

# Rodrigo Palacios

3631 Utah Ln.  
Clovis, CA 93619

559 930 0910

personal  
[rodrigopala91@gmail.com](mailto:rodrigopala91@gmail.com)  
<https://rodricios.github.io>

GitHub  
<https://github.com/rodricios>

## Projects

Maintaining the following (Python) repositories on GitHub:

libextract<sup>1</sup> - Libextract is a Python-based data extraction library that works on HTML and XML documents.

eatih<sup>2</sup> - Implemented novel unsupervised classification algorithm to extract content from web pages. Currently being integrated by folks from blockspring.com, papeur.com.

autocomplete<sup>3</sup> - Markov-chain-powered "autocomplete". Produced both as an API and as a learning resource for young kids and adults not familiar with practical AI.

## Skills, Languages, Tools, etc.

- Strong prototyping skills with Python
- Experienced with Python, C++, C
- Comfortable working in a Bash shell, Linux/Ubuntu environment
- Experienced with continuous-integration frameworks (Travis-CI)
- Source control: Git, Team Foundation
- SQL & NoSQL experience: SQL Server, MySQL, Elasticsearch, MongoDB

## Experience

Research Internship, University of Notre Dame, summer 2015, advisor: Tim Weninger

- Tasked with creating original, stochastic graph generation algorithms
- Written in Python, these novel machine learning algorithms can produce graphs similar to real-world graphs.

Contractual work at Visual Impact<sup>4</sup> included:

- Setup data pipelines between database and UI for large Ag. company.
- Multistage, agile work environment using C#, ASP.NET frameworks

59daysofcode competition, summer 2014

- Lead team of 3, created elasticsearch-based in-video search engine.
- Launched cluster of Linux/Ubuntu nodes running custom Python crawling scripts.

Project Leader - Fresno State, Software Architecture. CSci 152, Spring 2014

- Lead team of 4, tasked with implementing search engine of open-access journals (C#, ASP.NET frameworks).

## Publications

Tim Weninger, Rodrigo Palacios, Valter Crescenzi, Thomas Gottron and Paolo Merialdo  
**Web Content Extraction - a Meta-Analysis of its Past and Thoughts on its Future.**  
*ACM SIGKDD Explorations* Accepted. 2015.

## EDUCATION

**Fresno State**, 2013 to Currently

**Willow International, Fresno City and Fresno State (OpenU)**, 2010 to 2013

---

<sup>1</sup> <https://github.com/datalib/libextract>

<sup>2</sup> <https://github.com/rodricios/eatih>

<sup>3</sup> <https://github.com/rodricios/autocomplete>

<sup>4</sup> [info@v-i.us](mailto:info@v-i.us)