

BENOÎT GAILLARD

Adresse: 29, Boulevard Matabiau, 31000 TOULOUSE
Adresse électronique: benoit.gaillard@univ-tlse2.fr
téléphone: (+33)(0)628250217

PROFIL

INGENIEUR DE RECHERCHE EN INFORMATIQUE LINGUISTIQUE Docteur en intelligence artificielle

EXPERIENCE PROFESSIONNELLE

2010: CNRS, Umr 5263 Cile-erss, Université Toulouse 2 Le Mirail: Agent Contractuel au sein du projet ANR M3 (Postdoc).

- Constitution et analyse de graphes de synonymie et de traduction à partir de ressources lexicographiques et expérimentales, en vue de comparer les structures qui en émergent.

- Application à la typologie lexicale et sémantique de la théorie des graphes petit monde, de l'analyse formelle de concepts, de techniques de marches aléatoires et de compression pour la clusterisation.

2009: ORANGE LABS, Ingénieur de Recherche en recherche d'information multilingue (CLIR).

- Mise en œuvre de méthodes basées sur le Traitement Automatique de la Langue Naturelle (TALN) pour dépasser l'état de l'art du CLIR. Utilisation de thésaurus, de réseaux de mots, lexiques bilingues, d'analyse sémantique, de méthodes statistiques...

- Documentation de l'état de l'art, analyses de corpus de requêtes.

- Avec le projet européen QUAERO, implémentation de prototypes de CLIR basés sur la traduction automatique statistique (SMT), optimisations utilisant des outils de TALN.

2008: SYSTRAN, Ingénieur de développement, au sein de l'équipe R&D.

- Mises à jour, optimisations et maintenance du code des moteurs de traduction en C++, sous Linux. Développement d'outils de gestion de ressources lexicales pour répondre aux besoins des linguistes de l'équipe. Analyse et codage syntaxiques, post-édition statistique.

- Automatisation de l'évaluation (statistique et humaine) de la qualité des traductions. Scripts (Shell, Perl) de mesures de similarité et de mise en parallèle de phrases dont la traduction varie.

2007 à 2008: ARICIE, Ingénieur d'étude auprès de la direction technique. Développement dans l'environnement WEB 2.0 de ARICIE

- **Décembre 2007 à Mars 2008: Pour CrossKnowledge:** Implémentation d'un provider d'indexation et de recherche de documents à partir du moteur de recherche open source *Lucene.net*: Flux XML, indexation paramétrée, interfaces administrateur et recherche avancée.

- **Juillet à Décembre 2007: Pour SAGE:** Migration vers la technologie *DotNetNuke (DNN)* du système d'identification et de gestions des droits des revendeurs et utilisateurs de l'extranet: l'identification par Active Directory de DNN pour l'identification intégrée multi portail (SSO), administration de comptes, migration de bases de données d'utilisateurs.

- **Mai à Juillet 2007: Pour Wildlife Conservation Society (WCS):** Développement d'une interface d'aide à la traduction de fichiers de ressources. Intégration d'outils et ressources issus du TALN (WordNet, SharpNLP...) et de services web de traduction automatique. Création de dictionnaires de termes prioritaires pour améliorer les traductions automatiques à contrôler et éditer par les traducteurs professionnels.

- **Février 2007 à Avril 2007:** Développement d'un module de présentation de connaissances typées et hiérarchisées en s'appuyant sur le Mapping Entité/Relationnel de LLBLGen.

2003 à 2006 : Tuteur associé à l'Université de Sussex (Royaume Uni): Chargé de Travaux Dirigés, enseignant les "Outils Mathématiques pour les Sciences Cognitives".

2002 à 2005 : Élaboration d'une ontologie en ligne des publications de recherche en vision cognitive artificielle, au sein du projet Européen «ECvision».

2002 : Babelmedia Ltd (Brighton, UK): Spécialiste de la localisation Française de jeux vidéo.

FORMATION/RECHERCHE

2002 - 2006 : **Doctorat au Département Informatique, Université de Sussex, Brighton, (Royaume Uni)**

«Un modèle, inspiré de la biologie, de prise de décision contrôlée par l'activité neuronale aléatoire». Conception logicielle du modèle et d'un simulateur. Nouvelle approche basée sur les statistique du deuxième ordre pour expliquer la neurobiologie du contrôle du compromis «acuité vs. rapidité» de la prise de décision. Démonstration de l'importance du caractère étendu et distribué des systèmes neuronaux.

2001 - 2002 : **Master de Sciences en Systèmes Intelligents à COGS (Centre pour les sciences cognitives et l'informatique), Université de Sussex.**

Bio-informatique : Modélisation neuronale de la vision, bibliographie, modélisation et simulation (algorithmes en Matlab).

Vie Artificielle : Algorithmes génétiques pour montrer l'avantage évolutionniste de la faculté d'apprentissage (JAVA).

Traitement automatique de la langue naturelle : Résolution contextuelle d'ambiguïtés d'homographe, par apprentissage non supervisé (LISP).

Intelligence Artificielle : Implémentation d'un système expert (PROLOG).

1999 - 2002 : **École Nationale Supérieure des Télécommunications de Bretagne (ENSTB), Brest (29) Diplôme d'ingénieur en télécommunications.**

1996 - 1999 : **Classes préparatoires aux grandes écoles, Lycée Kérichen, Brest.**

1996: **Baccalauréat S, mention bien.**

DOMAINES DE COMPETENCES

Traitement du langage naturel et linguistique

Traduction Automatique, grammaires génératives, modèles statistiques du langage et leur entraînement sur corpus (mono- ou bi-lingue), analyses de corpus, codage (syntaxique et sémantique) automatique d'entrées de dictionnaires, post-édition statistique, analyse sémantique, traitements de requêtes (traduction, expansion...). *Linguistique* : Typologie lexicale et sémantique, analyses de corpus, lexicographie, étude crosslinguistique de la synonymie.

Neurosciences Théoriques et Intelligence Artificielle:

Modèles de neurones et de la neurologie de la décision, statistiques des trains de potentiels d'action neuronaux, modélisation numérique, réseaux de neurones, algorithmes génétiques, systèmes experts, sciences cognitives.

Langages de programmation : C++, C#, C, Vb.net, Java, SQL, Matlab, Maple, Caml, Lisp, Prolog, Pop-11, XML, HTML, PHP. Script: Python, Perl; Shells:BASH, CSH.

Langues étrangères (langue maternelle: Français):

Anglais : courant après 6 ans en Angleterre, thèse rédigée en Anglais.

Allemand : scolaire (10 ans), pratique orale conséquente

Espagnol : scolaire (2 ans)

Arabe : niveau scolaire (2 ans intensifs), certificat d'Arabe littéral de la Sorbonne.

Chinois : élémentaire.

Grec Ancien: 17/20 à l'oral du baccalauréat.

SITE WEB/PUBLICATIONS/MEMOIRE DE THESE: <http://benoit.gaillard.aricie.net>