

IRIT - UMR 5505

Institut de Recherche en Informatique de Toulouse  
CNRS - INP - UT3 - UT1 - UT2JUNIVERSITÉ  
TOULOUSE III  
PAUL SABATIERUNIVERSITÉ  
TOULOUSE I  
CAPITOLE

Aussenac-Gilles Nathalie

**Statut**

Permanent - DR au CNRS

**Equipe**Méthodes et ingénierie des Langues, des Ontologies et du Discours**Contact**[aussenac@irit.fr](mailto:aussenac@irit.fr)**Localisation**

IRIT2 / Niveau 2, Pièce: 275

**Téléphone**

+ 33 (0)5 61 55 82 93

**Publications****Recherche****Recherche****Encadrement****Collaborations****Logiciels****Projets****Enseignement****Animation  
scientifique****Raccourcis****Responsabilités au sein de l'IRIT**

- Responsable de l'équipe MELODI
- Membre du Conseil scientifique de l'IRIT
- Coanimatrice de l'axe "Masse de données et calcul"

**Domaine de recherche**

- Ingénierie des connaissances
- Domaines connexes :
  - Intelligence artificielle (représentation des connaissances)
  - Traitement automatique des langues
  - Gestion documentaire
  - Terminologie

**Thèmes de recherche**

- Démarches ascendantes pour la modélisation conceptuelle et l'ingénierie des connaissances, à partir de l'activité d'experts ou de textes
- Approches interdisciplinaires pour la modélisation de connaissances
  - Apport de la linguistique et du traitement automatique des langues à la modélisation à partir de textes
  - Approches ergonomiques d'analyse de la tâche, individuelle et collective ; modélisation des tâches et méthodes de résolution de problème
- Représentation des connaissances dans les modèles conceptuels.
- Ontologies et web sémantique
  - Méthodes et outils pour la construction de ressources ontologiques et terminologiques à partir de textes
  - Ontologies pour la gestion documentaire, la recherche d'informations et la gestion des connaissances
  - Ontologies pour le web sémantique
  - Annotation sémantique de documents à l'aide d'ontologies ou de bases de connaissances
- Extraction de connaissances à partir de textes
  - Extraction de relations sémantiques : approches par patrons ; combinaison d'approches linguistiques et d'apprentissage automatique
  - Peuplement de bases de connaissances à partir de textes
- Ontologies pour le repérage d'évolution de connaissances et de vocabulaires spécialisés

**Recherches actuelles**

- Plateforme DAFOE pour la construction de ressources terminologiques à partir de textes
- Extraction de relations sémantiques à partir de textes : généralisation de l'approche par patrons lexico-syntaxiques

- patrons pour le repérage de concepts et d'instances de concepts,
  - capitalisation de patrons spécialisés pour les relations spatiales et temporelles
  - extraction de relations non binaires,
  - extraction de règles,
  - patrons pour l'extraction d'information et l'annotation sémantique
- Construction automatique d'ontologies à partir de documents semi-structurés
  - Modèles de ressources terminologiques et ontologiques, représentation d'ontologies à composante lexicale
  - Processus d'évolution dans les ontologies pour l'annotation sémantique
  - Problématique d'évolution au sein des ontologies à composante lexicale

## Projet de recherche

### Modélisation des connaissances par et pour l'accès aux contenus textuels

- Document de présentation du projet
- Exposé de présentation (format .pdf)

## Parcours de recherche

### Parcours professionnel

#### Bilan complet de mon activité de recherche jusqu'à décembre 2015

- Depuis octobre 2010 : **Directeur de recherche au CNRS**, section 07 (Informatique) à l'Institut de Recherches en Informatique de Toulouse (IRIT), UMR 5505 du CNRS
- 1991 - 2010 : Chargée de recherche CR1 au CNRS, section 07 (Informatique) à l'Institut de Recherches en Informatique de Toulouse (IRIT), UMR 5505 du CNRS (1e classe à partir de 1995)
- 1989 - 1991 : Contrat post-doctoral cofinancé par le CNRS et l'entreprise Matra-Marconi Space, laboratoire mixte ARAMIIHS (Toulouse) UMR 155 du CNRS. Méthodes et outils d'analyse de tâche et d'acquisition de connaissances expertes.
- 1986 - 1989 : Docteur-Ingénieur titulaire d'une bourse du CNRS, laboratoire LSI de l'UPS, Toulouse.

