Daniel Macias-Galindo

birth date	July 4, 1982
address	Southbank VIC 3006, Australia
email	daniel.macias.galindo@gmail.com
website	http://daniel-maciasgalindo.info
github	deilanexe

I am interested in topics related to concept classification, from Taxonomies to Ontologies, from semi-automatic construction to data analysis, visualisation, representation and application. Also, I am interested in the software development process, system architecture, APIs and better ways to help communicating ideas. More particularly, in the quick prototyping and development of "proof-of-concepts" that can be the principle of new products. Recently I have been exploring and programming APIs and services for voice-operated devices like Alexa and Google Assistant.

Interests

Knowledge Representation and Reasoning, Computational Linguistics, Information Retrieval, Conversational Agents/Systems, Machine Learning.

Education

PhD in Computer Science (2014) - RMIT University

My research involved the automatic construction of *modular ontologies* representing the vocabulary of a topic (e.g. zoos, museums, food). These ontologies were incorporated in a conversational system capable of engaging in chat-like dialogue with users. I also investigated the effect of conversational domains in the human perception and automatic measures of semantic relatedness.

Master in Computer Science (2007) - Benemérita Universidad Autónoma de Puebla

BEng in Computer Engineering (2005) - Universidad Popular Autónoma del Estado de Puebla

Working Experience

Ontology Lead (May 2013 - Present)

Search Capability Team, Search and Knowledge Discovery | Sensis Pty Ltd

Projects

Chatbots and Voice-operated services (Jan 2017 - Present)

Some of the prototypes I have worked with include: skills for Amazon Alexa, actions for Google Home, chatbots for Facebook Messenger and Slack. The applications implemented range from search capabilities through our Sensis API, to work-related tasks such as validating input information that we require when servicing queries from sales representatives regarding our search engine. I have implemented these capabilities in Python Flask APIs using Natural Language Understanding platforms such as LUIS, API.ai and recast.ai.

SQUAT II (Jul 2016-Present)

The Search Quality Assessment Tool, or SQUAT, is an internally-developed framework where search analysts can debug queries to their search platform. In SQUAT, analysts create *gold sets*, which are collections of user *queries* with a set of expected results, also called *expectations*. In its second implementation, SQUAT enables map-based search, gold set result comparisons, and an API to connect to that provides the endpoints required to operate these functions.

As part of my contribution to SQUAT II, I got involved with the redesign of the user interface, including some of the functional components such as the creation of gold sets and importing/exporting gold sets into and from spreadsheets for easy operation outside of the tool. I was also in charge of documenting cards for an offshore developer who is in charge of implementing these changes.

Knowledge Network (Ontology 2.0; Mar-Sep 2016)

One of the major challenges with the Sensis Ontology program was the scale of the proof-of-concept system against the final production version, which was smaller in comparison. Consequently, some of the tasks of the construction process that were originally implemented sequentially could not be completed even in finite time.

The second version of the Sensis Ontology, which was also rebranded as the "Knowledge Network", consisted of re-thinking the original implementation that produced the database contents as tasks that could be executed in parallel. In addition to this, we reviewed and simplified some of these tasks to reduce dependencies and speed up the overall process to produce the network.

In the end, the process to construct the Knowledge Network passed from infinite time to less than six months (as with the original PoC), and a ten-fold increase in performance.

Ontology Opportunities (Jul 2015 - Jun 2016)

With the Sensis Ontology in production, the next step consisted of demonstrating the capabilities of our product to other business units within Sensis. This required face-to-face conversations with potential stakeholders, understanding their pain points and working towards solutions for them that could use our product.

By making a collection of endpoints available within the company, I was able to ask around the company for potential stakeholders with areas of improvement that could be addressed by developing quick "proof-of-concepts". Some of these prototypes are the following:

- Category Suggestion: As Yellow Pages operates under the existence of "headings", currently the process to assign a heading is manual and requires expertise of sales representatives on 2000+ categories where businesses can be classified into. Our PoC uses small snippets of text and information in the Sensis Ontology for suggesting the best heading an advertiser can appear into in order to be relevant. This prototype has been tested by different business units so far and internally within the Search and Knowledge Discovery team.

