

PréGDR TAL TALN-RECITAL 3 juillet 2019

<https://pregdr-tal.ls2n.fr>



Périmètre

Le GDR TAL s'intéresse à la **langue sous toutes ses formes : écrite, orale, signée.**

Il aborde les thématiques de la modélisation informatique de la langue, ses manifestations et ses applications.

Communautés

Les communautés centrales :

- traitement automatique des langues
- traitement automatique du langage parlé
- recherche d'informations

Les communautés secondaires :

- le traitement automatique du document
- la représentation des connaissances et le web sémantique
- l'interaction humain-machine
- les neurosciences et les sciences cognitives

Thématiques

1. Modèles computationnels de la langue écrite, de la langue orale et de la langue des signes
2. Ressources linguistiques (données et outils)
3. Apprentissage et modélisation statistique pour le TAL
4. Multilinguisme et multiplicité des langues
5. Intermodalité et multimodalité
6. Sémantique et compréhension
7. Accès à l'information et fouille de textes
8. Productions langagières

Missions du GDR TAL

1. Animation
2. Structuration des activités de recherche, et de coordination des équipes des unités CNRS
3. Veille scientifique et prospective

Membres

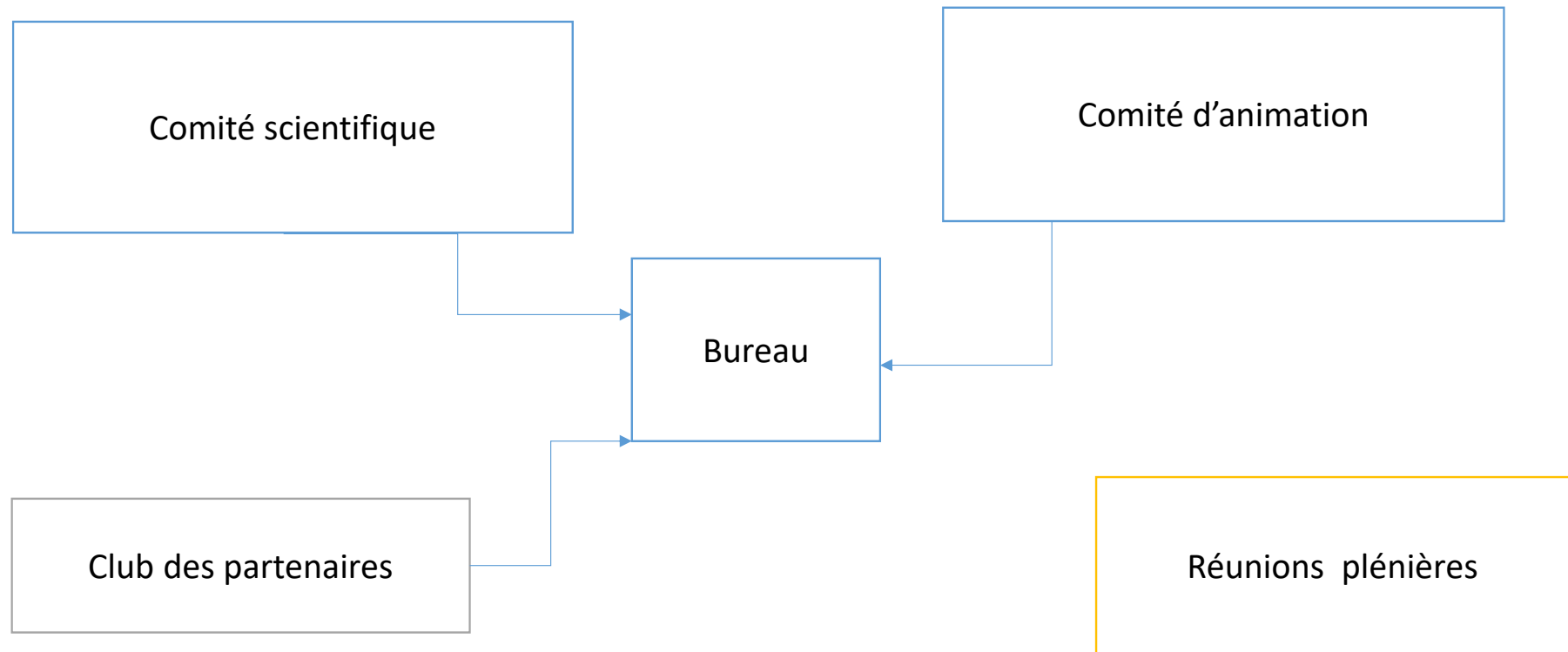
Demandes de rattachement officiel

16 UMR dont 3 INSHS - 28 équipes

+ EA : LIA, ERTIM, LI

Equipes INS2I	E/C 217	CNRS 67	TOTAL 284
---------------	---------	---------	-----------