### Diplômes

- **Habilitation à diriger des recherches**, Université Toulouse 3, Déc. 2005, « Méthodes ascendantes pour l'ingénierie des connaissances ». Téléchargez le manuscrit sur [Telnet](#). Jury :
  - Catherine Garbay, DR au CNRS, IMAG (Grenoble), rapporteur
  - Joost Breuker, PR, Université d'Amsterdam, SWL, rapporteur
  - Gilles Kassel, PR, Université d'Amiens, LaRIA (Amiens), rapporteur
  - Claude Chrisment, PR, Université Paul Sabatier, IRIT (Toulouse), président
  - Jean-Luc Soubie, IR, Habilité, INRIA, IRIT, directeur de recherche
  - Pierre Tchounikine, PR, Université du Mans, LIUM, examinateur
- **Doctorat en Informatique** Université de Toulouse 3. Oct. 1989, « Conception d'une méthodologie et d'un outil d'acquisition des connaissances expertes ».
- Diplôme d'Études Approfondies en Informatique, Université de Toulouse 3. 1986. Option algorithmique numérique.
- Ingénieur en Informatique de l'ENSEEIH, INPT Toulouse. 1986.

### Membre d'associations professionnelles

- Association de Recherche en Sciences Cognitives (ARCog) depuis 1990
- Association Française d'Intelligence Artificielle (AFIA) depuis 1991
- Association pour le Traitement Automatique des Langues (ATALA) depuis 1999

### Encadrement

#### Encadrement de thèses en cours

- Jordane DORNE : Représentation sémantique de données d'observation de la Terre pour l'analyse de changements. Depuis mai 2018, Financement CIFRE avec THALES ALIENA SPACE (TAS) (20% avec Cassia Trojahn et Catherine Comparot), IRIT-MELODI et Romain Hugues, TAS, Toulouse). Université Toulouse 2, école doctorale MITT (Mathématiques, Informatique et Télécommunication de Toulouse).