- Keyword Validation: Another manual process consists of assigning keywords to advertisers' *Business Profile Pages* that help them distinguish themselves from their competitors. Our prototype used information available for each advertiser in the Sensis Ontology to populate an initial set of relevant keywords. The original concept of this project has evolved over the past months and will be put into production in 2017.

- Website Leads: As part of the Marketing Acquisition strategy, we implemented a prototype that collected information from advertisers' websites based on quality display metrics, such as Google's Mobile Friendly Test. This tool was mostly used as a prototype, but was run as such intermittently over 2016, with casual specific requests over two more months.

- Product Search: Sensis Ontology can be used to classify and detect concepts representing brands, products and services. With classifications openly available in resources like Wikipedia, we could use these to arrange our concepts, thus enabling a new search capability that extends what the Sensis API can currently achieve. Product Search was tested with data from Skip, but due to clashing priorities was never put into production.

Ontologist (March 2013 - May 2014)

Search Capability Team, Search API | Sensis Pty Ltd

Projects

Search Capability program (Oct 2010 - Present) (http://developers.sensis.com.au/)

The Search Capability Program has built a heavily-tuned SOLR-based search engine currently servicing millions of requests per day, serving up highly relevant search results for Business, Government and Points of Interest for Yellow Pages, WhereIs (Mobile) and numerous external partnerships.

Working with the SAPI development team, the Search Capability team consists primarily of Data Scientists who leverage interaction data and content from across Sensis (Yellow Pages, White Pages and WhereIs) to automate the measurement and improvement of the SAPI Search algorithm and associated Ontologies for a dramatically improved User Experience and Advertiser ROI.

Sensis Ontology Program (Mar 2013 - Jun 2014)

The Sensis Ontology Program is an enabler for richer classification of Business, Government and Geospatial content and data across Sensis which aims to move classification away from manually managed print heading structures and into automated digital clusters of terms, concepts and categories.

This is designed to assist in providing fresher, more accurate and deeper content for more accurate retrieval of listings via SAPI Search, but also provides a significant boost in supporting data for business intelligence and data science teams.

Accountabilities

Problem-solving skills

- I enjoy finding solutions to new problems in my working group. From finding alternatives to approach problems differently to producing quick prototypes that can be later used for demonstrations and showcases.
- I am constantly challenged by designing and implementing tools in an Agile environment. An example of this is Sensis Ontologies, now in its second iteration, which has been constantly reviewed due to changing architecture and constrained hardware resources, while still being capable of producing solutions on time.
- I have started developing software products for personal use, following an Agile approach and sharing my code through Git platforms (e.g. Github, SourceTree, Gitorious).
- I have recently joined hackatons for good causes (Random Hacks of Kindness, GovHack), always with the idea of finding a new challenge. This has forced me to be faster at developing prototypes and deploying them online, in order to present demonstrable results.

Presentation skills

- As part of my formation as a researcher, I am capable of communicating ideas to any kind of audience, including classrooms for at least the past twelve years.
- When presenting, I tend to use visual materials as a way to engage with the audience rather than as supportive material for myself. I have tried different products, from PowerPoint and Beamer, to more visual and interactive platforms like Prezy and reveal.js
- I am continuously looking for opportunities to present to an audience, and to share my ideas with my peers.

