

# Mikel Egaña Aranguren, Ph.D.

Eurohelp Consulting S.L.  
Bilbo, Spain

Skype: mikel\_egana  
Email: mikel.egana.aranguren@gmail.com  
Homepage: <http://mikeleganaaranguren.com/>

## Education

2009 **Ph.D.** Computer Science, University of Manchester, UK  
2005 **M.Sc.** Bioinformatics, University of Manchester, UK  
2003 **B.Sc.** Biology, University of Basque Country, Spain  
2002 **Invited student** Evolutionary Ecology at Canterbury Christ Church University College, UK  
2002 **Erasmus student** Environmental Biology at Canterbury Christ Church University College, UK

## Employment

2016-Present **Analyst, Torres Quevedo fellow** Eurohelp Consulting, S.L.  
2015-2016 **Analyst** Eurohelp Consulting, S.L.  
2014-2015 **Post-doc researcher** Genomic Resources Group, UPV/EHU  
2011-2014 **Marie Curie Cofund fellow** Ontology Engineering Group (Computer Science);  
Biological Informatics Group (CBGP), UPM, Spain  
2009 **Researcher** OGO project, UM, Spain  
2006 **Marie Curie EST fellow** Computational Biology group, VIB, Belgium

## Research visits

2005 European Bioinformatics Institute (EBI), funded by the Network of Excellence on Semantic Interoperability and Data Mining in Biomedicine (EU)

## Funding obtained

2011-2014 Marie Curie Cofund (EU)  
2006 Marie Curie EST (EU)  
2005 EPSRC (UK): Ph.D. fees  
2005 - 2008 University of Manchester (UK): Ph.D. maintenance allowance  
2002 Erasmus (EU)

## Participation in funded projects

2016 - 2020 REPLICATE. Renaissance of Places with Innovative Citizenship And TEchnology (Project 691735). EU. (Consortium: 24M EUR; Eurohelp: 300.000 EUR).  
2015 - 2016 Linking Open Domains, Plataforma para la generación de datos enlazados (LODGen). Ministerio de Industria, Energía y Turismo, Spain (TSI-100105-2015-0012). (40.000 EUR)  
2015 Enlazando Gipuzkoa con el Mundo (ENGIMU). Gipuzkoako Foru Aldundia, Spain. (40.000 EUR)

## Publications

### Refereed Journal Articles

- |      |  |
|------|--|
| 2015 | Aranguren, M. E. and Wilkinson, M. D. (2015). Enhanced reproducibility of SADI Web service workflows with Galaxy and Docker. <i>GigaScience</i> , 4(59)  |
| 2015 | Pawluczyk, M., Weiss, J., Links, M. G., Aranguren, M. E., Wilkinson, M. D., and Egea-Cortines, M. (2015). Quantitative evaluation of bias in PCR amplification and Next Generation Sequencing derived from metabarcoding samples. <i>Analytical and Bioanalytical Chemistry</i> , 407(7):1841–1848 |
| 2014 | González, A. R., Callahan, A., Toledo, J. C., García, A., Aranguren, M. E., Dumontier, M., and Wilkinson, M. D. (2014a). Automatically exposing OpenLifeData via SADI semantic Web Services. <i>Journal of Biomedical Semantics</i> , 5(1):46+   |
| 2014 | Aranguren, M. E., González, A. R., and Wilkinson, M. D. (2014). Executing SADI services in Galaxy. <i>Journal of Biomedical Semantics</i> , 5(1):42+   |
| 2014 | José Antonio Miñarro Giménez, Mikel Egaña Aranguren, Boris Villazón Terrazas, and Jesualdo Tomás Fernández Breis (2014). Translational research combining orthologous genes and human diseases with the OGOLOD dataset. <i>Semantic Web Journal</i> , 5(2):145–149                                 |
| 2014 | Mikel Egaña Aranguren, Jesualdo Tomás Fernández Breis, and Michel Dumontier (2014). Special issue on Linked Data for Health Care and the Life Sciences. <i>Semantic Web Journal</i> , 5(2):99–100  |
| 2013 | Duque-Ramos, A., Fernández-Breis, J. T., Iniesta, M., Dumontier, M., Egaña Aranguren, M., Schulz, S., Aussenac-Gilles, N., and Stevens, R. (2013). Evaluation of the OQuaRE framework for ontology quality. <i>Expert Systems with Applications</i> , 40(7):2696–2703                              |
| 2013 | Egaña Aranguren, M., Fernández-Breis, J. T., Antezana, E., Mungall, C., Rodríguez González, A., and Wilkinson, M. D. (2013). OPPL-Galaxy, a Galaxy tool for enhancing ontology exploitation as part of bioinformatics workflows. <i>Journal of biomedical semantics</i> , 4(1):2                   |
| 2012 | Minarro-Gimenez, J., Egana-Aranguren, M., Villazon-Terrazas, B., and Fernandez-Breis, J. (2012). Publishing Orthology and Diseases Information in the Linked Open Data Cloud. <i>Current Bioinformatics</i> , 7(3):255–266   |
| 2011 | Mironov, V., Antezana, E., Egaña, M., Blondé, W., De Baets, B., Kuiper, M., and Stevens, R. (2011). Flexibility and utility of the Cell Cycle Ontology. <i>Applied Ontology</i> , 6(3):247–261   |

