

17 rue de Châteaudun

35000 Rennes

☎ 06 73 62 80 46

✉ toma.girault@gmail.com

né le 16 novembre 1981 (29 ans)



Thomas Girault

Formation universitaire

- 2010 **Thèse de doctorat**, *IRISA, Université de Rennes 1*, Mention très honorable.
Spécialités : traitement automatique des langues, apprentissage automatique
Sujet de recherche : « apprentissage incrémental pour la construction de bases lexicales évolutives : application en désambiguïsation d'entités nommées »
Encadrement : Pascale Sébillot
- 2005 **Master 2 recherche**, *GREYC, Université de Caen*, mention AB.
Spécialités : « langage, image, document »
Sujet de recherche : « annotation et structuration automatique de corpus de dialogue »
Encadrement : Anne Nicolle
- 2004 **Maitrise d'informatique ERASMUS**, *University of Sheffield (UK)*, mention AB.
Spécialités : intelligence artificielle, traitement automatique des langues et de la parole
Sujet de recherche : « symbolic machine learning for phrase chunking and robust parsing »
Encadrement : Mark Hepple
- 2003 **Licence d'informatique**, *Université de Caen*.
Conception et réalisation d'un système d'indexation multilingue de la presse en ligne
- 2002 **DEUG MIAS**, *Université de Caen*.

Expérience professionnelle

- 2005–2008 **Ingénieur de recherche**, *France Télécom R&D*, Lannion.
Spécialités : fouille de données textuelles, recherche d'information, extraction d'information
Encadrement d'un stage de Master 2 professionnel : « Extraction de connaissances issues de Wikipédia pour la constitution automatique d'ontologies multilingues »
- 2007 et 2008 **Enseignement universitaire (vacataire)**, *UCO*, Guingamp.
Deux semestres d'enseignement (50 heures : CM+TD+TP)
Deuxième année de licence mathématiques et informatique
« Programmation orientée objet avancée » (multithreading en Java, génie logiciel avec UML)
- Été 2003 **Stage de maîtrise**, *Normanet*, Caen.
Conception et réalisation d'un module générique de cartographie (éditeur SVG en Flash)

Compétences en informatique

| | | | |
|------------------------------|--|--------------------------------|-----------------|
| Programmation orientée objet | C++, Java, Python | Parallélisme et multithreading | Java et C Posix |
| Web | Php, MySQL, PostgreSQL, XML, HTML/CSS, JavaScript, Flex/ActionScript | | |
| Autres langages | Perl, Prolog, Haskell, Scheme, Lisp, Shell, Pure Data | | |
| Divers | Programmation et visualisation scientifique (SciPy, notions en Matlab et R) Graphisme et animation 2D (Flash) et 3D (Blender) Administration et programmation système (Linux, Windows) | | |

Langues

| | |
|---------|---|
| Anglais | Bonne compréhension et communication à l'oral, expérience de l'écriture académique Notions élémentaires en allemand, japonais et indonésien Plus d'une dizaine de séjours à l'étranger dont une année scolaire au Royaume Uni |
|---------|---|

Divers

| | |
|---------|---|
| Sports | Natation, surf, escalade |
| Loisirs | Basse électrique, MAO, dessin, peinture, photographie pour l'animation « stop motion », jonglerie |
| Autres | Titulaire de l'attestation de formation aux premiers secours Permis B avec véhicule |

Publications

Olivier Collin, Benoît Gaillard, Jean-Léon Bouraoui, and Thomas Girault. Semantic resource extraction from the wikipedia category lattice. In *Workshop on Methods for the automatic acquisition of Language Resources and their evaluation methods, LREC 2010*, Valletta, Malta, may 2010.

Thomas Girault. Concept lattice mining for unsupervised named entity disambiguation. In *Concept Lattices and their Applications, CLA'08*, pages 32–43, Olomouc, Czech Republic, 2008.

Thomas Girault. Exploitation de treillis de Galois en désambiguïsation non supervisée d'entités nommées. In *15ème conférence sur le Traitement Automatique des Langues Naturelles, TALN'08*, pages 260–269, Avignon, France, 2008.

Thomas Girault. *Apprentissage incrémental pour la construction de bases lexicales évolutives : application en désambiguïsation d'entités nommées*. PhD thesis, Université de Rennes 1, France, 2010.