Team member

- In the Sensis Search API team, due to its Agile formation, I constantly must communicate our progress and roadblocks that we can devise before these become issues. This includes not only the Sensis Ontology, but also other products and projects that have been under my charge (refer to **Projects** section).
- Documentation is an important part of any project, and I understand passing an idea through text sometimes does not reflect the whole picture to other team members. Therefore I tend to create thorough and informative cards, especially when these will be taken by developers. When these are taken, I join the involved team, and explain to them the business benefits that the card represents.
- I am also proactive, by following up current status of cards and by proposing solutions, in case of any roadblock affecting developers and/or testers. Given that at some point I have executed roles as both developer and tester, I am familiar with the product life cycle.
- When possible, I would rather work in a team than on my own. But I can deliver results in both scenarios.

Voluntary Work

- Webmaster of the IEEE Australia and New Zealand Student Conference (ANZSCON'2010). The site was implemented in HTML and PHP.
- Web development leader for the IEEE RMIT Student Branch during 2010. The site was developed in WordPress.
- Webmaster of the RMIT Computer Science Research Student Conference 2009. Also conducted logistic duties before and during the conference. The site was developed in both PHP and Flash.
- Volunteer in the Summerfest 2009, University of New South Wales, Sydney, NSW. December 2009. In charge of setting up projectors and rooms for presenters and session chairs.
- Volunteer in the ANU Summer School of Logic and Machine Learning, Canberra, ACT, Australia. January 2009. In charge of setting up presentation materials in rooms for lecturers.

Latest Publications

(only the last five publications are shown; refer to my website for the full list)

Effects of Domain on Measures of Semantic Relatedness. *Daniel Macias-Galindo, Lawrence Cavedon, Wilson Wong, John Thangarajah*. In *Journal of the American Society for Information Science and Technology*. (Journal, 2015). Volume 66, Number 10. Pages: 2116-2131. ISSN: 2330-1643.

Coherent Topic Transition in a Conversational Agent. *Daniel Macias-Galindo, Wilson Wong, John Thangarajah, Lawrence Cavedon*. In INTERSPEECH. (Poster, 2012). Portland, OR.

Using a lexical dictionary and a Folksonomy to automatically construct domain ontologies. *Daniel Macias-Galindo, Lawrence Cavedon, Wilson Wong, John Thangarajah*. In *Australasian Joint Conference on Artificial Intelligence*. (Conference proceedings, 2011). In Perth, WA.

Building Modular Knowledge Bases for Conversational Agents. *Daniel Macias-Galindo, Lawrence Cavedon, John Thangarajah*. In Knowledge and Reasoning in Practical Dialogue Systems. (Conference proceedings, 2011). Workshop during IJCAI'11, in Barcelona, Spain.

Computer Skills

Programming Languages

- Python (2.7) (Advanced | Self-taught, Online courses, Used for work)
- Java (1.7) (Advanced | Learned at school, Used for research)
- C/C++ (Intermediate | Learned at school)
- PHP (5) (Intermediate | Learned at school, Self-taught, Used for research)
- LaTeX (Intermediate | Self-taught, Used for research)
- R (Intermediate | Self-taught, Used for research)
- JavaScript (Intermediate | Self-taught, Used for work)
- Bash (awk, grep, sed) (Beginner | Self-taught, Used for research, Used for work)

DB Managers

- MySQL (Advanced | Learned at school, Used for research, Used for work)
- Oracle (Intermediate | Self-taught, Used for work)
- PostgreSQL (9.5) (Intermediate | Self-taught, Used for work)
- MongoDB (Intermediate | Self-taught, Used for work)
- Solr (5.0) (Intermediate | Self-taught, Used for research, Used for work)
- ElasticSearch (2.3) (Intermediate | Self-taught, Used for work)

Operating Systems

- Microsoft Windows (XP, Vista, 7, 8, 10) (Advanced | Self-taught, Online courses)
- Linux (Xubuntu, Ubuntu, AMI for AWS) (Advanced | Self-taught, Used for research, Used for work)

Cloud Service Providers

- Amazon Web Services (EC2, RDS, IAM) (Advanced | Self-taught, Used for work)

Natural Language Understanding platforms

- API.ai (Intermediate | Self-taught, Used for work)

Closing date: Mar 19, 2017