- 2011 Miñarro-Gimenez, J., Aranguren, M., Béjar, R., Fernández-Breis, J., and Madrid, M. (2011). Semantic integration of information about orthologs and diseases: The OGO system. *Journal of biomedical informatics*, 44:1020–1031
- 2009 Antezana, E., Egaña, M., Blondé, W., Illarramendi, A., Bilbao, I., De Baets, B., Stevens, R., Mironov, V., and Kuiper, M. (2009b). The Cell Cycle Ontology: an application ontology for the representation and integrated analysis of the cell cycle process. *Genome Biol*, 10(5):R58
- 2009 Antezana, E., Blondé, W., Egaña, M., Rutherford, A., Stevens, R., De Baets, B., Mironov, V., and Kuiper, M. (2009a). BioGateway: a semantic systems biology tool for the life sciences. *BMC bioinformatics*, 10(Suppl 10):S11
- 2008 Egaña Aranguren, M., Wroe, C., Goble, C., and Stevens, R. (2008). In situ migration of handcrafted ontologies to reason-able forms. *Data & Knowledge Engineering*, 66(1):147–162
- 2008 Aranguren, M., Antezana, E., Kuiper, M., and Stevens, R. (2008a). Ontology Design Patterns for bio-ontologies: a case study on the Cell Cycle Ontology. *BMC bioinformatics*, 9(Suppl 5):S1
- 2008 Antezana, E., Egaña, M., De Baets, B., Kuiper, M., and Mironov, V. (2008b). ONTO-PERL: an API for supporting the development and analysis of bio-ontologies. *Bioinformatics*, 24(6):885
- 2007 Stevens, R., Egaña Aranguren, M., Wolstencroft, K., Sattler, U., Drummond, N., Horridge, M., and Rector, A. (2007). Using OWL to model biological knowledge. *International Journal of Human-Computer Studies*, 65(7):583–594
- 2007 Aranguren, M., Bechhofer, S., Lord, P., Sattler, U., and Stevens, R. (2007). Understanding and using the meaning of statements in a bio-ontology: recasting the Gene Ontology in OWL. *BMC bioinformatics*, 8(1):57

## Book Chapters

- 2010 Aranguren, M., Stevens, R., Antezana, E., Fernández-Breis, J.T., Kuiper, M., and Mironov, V. (2010). Technologies and Best Practices for Building Bio-Ontologies. In *Knowledge-Based Bioinformatics*, volume Gil Alterovitz and Marco Ramoni (Eds.), pages 67–86. Wiley Online Library

## Books

- 2010 Aranguren, M. (2010). *Role and application of ODPs in bio-ontologies*. Lambert Academic Publishing