## Encadrement de thèses terminées

- Adel Ghamnia : Extraction de relations à partir de textes. Sept. 2015 à Sept 2018 (non soutenue), bourse PRES-région, directrice de thèse et co-encadrement (30% avec Mouna Kamel, Cassia Trojhan, IRIT-MELODI et Cécile Fabre, CLLE-ERSS, Toulouse). Université Toulouse 2, école doctorale MITT (Mathématiques, Informatique et Télécommunication de Toulouse).
- Jihen Karoui : Analyse de l'humour dans les textes courts en français et en arabe. Depuis sept.2013, cotutelle avec la Tunisie, financement bourse tunisienne, directrice de thèse et co-encadrement 30 % avec Farah Benamara, IRIT-MELODI). Université Toulouse 3, école doctorale MITT (Mathématiques, Informatique et Télécommunication de Toulouse) et université de Sfax (Tunisie).
- Jean-Philippe Fauconnier : Acquisition de liens sémantiques à partir d'éléments de mise en forme des textes : exploitation des structures énumératives. 2012 - 2016, bourse MNRT, directrice de thèse et encadrement (30% avec M. Kamel, IRIT-MELODI). Soutenue le 25 janvier 2016. Université Toulouse 3, école doctorale MITT (Mathématiques, Informatique et Télécommunication de Toulouse).
- Laurent Sorin : Contribution des architectures textuelles à l'accessibilité non-visuelle des documents numériques. 2012 - 2015. Bourse PRES-région, directrice de thèse et encadrement (30% avec M. Mojahid, IRIT-ELIPSE et Julie Lemarié, CLLE-LTC, Toulouse). Soutenue le 15 déc. 2015. Université Toulouse 2, école doctorale MITT.
- Julien Corman : Knowledge base ontological debugging, guided by linguistic evidence. 2012 - 2015. Bourse MNRT, co-encadrement à 30% avec Laure Vieu, directrice de thèse. Soutenue le 14 déc. 2015, université Toulouse 3, école doctorale MITT.
- Roger Leitzke Granada : évaluation de méthodes pour l'extraction de relations taxonomiques à partir de textes. 2014\2015, co-directrice en co-tutelle avec Renata Vieira (Pontificia Universidade Católica do Rio Grande do Sul (PUCRS)) et C. Trojahn (IRIT). Soutenue le 28 sept. 2015.
- Anis Tissaoui : « Annotation sémantique de corpus dynamiques », 2008\2013, directrice de thèse et co-encadrement avec P. Laublet, Lalic-Paris4 et N. Hernandez, IRIT-IC3. Université Toulouse 3, spécialité Informatique, école doctorale EDIT.
- Zied Sellami : « Gestion dynamique d'ontologies à partir de textes par systèmes multi-agents adaptatifs ». 2008 - 2012, co-directrice de thèse avec V. Camps (IRIT-SMAC). Université Toulouse 3, spécialité Informatique, école doctorale EDMITT, 9 Juillet 2012.
- Nacim Chikhj : « Calcul de centralité et identification de structures de communautés dans les graphes de documents », 2006 - Dec 2010, directrice de thèse et co-encadrement (avec B. Rothenburger, IRIT-IC3). Université Toulouse 3, spécialité Informatique, école doctorale EDIT, le 17 décembre 2010.
- Aurélie Picton : « Diachronie en langue de spécialité. Définition d'une méthode linguistique outillée pour repérer l'évolution des connaissances en corpus. Un exemple appliqué au domaine spatial », déc. 2004 - oct. 2009, directrice de thèse A. Condamines (CLLE-ERSS), Université Toulouse 2, spécialité Sciences du langage, 20 octobre 2009.
- Axel Reymonet : « Modélisation de connaissances à partir de textes pour une Recherche d'Information Sémantique », 2004 - 2008. Contrat Cifre avec l'entreprise ACTIA (Toulouse). Université Toulouse 3, spécialité Informatique, école doctorale MITT, 23 septembre 2008.
- Kévin Ottens : « Un système multi-agent adaptatif pour la construction d'ontologies à partir de textes », 2004 - 2007, directrice de thèse : M.-P. Gleizes (IRIT-SMAC). Université Toulouse 3, spécialité Informatique, école doctorale EDIT, 2 octobre 2007.
- Mustapha Baziz : « Apport des ontologies à la recherche d'information à l'aide du moteur Mercure », 2002 - 2005. Directeur de thèse: M. Boughanem (IRIT-SIG), Université Toulouse 3, spécialité Informatique, école doctorale EDIT, 14 décembre 2005.
- Nathalie Hernandez : « Ontologies pour l'aide à une activité de veille ou d'exploration d'un domaine », 2002 - 2005 . Directrice de thèse : J. Mothe (IRIT-SIG), Université Toulouse 3, spécialité Informatique, école doctorale EDIT, 6 décembre 2005.
- Patrick Séguéla : « Extension du logiciel REX par des outils et techniques d'analyse de texte et de traitement automatique du langage naturel. Développement du système Caméléon », 1996 - 2001. Thèse Cifre au CEA, CEN de Cadarache. Directeur de thèse : Mario Borillo, Université Toulouse 3, école doctorale Informatique, mars 2001.
- Nada Matta : « Représentation des connaissances pour l'acquisition des connaissances

», 1991 - 1995. Directeur de thèse : Mario Borillo, Université Toulouse 3, école doctorale Informatique, octobre 1995.

- David Macchion : « Intégration de plusieurs modes de raisonnement pour le diagnostic », 1991 - 1994. Directeur de thèse : J.P. Macchion, Université Toulouse 3, Ecole Doctorale Informatique, nov. 1994.



## Collaborations

### Collaborations au sein de l'IRIT

- SIG-EVI et SIG-BD, co-encadrement des thèses de M. Baziz et N. Hernandez, plates-formes de recherche d'information EXPRIM et PLEXIR, projet BQR "ontologies et textes"
- SMAC, co-encadrement de la thèse de K. Ottens et de Z. Sellami, projets Dynamo et BQR "ontologies et textes"
- LaLIC, P. Muller et Farah Bénomara
- Groupe de travail "langage naturel"

