/coria.2017.41>
- Mahdi Washha, Aziz Qaroush, and Florence Sèdes. Impact of Time on Detecting Spammers in Twitter (regular paper). In Gestion de Données Principes, Technologies et Applications (BDA), Poitiers, 15/11/2016-18/11/2016, pages 1–10, <http://hal.inria.fr>, novembre 2016. HAL-INRIA
- Mahdi Washha and Florence Sèdes. Toward free spam social networks: detecting and tracking spammers in Twitter (short paper). In INFORSID Forum Jeunes Chercheurs, Grenoble, 31/05/2016-03/06/2016, pages 1–4, <http://praxinsa.insalyon.fr>, juin 2016. INFORSID

Travaux de Thèse de Sirinya ON-AT (Florence SEDES, florence.sedes@irit.fr)

Activités scientifiques réalisées sur la plateforme OSIRIM :

Utilisation de la plateforme OSIRIM pour effectuer les expérimentations sur la validation des travaux de thèse de Sirinya ON-AT qui portent sur la construction du profil social de l'utilisateur en prenant en compte l'évolution du réseau social de l'utilisateur.

Utilisation technique :

- Acquisition des données depuis Twitter pour utiliser comme échantillons de test dans les travaux via l'API officielle de Twitter : Java
- Calcul parallèle pour construire le profil social avec différents paramètres (utilisation du Threads) : Java + SQLite
- Calcul et analyse des résultats (calcul de similarités) : R

Résultats obtenus grâce aux services de la plateforme OSIRIM :

Les profils utilisateur construits avec le temps de calcul beaucoup plus court par rapport à l'utilisation de l'ordinateur personnel

La validation de résultats qui permet de valider l'hypothèse des travaux de thèse

Publications :

- Mahdi Washha, Aziz Qaroush, Manel Mezghani, and Florence Sèdes. Information quality in social networks: A collaborative method for detecting spam tweets in trending topics. In *Advances in Artificial Intelligence: From Theory to Practice - 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2017, Arras, France, June 27-30, 2017, Proceedings, Part II*, pages 211–223, 2017. doi: 10.1007/978-3-319-60045-1_24. URL https://doi.org/10.1007/978-3-319-60045-1_24
- Sirinya On-At, Arnaud Quirin, André Péninou, Nadine Jessel, Marie-Françoise Canut, Florence Sèdes. Taking into Account the Evolution of Users Social Profile: Experiments on Twitter and some Learned Lessons (regular paper). Dans : *IEEE International Conference on Research Challenges in Information Science (IEEE RCIS 2016)*, Grenoble, 01/06/2016-03/06/2016, Sergio Espanéa, Jolita Ralyté, Carine Souveyet (Eds.), IEEE, p. 87-98, juin 2016.
- Sirinya On-At, Arnaud Quirin, André Péninou, Nadine Jessel, Marie-Françoise Canut, Florence Sèdes. A parametric study to construct time-aware social profiles. Dans : *Trends in Social Network Analysis: Information Propagation, User Behavior Modelling, Forecasting, and Vulnerability Assessment*. Rokia Missaou, Talel Abdesslem, Matthieu Latapy (Eds.), Springer, p. 21-50, 2017.
- Sirinya On-At, André Péninou, Marie-Françoise Canut, Florence Sèdes. Toward a combinatorial analysis and parametric study to build time-aware social profiles (regular paper). Dans :

International Conference on Management of Emergent Digital EcoSystems (MEDES 2017),
Bangkok - Thailand, 07/11/2017-10/11/2017, 2017 (à paraître).