## Conference Proceedings

- |      |   |
|------|---|
| 2014 | González, A. R., Romero, M. M., Aranguren, M. E., and Wilkinson, M. D. (2014b). Nanopublishing clinical diagnoses: tracking diagnostic knowledge base content and utilization. In <i>27th International Symposium on Computer-Based Medical Systems (CBMS)</i> , pages 335–340                                    |
| 2013 | Iglesias, A. R., Aranguren, M. E., González, A. R., and Wilkinson, M. D. (2013). Plant Pathogen Interactions Ontology (PPIO). In Rojas, I. and Guzman, F. M. O., editors, <i>IWBBIO</i> , pages 695–702. Copicentro Editorial   |
| 2011 | Aranguren, M., Fernández-Breis, J., and Antezana, E. (2011). OPPL-Galaxy: enhancing ontology exploitation in galaxy with OPPL. In <i>Proceedings of the 4th International Workshop on Semantic Web Applications and Tools for the Life Sciences</i> , pages 12–19. ACM  |
| 2010 | Miñarro-Giménez, J., Aranguren, M., García-Sánchez, F., and Fernández-Breis, J. (2010). A semantic query interface for the OGO platform. In <i>Information Technology in Bio-and Medical Informatics, ITBAM 2010</i> , pages 128–142. Springer  |
| 2008 | Egaña, M., Rector, A., Stevens, R., and Antezana, E. (2008). Applying ontology design patterns in bio-ontologies. In Gangemi, A. and Euzenat, J., editors, <i>Knowledge Engineering: Practice and Patterns</i> , volume 5268 of <i>Lecture Notes in Computer Science</i> , pages 7–16. Springer Berlin Heidelberg |

## Preprints, Workshop Proceedings and other publications

- |      |  |
|------|--|
| 2015 | Aranguren, M. E. (2015). Merging OpenLifeData with SADI services using Galaxy and Docker (DOI 10.1101/013615). <i>BioRxiv, Cold Spring Harbor Labs</i>   |
| 2012 | Horridge, M., Aranguren, M., Mortensen, J., Musen, M., and Noy, N. (2012). Ontology Design Pattern Language Expressivity Requirements. In <i>WOP (Co-located with ISWC), Boston, USA</i>                     |
| 2012 | Gimenez, J., Aranguren, M., and Tomas, J. (2012). NCBO-Galaxy: bridging the BioPortal web services and the Galaxy platform. In <i>ICBO, Graz, Austria</i>  |
| 2009 | Fernandez-Breis, J., Aranguren, M., and Stevens, R. (2009). A quality evaluation framework for bio-ontologies. In <i>ICBO, Buffalo, USA</i>  |
| 2009 | Aranguren, M., Stevens, R., and Antezana, E. (2008b). Transforming the Axiomisation of Ontologies: The Ontology Pre-Processor Language. In <i>OWL Experiences and Directions (OWLEd), Washington DC, USA</i> |

2008	Iannone, L., Egana, M., Rector, A., and Stevens, R. (2008). Augmenting the expressivity of the ontology pre-processor language. In <i>OWL Experiences and Directions (OWLEd)</i> , co-located with ISWC, Karlsruhe, Germany
2008	Antezana, E., Blondé, W., Egana, M., Rutherford, A., Stevens, R., De Baets, B., Mironov, V., and Kuiper, M. (2008a). Structuring the life science resourceome for semantic systems biology: lessons from the BioGateway Project. In <i>Proceedings of the Workshop on Semantic Web Applications and Tools for Life Sciences (SWAT4LS): November 28, 2008; Edinburgh, United Kingdom</i>
2012	Marshall, M., Boyce, R., Deus, H., Zhao, J., Willighagen, E., Samwald, M., Pichler, E., Hajagos, J., Aranguren, M., Miller, M., Prudhommeaux, E., Dumontier, M., and Stephens, S. (2012). Health Care and Life Science (HCLS) Linked Data Guide ( <a href="http://www.w3.org/2001/sw/hcls/notes/hcls-rdf-guide/">http://www.w3.org/2001/sw/hcls/notes/hcls-rdf-guide/</a> )
2007	Aranguren, M. (2007). ¿Qué puede hacer la web semántica por la biología? In <i>BioGaia 7</i>
2003	Aranguren, M. (2003). Software libre (GNU/linux) para biólogos. In <i>BioGaia 3</i>