### Collaborations avec des entreprises

- CNES (Toulouse), Agence Spatiale, contrats Recherche et Technologie sur la gestion des connaissances
- ACTIA (Toulouse), PME diagnostic automobile, Dynamo et laboratoire Autodiag
- ARTAL (Toulouse), SSI informatique, Dynamo
- Upetech (Toulouse), systèmes multi-agents, Dynamo
- Editions Epistèmes (Paris), édition scientifique, Corpus logicistes et Arkeotek
- Mondeca (Paris), gestion documentaire, DAFOE4App et construction d'ontologies
- Synapse-Développement (Toulouse), PME traitement automatique du langage

### Collaborations nationales

- Méthodes de modélisation : LRI, IRIN, INRIA-Acacia, TechCico, DER-EDF, LaRIA
- Construction d'ontologie à partir de textes : CLLE-ERSS, LIPN, INSERM-S\_852, INRIA-Edelweis
- Extraction de relations sémantiques : CLLE-ERSS, LIPN, LIUPPA, COGIT
- Annotation sémantique : « Préhistoire et Technologie », LaLIC, Mondeca

### Collaborations internationales

- LOA (Trento), laboratoire ILIKS
- Departamento de Ingeniería de la Información y las Comunicaciones, Facultad de Informática, Universidad de Murcia (Espagne), enseignements et collaboration avec J. Fernandez Breis, sur "ontologies et fouille de textes sur le web"
- DFKI : invitation de P. Cimiano, P. Buitelaar (DFKI) et T. Declerck
- Université de Laval à Québec, Nicole Tourigny, préparation de cours communs en ingénierie des connaissances, deux séjours de 15 jours en 2007.



## Logiciels

- Caméléon : aide au repérage de relations sémantiques par projection de patrons lexico-syntaxiques sur des textes
- Terminae (développé au LIPN par S. Szulman et B. Biébow) : plate-forme de modélisation de ressources terminologiques ou ontologiques à partir de textes
- Géditerm : éditeur de bases de connaissances terminologiques
- Macao : méthode et logiciel de modélisation de connaissances expertes



## Projets

## Projets en cours

### Coopérations industrielles

#### Projet CRE-Orange-Labs- 2018 : extraction de relation n-aires à l'aide de Frames

Exploitation de frames pour identifier des relations n-aires à partir de textes et enrichie une base de connaissances en RDF. Intégration des résultats dans la plate-forme TILT de Orange-Labs. Porteur : Farah Benamara.

#### Collaboration avec Thalès Alenia Space (2017-2020) Encadrement d'une bourse CIFRE

Représentation sémantique de grands volumes de données de géo-information : Application à l'analyse de changements paysagers et environnementaux. Thèse CIFRE de Jordane Dorne.

### Projets régionaux

#### Projet AVI-REX (2018-2020) construction d'une ontologie de diagnostic technique et de retour d'expérience

avec SimSoft3D, PME spécialisée dans les assistants virtuels vocaux : Méthode de construction d'une ontologie de diagnostic technique et de retour d'expérience permettant de construire une base de connaissances pour un assistant virtuel. Projet financé par la région Occitanie. Encadrement d'une post-doctorante (Amina Annane) et porteur du projet pour l'IRIT.

#### SMEPEDIA : Projet COMUE-égion Midi-Pyrénées (2015-2018) Intégration de différentes techniques linguistiques et statistiques d'extraction de relations sémantiques pour enrichir DBPedia en Français

co-Financement de thèse par la région Midi-Pyrénées et la COMUE de Toulouse. Intégration de différentes techniques linguistiques et statistiques d'analyse du langage, ainsi que de connaissances du web des données, pour l'extraction de relations sémantiques. Partenaires : IRIT-MELODI (M. Kamel porteur) et laboratoire CLLE-ERSS (C. Fabre). Co-direction de la thèse d'A. Ghamnia.

#### DataNoos (2018-2020) Projet du RTRA STAE : Vers une alliance académique pour le partage de données et la production de connaissances.

Porteur du projet, 10 laboratoires partenaires.

### Projets Européens

#### CANDELA (2018-2020) (H2020) Copernicus Access Plaform Intermiediate Layers Small Scale Demonstrator

. <http://www.candela-h2020.eu/> . Réalisation d'un module de recherche sémantique d'images d'observation de la Terre enrichies par des données extraites des images et par des données ouvertes. Encadrement d'un post doctorant (Ba-Huy TRAN).