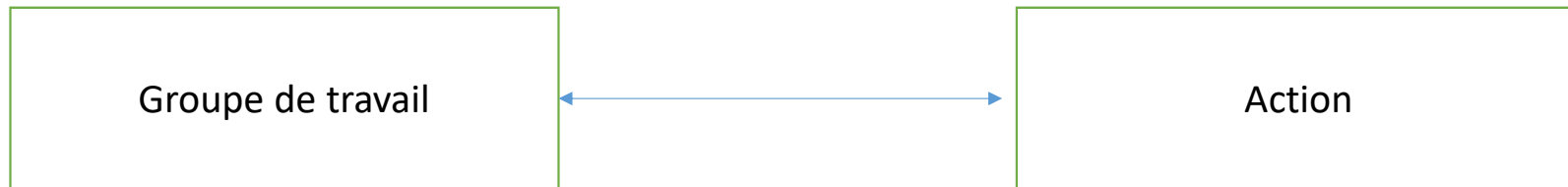
Gouvernance



Comité scientifique

- Béatrice Daille, Pr, LS2N (présidente)
- Nathalie Aussenac, DR, IRIT
- Patrice Bellot, Pr, LIS
- Laurent Besacier, Pr, LIG
- Mohand Boughanem, Pr, IRIT
- Claire Gardent, DR, LORIA
- Guillaume Gravier, DR IRISA
- Francois Yvon, DR, LIMSI
- Pierre Zweigenbaum, DR, LIMSI

Structuration



Enjeux

Débats sur les modèles de langues

Explicabilité

représentations symboliques, voire logique

Prédictabilité

représentations autoconstruites à partir de réseaux neuronaux ou autres méthodes non supervisées

Enjeux

Généralisation

Comment garantir la généralisation des approches TAL à différentes modalités, langues de domaines de spécialités ?

Robustesse

Approches peu ou non supervisées
Approches Zéro ressources

Enjeux

Responsabilité

Responsable face aux données
Transparence des données produites

Auditabilité

Contrôler la qualité des sorties des systèmes
Détecter les biais dans les données, les incertitudes

Groupes de travail

- GT1 Apprentissage et modèles pour le TAL
N. Asher, F. Bechet, A. Allauzen
- GT2 Intermodalité, mulimodalité et production langagière
P. Sébillot, B. Fabre, D. Lolive
- GT3 Multilinguisme, Multiplicité des langues
L. Besacier, F. Yvon
- GT4 Accès à l'information et fouille de textes
T. Charnois, J.-P. Chevallet

Action

A1 Ressources

G. Ada, P. Muller

Programmation

- Journée inaugurale en mars 2019
- Ecole d'été
 - Ecole jeunes chercheurs ERIA 2019
 - Ecole thématique CNRS ETAL 2020 en partenariat avec GDR Madics
- Journée des GT2, GT3 et GT4 en octobre 2019
- Journée Grand Public 2020
 - Compréhension et systèmes de TAL

Journée GDR TAL 24 octobre 2019

IMAG LIG Président comité organisation L. Besacier.

9h15-10h30 Axe multimodalité

- Présentation du travail de l'axe (30mn, responsables d'axe : D. Lolive, Pascale Sébilot, Benoit Fabre)
- Conférence Invitée (45mn) : en cours

10h45-12h00 Axe fouille & RI

- Présentation du travail de l'axe (30mn, responsables d'axe : J.P. Chevallet et Thierry Charnois)
- Conférence Invitée (45mn) : Benjamin Piwowarski (Lip6)

13h30-14h45 Axe multilinguisme

- Présentation du travail de l'axe (30mn, responsables d'axe F. Yvon et L. Besacier)
- Conférence Invitée (45mn): Matthias Gallé (Naver Labs)

14h45-16h00 Travail en 3 sous-groupes (par axes)

16h-16h30 Restitution pour chaque axe (10mn chacun)

ETAL - Nouvelles interactions pour l'accès à l'information

ETAL : école de formation du GDR TAL – tous les 2 ans

Ecole thématique INS2I CNRS en collaboration avec GDR MADICS

6 juillet au 10 juillet 2020 – ENSSAT Lannion

Nombre de participants : 30 Niveau Master informatique

Focus : apprentissage statistique et approches neuronales pour l'écrit, l'oral et la RI

Plusieurs types de sessions : cours théoriques, travaux pratiques, travaux en groupe

Intervenants : X. Tannier, Y. Estève, A. Allauzen, L. Soullier, S. Rosset, K. Fort, L.M. Rojas Barahona