## Invited Talks

2016	Los Datos Enlazados y la Web Semántica. Tikitalka, VE Interactive, Spain
2014	Building reasonable biomedical ontologies for a Life Sciences Semantic Web. 3S (Systems, Synthetic, and Semantic) Biology summer school. CIBIO (Centre for Integrative Biology), University of Trento, Italy
2011	Linked Data for Functional Genomics. NTNU, Trondheim, Norway
2010	Aplicación de la Web Semántica en Biología Molecular. Universidad de Deusto, Facultad de Ingeniería, Spain
2008	Aplicación de la Web Semántica en Bioinformática. UM, Facultad de Informática, Spain
2004	Métodos y resultados actuales en Bioinformática: know-how y know-what de las redes tecnocientíficas en Bioinformática. EHU, Facultad de Filosofía, Spain

## Teaching Experience

2016	Life Sciences Semantic Web. MSc Bioinformatics, UM. Spanish
------	---

2015	Linked Open Data tutorial. IZFE. Spanish
2015	Life Sciences Semantic Web. MSc Bioinformatics, UM. Spanish
2014	Semantic biology tutorial: Use of Semantic Web resources for knowledge discovery. 3S (Systems, Synthetic, and Semantic) Biology summer school. CIBIO (Centre for Integrative Biology), University of Trento, Italy. English
2014	Galaxy tutorial. Erasmus mundus MSc in Marine Environment and resources, UPV-EHU. English
2014	Life Sciences Semantic Web. MSc Bioinformatics, UM. Spanish
2013	Introductory talk on bioinformatics for high school students visiting the CBGP. Spanish
2013	Galaxy tutorials at CBGP. English and spanish
2013	Life Sciences Linked Data. MSc Bioinformatics, UM. Spanish
2012	OWL, as part of ATHENS course (UPM). English
2012	OWL, as part of ATHENS course (UPM). English
2011	Populous tutorial at SWAT4LS (London, UK), English
2011	OWL, as part of ATHENS course (UPM). English
2011	Web Ontology Language (OWL), as part of Artificial Intelligence MSc (UPM). English
2011	OWL/Description Logics, as part of the Artificial Intelligence course (UPM). Spanish
2005-2008	OWL tutorials for biologists (University of Manchester, UK). English

## Ph.D. panels

2013	Meifania Monica Chen, “Lipoprotein Ontology: A Formal Representation of Lipoproteins”, Curtin University. Australia
2012	Jose Antonio Miñarro-Giménez, “Entorno para la gestión semántica de información biomédica en investigación traslacional”. UM, Spain
2011	Doris Mejía Ávila, “Estrategia de interoperabilidad semántica en el contexto de integración de conocimiento geográfico y ambiental. Caso de aplicación: Biodiversity Ontology”. UPM, Spain

## Student supervision

2015	Salvador Alonso Martnez, “Imagen Docker para pipelines de Metagenómica”, Bioinformatics MSc project. UM, Spain
------	--

## Service

2015	<b>Program Committee Member</b> at Linked Data workshop (CAEPIA 2015)
2015	<b>Reviewer</b> for BMC Medical Informatics and Decision Making
2013	<b>Special issue editor</b> for Semantic Web Journal (SWJ): Special issue on Linked Data for Health Care and the Life Sciences
2013	<b>Reviewer</b> for PeerJ
2013	<b>Reviewer</b> for Data and Knowledge Engineering (DKE)
2012	<b>Program Committee Member</b> at Managing Interoperability and compleXity in Health Systems. In conjunction with the ACM International Conference on Information and Knowledge Management
2012	<b>Program Committee Member</b> at Joint Workshop on Semantic Technologies Applied to Biomedical Informatics and Individualized Medicine (SATBI + SWIM 2012). In conjunction with International Semantic Web Conference (ISWC)
2012	<b>Reviewer</b> for BMC Bioinformatics
2012	<b>Reviewer</b> for Journal of Biomedical Informatics (JBI)
2012	<b>Reviewer</b> for Computational and Mathematical Methods in Medicine (CMMM)
2012	<b>Reviewer</b> for Journal of Medical Systems (JOMS)
2012	<b>Reviewer</b> for Journal of Biomedical Semantics (JBS)
2011	<b>Program Committee Member</b> at Managing Interoperability and compleXity in Health Systems. In conjunction with the ACM International Conference on Information and Knowledge Management
2011	<b>Program Committee Member</b> at Knowledge Capture (K-CAP)
2011	<b>Program Committee Member</b> at Semantic Applied Technologies on Biomedical Informatics (SATBI 2011). In conjunction with the ACM International Conference on Bioinformatics and Computational Biology
2011	<b>Reviewer</b> for Semantic Web Journal (SWJ)