### Autres Projets

#### Projet RelExt (2016-2017) Extraction de relations dans les corpus ISTEEX

Appel d'offre ISTEEX de l'ATILF- exploitation de corpus scientifiques pour la mise au point de différents logiciels d'extraction de relations sémantiques. Porteur du projet. Partenaire : C. Fabre de CLLE-ERSS. Co-encadrement du post-doctorat d'E. Manishina.

#### Projet FUI SparkInData (2015 - 2018) - appel d'offre - Cloud Computing & Big Data

SparkIndata : Plateforme fédérant les sources de données d'observation de la Terre et favorisant l'émergence d'un écosystème riche de nouveaux services et usages (agriculture, urbanisme, sécurité, climat, santé, etc.). Porteur : ATOS, 11 partenaires. Je coordonne la participation de 4 équipes de l'IRIT (dont MELODI) et suis responsable du Sous-projet 4 - Apport des technologies sémantiques à la gestion d'images satellites.

### Projets ANR terminés



**Projet ASFALDA ♦ ANR-11-CONTINT (2012-2015) Analyse Sémantique en Frames: Annotation, Lexique, Discours et Automatisation**

Porteur : Marie Candito (ALPAGE). Participation au lot 1 (t♦che 1.3) sur l'acquisition de relations entre frames ou entre rôles.

**Projet MOANO ♦ ANR-10-CORD-024-03 (2011♦début 2015) Modèles et Outils pour Applications NOMades de découverte de territoire**

Porteur : Philippe Roose (LIUPPA). Partenaires : LIUPPA, LIG, LIFL, IRIT. Participation au WP1 sur la recherche sémantique d'information. Co-encadrement de Davide Buscaldi en post-doctorat de Nov. 2011 à août 2012, de 2 stages L3 et M1, et d'un master en 2015.

**Corpus Logicistes - ANR-07-CORP-006 (2008-2010) : Corpus logicistes en archéologie des techniques**

Site : <http://www.arkeotek.org/>

Responsable : V. Roux (« préhistoire et technologies », MAE Nanterre)

Partenaires : laboratoires « préhistoire et technologies » de Paris X, Editions Epistèmes

Le projet a trois objectifs au service de la cumulativité des connaissances dans les sciences humaines : (1) développer des méthodes et outils permettant la constitution de corpus dits "logicistes", composés (...)

[Lire la suite](#) ☐

**DAFOE4App - ANR-06-TLOG-010 (2007-2009) : Differential and formal ontology editor**

Site : <http://dafoe4app.fr/>

Responsable : J. Charlet (INSERM)

Partenaires : INSERM UMR\_S 872 équipe 20 (Paris 6/Paris 5), ENST/GET (Paris), IRIT (Toulouse 3), LIPN (Paris 13), LISI (Poitiers), Mondeca (Paris), Supélec (Saclay), UTC (Compiègne)

L'objectif du projet est de proposer une méthode complète associée à une plateforme technique pour concevoir des ontologies, de la modélisation à partir du domaine à leur (...)

[Lire la suite](#) ☐

**Dynamo - ANR-07-TLOG-004 (2008-2010) : Dynamic Ontologies**

site : <https://www.irit.fr/dynamo/>

Responsable : J. Thomas (ACTIA) - Partenaires : ACTIA, ARTAL, LaLIC (Univ. Paris 4), IRIT éq. IC3, SIG-EVI et SMAC, Préhistoire et Technologie (Paris X).

Objectifs : définir une méthode et des outils pour améliorer la recherche d'information et la satisfaction des utilisateurs en prenant en compte l'évolution des collections documentaires consultées. L'objectif est de définir une recherche (...)

[Lire la suite](#) ☐

**Geonto - ANR-07-MDCO-005 (2008 - 2010) : Constitution, alignement, comparaison et exploitation d'ontologies géographiques hétérogènes**

site : <http://geonto.lri.fr/>

Responsable : C. Reynaud, LRI - Partenaires : LRI éq. IASI/gemo (Univ. Paris Sud), IRIT éq. IC3 (Toulouse 3), LIUPPA éq. DESI (Univ. de Pau) et CoGIT (IGN).

Ce projet vise l'interopérabilité de données géographiques hétérogènes via deux objectifs applicatifs : (1) l'intégration de bases de données géographiques hétérogènes, par la mise en (...)

