AfIA

Association française
pour l'Intelligence Artificielle

TALN-RECITAL

Conférence sur le Traitement Automatique des Langues Naturelles

PFIA 2019

Table des matières

Éditorial

La 26e édition de la conférence TALN et la 21e édition de la session jeunes chercheuses et chercheurs RECITAL se déroulent cette année à Toulouse au sein de la Plateforme française d'intelligence artificielle (PFIA). TALN a une longue tradition de tenue conjointe avec des conférences de domaines proches. Cette pratique a été initiée avec les Journées d'étude sur la parole (JEP) en 2002 à Nancy puis depuis 2008 tous les quatre ans (2008 : Avignon, 2012 : Grenoble, 2016 : Paris). Elle s'est diversifiée avec la Conférence de recherche d'information et applications (CORIA) en 2018 à Rennes. Elle innove cette année avec un hébergement à Toulouse au sein de PFIA. Ces événements sont l'occasion de rencontres enrichissantes pour tous. Cette année, ce ne sont pas moins de huit conférences, sans compter les ateliers associés, aux sessions desquelles les participants à TALN-RECITAL pourront se mêler : APIA (5e Conférence sur les Applications Pratiques de l'Intelligence Artificielle), CAp (21e Conférence sur l'Apprentissage Automatique), IC (30es Journées Francophones Ingénierie des Connaissances), JFPDA (14es Journées Planification, Décision et Apprentissage), JFSMA (27es Journées Francophones sur les Systèmes Multi-Agents), JIAF (13es Journées d'Intelligence Artificielle Fondamentale), RJCIA (17e Rencontre des Jeunes Chercheurs en Intelligence Artificielle), ainsi que CNIA (22e Conférence Nationale en Intelligence Artificielle), qui regroupe les thématiques de l'intelligence artificielle non couvertes par les conférences précédentes.

Les conférences invitées plénières, les sessions de présentations affichées et de démonstrations, les déjeuners et pauses café, les dîners de la conférence sont autant de moments programmés pour que se retrouvent les participants de toutes les conférences. Nous tenons à saluer la qualité de la planification et du suivi du comité scientifique de la plateforme ainsi que le grand travail du comité d'organisation, le tout visant à assurer que l'ensemble des conférences se tienne dans les meilleures conditions et au meilleur coût.

Pour la deuxième année consécutive, les modalités de soumission à TALN se faisaient avec un appel unique et un seul format de soumission en article court pouvant être étendu en article long sur proposition du comité de programme (et demande préalable des auteurs). Nous avons ainsi reçu soixante cinq articles courts et le comité de programme a proposé à dix articles le passage en format long (15 %) et a retenu trente et un articles en format court (48 %). Chaque article a été relu par trois membres du comité de lecture en s'appuyant le cas échant sur des relecteurs additionnels. Le comité de programme s'est appuyé sur ces relectures pour sélectionner lors d'une réunion pleinière les articles composant le programme. C'est un fonctionnement auquel nous sommes profondément attachés pour assurer une diversité dans les thématiques abordées. L'ensemble des évaluations ont été réalisées en double aveugle. Nous remercions les membres des comités de programme et de lecture (à parité femme – homme) pour leur contribution indispensable à ce processus. Le programme de la conférence est complété par quatre démonstrations sélectionnées par le comité de programme. Les titres des sessions donnent une idée des thématiques abordées par la conférence. Ils comprennent des paliers et tâches habituels du TAL (Morphologie et Syntaxe, Syntaxe, Résolution d'anaphores, Multilinguisme), reflètent la place prise par l'apprentissage (Apprentissage par transfert et modèles de langue, Plongements de mots), l'importance fondamentale que continuent à jouer les corpus et bases de données lexicales (Ressources), et l'intérêt du TAL pour des domaines particuliers (Langues spécialisées, Traitement de la langue biomédicale). Comme chaque année, l'ATALA a décerné un prix de thèse dont la récipiendaire présentera son travail en session plénière. La conférence a invité la présentation d'instruments récents du CNRS par leurs coordinatrices : d'une part le pré-GDR TAL (INS2I / informatique), qui adopte une vision inclusive du traitement de la langue (écrite, orale, signée), couvrant les communautés du traitement automatique des langues, du traitement automatique du langage parlé et de la recherche d'information ; d'autre part le GDR LIFT (INSHS / sciences du langage) sur la linguistique informatique, formelle et de terrain.

