

# John DeSanto

415-420-6742 (m)

[desanto@mac.com](mailto:desanto@mac.com)

San Francisco, CA

---

## Qualifications

- Over 20 years' experience in information technology, including 5 years in leadership roles
- Information retrieval, statistical NLP, search relevance, sentiment analysis
- Ontology and knowledge base development
- Graduate degrees in mathematics and philosophy

## Technology

- Python, Go, Lisp, Scala
  - MySQL, Redshift, Snowflake, RDS, MongoDB, DynamoDB, AWS, Redis
  - OWL, RDF/S, SPARQL
  - PySpark, TensorFlow, Pandas, NLTK
- 

## Employment History

### Ontologist (Consulting)

July 2011 – Present

*Aevea Consulting (Self-employed) - San Francisco, CA*

- Developed topic specific ontologies and related tools for multiple clients. Select ontologies: academic psychology, agriculture, clothing and merchandise.
- Defined scope and methods for ontology representation, inferential reasoning, and search APIs. Responsible for all design, ontology representation, and code.

### Staff Engineer

November 2018 – Jan. 2020

*Medium - San Francisco, CA*

- Member of the data engineering team supporting Medium's data infrastructure.
- Improved and maintained 200+ Scala/Spark ETL pipelines, including throughout a full database migration of ~20TB from Redshift to Snowflake.
- Owned development for improving our in-house job orchestration framework to support PySpark ETL pipelines and on-demand, single-instance Python machine learning jobs. Responsible for requirements and implementation.

### Principal Developer, Foundation

May 2016 – September 2018

*EvengX (formerly TrustPipe Security) - San Francisco, CA*

- Co-created and led development of EvengX's probabilistic method for identifying malware. Built initial prototype and owned requirements for production implementation
- Utilized traditional NLP techniques to create custom k-means clustering methods for identifying malware and non-malware in network packet data.
- Developed and maintained processing/testing pipelines to assess model performance.

### Staff Software Engineer, Architecture

April 2015 – May 2016

*Vungle - San Francisco, CA*

- Developed and implemented strategies for capturing, storing and utilizing data generated by Vungle's platform and SDKs. Created data pipelines to improve ad customization and data reconciliation.
- Team lead for designing and implementing Vungle's advertising exchange in Go. We built an advertising exchange from the ground up in two quarters.

## **Software Architect**

April 2011 – March 2015

*AdPredictive - San Francisco, CA*

- Created and implemented strategies for improving AdPredictive's programmatic advertising platform. Optimized bidding engines for large-scale ad arbitrage, including real-time Bayesian analysis for adjusting bid prices.
- Wrote high-capacity, low latency code in Python for real-time bidding (RTB) on ad impressions and offline analytical tools for processing user data using DynamoDB and Amazon EMR.

## **Senior Ontologist**

August 2007 – July 2008

*NASA / Intrinsic - Mountain View, California*

- Created and tested OWL/RDFS ontologies for NASA data systems using TopBraid Composer and XML Schema tools.
- Worked with SMEs to establish model validity. Created and presented a common data model for integrating multiple ontology formats.

## **Director of Research and Development**

August 2004 – October 2006

*Kozoru - Overland Park, Kansas*

- Led Kozoru's research and development efforts and contributed to the newLISP open source project. Managed junior developers. Developed answer ranking algorithms, distributed source acquisition code and methods for testing query response changes
- Created and released a customizable chat-based search product, BYOMS (build your own mobile search), using in-memory data store generated in real time. Implemented a distributed statistical parser using a customized map/reduce framework

## **Ontologist**

October 2002 – July 2004

*Unicru - San Francisco, California*

- Led development of the SmartMatch ontology. Supervised knowledge base content editors, developed ontology content and KB functional requirements

## **Ontologist, Project Manager**

September 1998 – March 2001

*Cycorp - Austin, Texas*

- Created and tested new areas of the ontology for the Cyc knowledge base, including CycSecure, DARPA High Performance Knowledge Base (HPKB), Rapid Knowledge Formation (RKF), Hotbot/Lycos. Wrote DARPA proposals, including for DARPA's AML
- Supervised four ontologists and programmers and managed project KB development. Oversaw all subject matter experts and reported to DARPA program manager.

## **Education**

- **The Johns Hopkins University**  
M.A. Philosophy, 1995
- **The University of Colorado**  
M.A. Mathematics, 1991
- **The University of Michigan**  
B.S. Mathematics and Philosophy, 1989

## Publications

- Dan Crow and John DeSanto, [\*A Hybrid Approach to Concept Extraction and Recognition-based Matching in the Domain of Human Resources\*](#). 16th. IEEE International Conference on Tools with Artificial Intelligence (ICTAI-04), Boca Raton 2004.
- Martin Quiroga, John DeSanto and John Flowers, "[Natural language based search engine and methods of use therefore](#)", U.S Patent #7,447,683.