[Lire la suite](#) ☐



## Enseignement

- Septembre 2013 : ENSEEIHT (Toulouse), 3e année Informatique, " Ontologies et Web sémantique ", 14h de cours par an.
  - Supports de cours : [Ontologies et web sémantique \(format.zip\)](#)
  - Archive de [sujets d'examens](#)
- p>
- Septembre 2008 : Universidad de Murcia (Espagne), departamento de Informática, Master 2 recherche, " De textos a ontologías : herramientas de procesamiento del lenguaje natural para el modelado del conocimiento ", professeur invitée, 10h cours et 10h encadrement.
  - [support de cours en espagnol](#)
- De 2007 à 2013: Université P. Sabatier, M2R MIT en Informatique et Télécommunication, responsable du module BDRI8 - ALRI (Analyses Linguistiques pour la Recherche d'Information), intervention sur « ontologies et web sémantique » (8h/an)
  - Supports de cours : [Ontologies et web sémantique \(format.zip\)](#)
  - Archive de [sujets d'examens](#)
- Depuis 2006 : Université Toulouse 2, M2R ECIL en Ergonomie Cognitive et Informatique linguistique, intervention sur « Ontologies et Traitement Automatique des Langues » (6h/an)
  - [support de cours \(format.zip\)](#)
- Depuis 2006 : Université P. Sabatier, M2R ICMST en Information Communication et Médiations Socio-techniques, intervention sur « Ontologies et web sémantique » (6h/an)
  - [support de cours \(format.zip\)](#)
- Depuis 2006 : URFIST, Université Toulouse1, Formation continue, « ontologies et web sémantique » (6h/an).
  - support de cours : [support1](#), [support2](#)



## Animation scientifique

### Groupes de travail

- co-fondatrice et animatrice du [GRACQ](#) de 1999 à 2005 ; modératrice de la liste info-ic depuis sa création en 1990
- thème 7 du [GDR-I3](#)
- coanimatrice du groupe "[Terminologie et IA](#) "
- AS du RTP-DOC " Terminologie et Corpus " [ASSTICCOT](#)
- Comité de pilotage des conférences IC, TIA et EKAW

### Comité éditorial de revues

- [Revue I3](#)
- [Applied Ontology](#)
- [IJHCS](#) : International Journal of Human Computer Studies

### Organisation de conférences et workshops

- [TIA 2009 : Conférence Internationale Terminologie et Intelligence Artificielle](#) à Toulouse du 18 au 20 novembre 2009
- [TALN 2007](#) à Toulouse, juin 2007
- [SdC2006 : Semaine de la Connaissance](#) à Nantes, juin 2006.
- [IC 2000 : Conférence francophone d'ingénierie des connaissances](#) à Toulouse, mai 2000.
- [EKAW 1993 : 4th European Knowledge Acquisition Workshop](#), Toulouse et Caylus, Août 1993

### Evaluation de la recherche

- Demandes ponctuelles d'évaluation en tant qu'expert :
  - Projets régionaux : Région Champagne Ardenne (2004), Région Normandie (2003), région Ile de France (2009)

- Bourses post-doctorales: Université Technologique de Troyes (2003, 2004)
- Programmes internationaux : GODARD, collaboration scientifique France-Argentine (2004)
- Programmes nationaux : ACI puis ANR Masse de données (2004, 2006, 2007, 2008) ; ANR « blanc SHS » (2007, 2008) ; ANR - Technologies Logicielles (2006, 2007)
- ANVAR - OESO (projet en 2005, bourses Cifre : 2005, 2006, 2008)
- Evaluation d'unité : unité INRA Méta-Risk, 2007 ; Unité INRA IATE, 2010
- Membre du comité d'évaluation de programmes scientifiques
  - TCAN (Traitement des Connaissances, Apprentissage et NTIC) (2003, 2004)
  - RTP-DOC (Réseau Thématique Pluridisciplinaire du département STIC du CNRS sur le « document numérique », N. 33) : bureau, comité de pilotage (2002 à 2005).
  - GDR-I3 : membre du comité directeur depuis janvier 2007
  - ANR, programmes DEFIS (2008)
  - CNRS, comité national : membre nommé de la CID 42