Cette année, dix-sept articles ont été soumis à RECITAL. Après avoir été chacun évalué par deux membres du comité de programme, quatre articles ont été retenus pour une présentation orale (soit un taux de sélection pour présentation orale de 24 %), et sept autres ont été retenus pour une présentation sous forme de poster (taux de sélection global de 65 %). Nous avons ainsi pu donner l'opportunité à douze jeunes chercheuses et chercheurs, en grande majorité en début de thèse, de présenter leurs travaux à la communauté. Nous remercions le comité de programme (également à parité femme – homme) pour leur minutieux travail de relecture.

Nous souhaitons pour finir au public de ces conférences une semaine riche en découvertes scientifiques et en rencontres de nouveaux collègues, dans une ambiance assurément chaude pour toute la semaine.

Emmanuel Morin, Sophie Rosset et Pierre Zweigenbaum (TALN)
Anne-Laure Ligozat et Sahar Ghannay (RECITAL)

Comités

Présidents de TALN

— Emmanuel Morin (LS2N, Université de Nantes)
— Sophie Rosset (LIMSI, CNRS, Université Paris-Saclay)
— Pierre Zweigenbaum (LIMSI, CNRS, Université Paris-Saclay)

Membres du CP de TALN

— Delphine Bernard (LiLPa, Université de Strasbourg)
— Chloé Braud (LORIA, CNRS)
— Nathalie Camelin (LIUM, Le Mans Université)
— Peggy Cellier (IRISA, INSA Rennes)
— Benoît Crabbé (LLF, Université Paris Diderot)
— Iris Eshkol-Taravella (MoDyCo, Université Paris Nanterre)
— Cécile Fabre (CLLE-ERSS, Université Toulouse - Jean Jaurès)
— Núria Gala (LPL, Aix Marseille Université)
— Thierry Hamon (LIMSI, Université Paris Nord)
— Philippe Langlais (RALI/DIRO, Université de Montréal)
— Gwénolé Lecorvé (IRISA, Université de Rennes 1)
— Aurélie Névéol (LIMSI, CNRS, Université Paris-Saclay)
— Damien Nouvel (ERTIM, INaLCO)
— Didier Schwab (LIG, Université Grenoble Alpes)
— Xavier Tannier (LIMICS, Université Pierre et Marie Curie)

Comité de lecture de TALN

— Gilles Adda (LIMSI, CNRS, Université Paris-Saclay)
— Salah Ait-Mokhtar (Naver Labs Europe)
— Alexandre Allauzen (LIMSI, CNRS, Université Paris-Saclay)
— Maxime Amblard (LORIA, Université de Lorraine)
— Jean-Yves Antoine (LIFAT, Université de Tours)
— Loïc Barrault (LIUM, Le Mans Université)
— Denis Béchet (LS2N, Université de Nantes)
— Frederic Béchet (LIS, Aix-Marseille Université)
— Patrice Bellot (LIS, Aix-Marseille Université)
— Asma Ben Abacha (Lister Hill Center, National Library of Medicine)
— Laurent Besacier (LIG, Université Grenoble Alpes)
— Yves Bestgen (ILC, Université catholique de Louvain)
— Philippe Blache (LPL, CNRS, Aix-Marseille Université)
— Fethi Bougares (LIUM, Le Mans Université)
— Thierry Charnois (LIPN, Université Paris 13)
— Vincent Claveau (IRISA, CNRS)
— Chloé Clavel (LTCI, Télécom ParisTech)
— Kevin Bretonnel Cohen (University of Colorado School of Medicine)
— Béatrice Daille (LS2N, Université de Nantes)
— Géraldine Damnati (Orange Labs)
— Gaël Dias (GREYC, Normandie Université)
— Marco Dinarelli (LIG, CNRS)
— Patrick Drouin (OLST, Université de Montréal)
— Dominique Estival (MARCS, Western Sydney University)
— Yannick Estève (LIUM, Le Mans Université)
— Olivier Ferret (CEA LIST)
— Karën Fort (STIH, Sorbonne Université)
— Thomas Francois (CENTAL, Université catholique de Louvain)
— Éric Gaussier (LIG, Université Grenoble Alpes)
— Jérôme Goulian (LIG, Université Grenoble Alpes)