2011	<b>Reviewer</b> for Journal of Research and Practice in Information Technology (JRPIT)
2008	<b>Program Committee Member</b> at ONTORACT

## Collaboration with research groups in publications

2015	Computer Science. University of Saskatchewan, Canada
2015	Genetics. Technical University of Cartagena, Spain
2015	Biological Informatics Group. CBGP (UPM), Spain
2014	Biological Informatics Group. CBGP (UPM), Spain
2014	Dumontier lab. Carleton University, Canada
2014	Bio-Medical Informatics Research Group. Stanford University, USA
2013	Berkeley Bioinformatics Open-source Projects. Lawrence Berkeley National Laboratory, USA
2013	Methods and Engineering of Language, Ontology and DIscourse. Toulouse Institute of Computer Science Research, France
2013	Institute for Medical Informatics, Statistics and Documentation. Medical university of Graz, Austria
2012	Dumontier lab. Carleton University, Canada
2012	Bio-Medical Informatics Research Group. Stanford University, USA
2012	Isoco Labs. Isoco, Spain
2009	Noray bioinformatics, Spain
2009	Tecnologías de Modelado, Procesamiento y Gestión del Conocimiento. Universidad de Murcia, Spain
2008	Semantic Systems Biology. Trondheim University, Norway
2008	KERMIT. Gent University, Belgium
2005	Information Management Group. University of Manchester, UK
2005	Bio-Health Informatics Group. University of Manchester, UK



## Technical skills

Semantic Web and Linked (Open) Data (Advanced): RDF, RDFS, SPARQL, OWL, SWRL, JSON-LD, OWL API, RDFLib, ONTO-PERL, Jena, Sesame, Virtuoso, Stardog, Blazegraph, Pubby, Protégé 4, TopBraid composer, CKAN, D2RQ

Programming languages (Medium): Java, Perl, Python

Programming languages (Basic): JavaScript, JSP, Bash

UNIX systems (Medium): GNU/Linux (Debian, Ubuntu, CentOS, Red Hat)

Markup languages (Medium): XML, HTML 5, CSS, L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub>, Markdown

Web Services (Basic): SADI framework, NCBO services

Web (Basic): Galaxy (Bioinformatics workflows), Apache, Nginx, Tomcat, Jetty, lighttp, Solr/Lucene, Wordpress, Drupal, Jekyll, NodeJS, Spring-MVC

Software development (Basic): Maven, Ant, SCons, Eclipse, Subversion, Mercurial, Git

Project management (Basic): Jira, Trac, Scrum

SQL and NoSQL Databases (Basic): MySQL, PostgreSQL, Neo4j

File based data storage (Basic): HDF5, YAML, JSON

Statistical analysis (Basic): R

Cloud computing (Basic): Amazon EC2

Virtualisation (Basic): Docker, Virtual Box

## Other merits

ANECA certificate: Ayudante Doctor

Member of Open Knowledge Foundation Spain (OKFN-ES) and Basque Association of Biologists (COBE)

Former member of the W3C Semantic Web Health Care and Life Sciences Interest Group and the Spanish Association of Linked Data (AELID)

## Useful profiles

GitHub: <http://github.com/mikel-egana-aranguren>

Google Scholar: <http://scholar.google.com/citations?user=JsMMKnoAAAAJ>

Scopus: <http://www.scopus.com/authid/detail.url?authorId=16038705500>

ResearchGate: [http://www.researchgate.net/profile/Mikel\\_Egana](http://www.researchgate.net/profile/Mikel_Egana)

Last updated: February 29, 2016