— Natalia Grabar (STL, CNRS)
— Cyril Grouin (LIMSI, CNRS, Université Paris-Saclay)
— Olivier Hamon (Syllabs)
— Nabil Hathout (CLLE-ERSS, CNRS)
— Amir Hazem (LS2N, Université de Nantes)
— Nicolas Hernandez (LS2N, Université de Nantes)
— Stéphane Huet (LIA, Université d'Avignon et des Pays de Vaucluse)
— Christine Jacquin (LS2N, Université de Nantes)
— Sylvain Kahane (Modyco, Université Paris Nanterre)
— Olivier Kraif (LIDILEM, Université Grenoble Alpes)
— Mathieu Lafourcade (LIRMM, Université de Montpellier)
— David Langlois (LORIA, Université de Lorraine)
— Eric Laporte (LIGM, Université Paris-Est Marne-la-Vallée)
— Thomas Lavergne (LIMSI, Université Paris Sud, Université Paris-Saclay)
— Joseph Le Roux (LIPN, Université Paris 13)
— Benjamin Lecouteux (LIG, Université Grenoble Alpes)
— Yves Lepage (Waseda University)
— Denis Maurel (LIFAT, Université de Tours)
— Richard Moot (LIRMM, CNRS)
— Véronique Moriceau (IRIT, Université Paul Sabatier)
— Philippe Muller (IRIT, Université Paul Sabatier)
— Alexis Nasr (LIS, Aix Marseille Université)
— Adeline Nazarenko (LIPN, Université Paris 13)
— Luka Nerima (Université de Genève)
— Jian-Yun Nie (RALI/DIRO, Université de Montréal)
— Yannick Parmentier (LORIA, Université de Lorraine)
— Sebastian Peña Saldarriaga (Dictanova)
— Thierry Poibeau (Lattice, CNRS)
— Alain Polguère (ATILF, Université de Lorraine)
— Jean-Philippe Prost (LIRMM, Université de Montpellier)
— Solen Quiniou (LS2N, Université de Nantes)
— Christian Raymond (IRISA, INSA Rennes)
— Christian Retoré (LIRMM, Université de Montpellier)
— Djamé Seddah (ALMAnaCH, Paris Sorbonne Université)
— Gilles Serasset (LIG, Université Grenoble Alpes)
— Michel Simard (NRC, Canada)
— Kamel Smaili (LORIA, Université de Lorraine)
— Pascale Sébillot (IRISA, INSA Rennes)
— Ludovic Tanguy (CLLE-ERSS, Université Toulouse - Jean Jaurès)
— Juan-Manuel Torres-Moreno (LIA, Université d'Avignon et des Pays de Vaucluse)
— Guillaume Wisniewski (LIMSI, Université Paris-Sud, Université Paris-Saclay)
— François Yvon (LIMSI, CNRS, Université Paris-Saclay)

Relecteurs additionnels de TALN

— Jingshu Liu (Dictanova)
— Emile Chapuis (LTCI, Télécom ParisTech)
— Caroline Langlet (LTCI, Paris Sorbonne Université)
— Joseph Lark (Dictanova)
— Alexandre Garcia (LTCI, Télécom ParisTech)

Présidentes de RECITAL

— Anne-Laure Ligozat (LIMSI, CNRS, Université Paris-Saclay)
— Sahar Ghannay (LIMSI, CNRS, Université Paris-Saclay)

Membres du CP de RECITAL

— Jean-Yves Antoine (LIFAT, Université de Tours)

— Ismail Badache (ESPE / LIS, Aix-Marseille Université)
— Amira Barhoumi (LIUM, Université du Maine - MIRACL Sfax)
— Rachel Bawden (University of Edinburgh)
— Aurélien Bossard (LIASD, Université Paris 8)
— Chloé Braud (LORIA, CNRS)
— Nathalie Camelin (LIUM, Université du Maine)
— Rémi Cardon (STL, Lille)
— Peggy Cellier (IRISA, INSA Rennes)
— Antoine Doucet (L3i, Université de la Rochelle)
— Maha Elbayad, LIG/ Inria
— Arnaud Ferré (LIMSI-CNRS/MaIAGE-INRA, Université Paris-Saclay)
— Amel Fraisse (Gériico, Lille)
— Thomas François (CENTAL, Université catholique de Louvain)
— Nicolas Hernandez (LS2N, Université de Nantes)
— Yann Mathet (Greyc, Université de Caen)
— Alice Millour (STIH, Université Paris-Sorbonne)
— Anne-Lyse Minard (LLL, Orléans)
— Jose Moreno (IRIT, UPS)
— Tsanta Randriatsitohaina (LIMSI, Université Paris-Sud, Université Paris-Saclay)
— Loïc Vial (LIG, Université Grenoble Alpes)

Volume III : RECITAL