Aubrey J. Rembert

CONTACT
INFORMATION

1495 Fuji Drive
Melbourne, FL 32940

phone: (720) 989-8400
email: info@aubreyjrembert.com

PROFILE

Aubrey J. Rembert is an *Innovative Data Scientist* with nearly 20 years of experience in academic and industrial research. He is passionate about designing and directing the implementation of large-scale data-driven decision-making tools that help solve *Real-world Business Problems*.

PROFESSIONAL
EXPERIENCE

Pandora Media, Inc, New York, NY

<u>Senior Data Scientist</u> - *Content Science Group (Remote Employee)* *November, 2018 - Present*

- *Researcher for Artist Disambiguation Project*: This problem refers to the situation in which multiple artists have the same name. I reformulated this problem as binary classification using Gradient Boosted Trees. This led to an improvement in prediction accuracy as well as a 3X improvement in the number of artists that could be disambiguated.
- *Researcher and Developer for Social Entity Resolver System*: This system determines if different social profiles represent the same logical artist. I re-architected the task dependency model, and improved the scalability and prediction model of this system. This system is used to enable the computation of the social metrics of an artists' Next Big Sound profile.

<u>Senior Data Scientist</u>, *Next Big Sound Group (Remote Employee)* *May, 2015 - November, 2018*

- *Research and Development Lead for Large Scale Song Resolution System*: This system provides a solution to the resolution, disambiguation, and integration of songs from multiple streaming sources and social networks. I conceived of, prototyped, and productionized this system. Today it is used to power all of the song based metrics at Next Big Sound.
- *Research Lead for Venue Capacity Prediction and Recommendation System*: This system predicts the venue capacity and recommends the city and venue in which a particular artist should perform. This research led to many tour managers for high profile artists reaching out to our team for city and venue suggestions.
- *Development Lead for Charts ETL Pipelines*: I architected, prototyped, and implemented the ETL pipelines for all of Pandora's public facing charts (Trendsetters, Predictions, Top Spins). A few of these charts power popular Pandora stations.
- *Researcher for the Song Streaming Behavior Prediction Model*: I led the research and algorithm development of an autoregressive model to predict the future streaming behavior of songs.

Florida Institute of Technology, Melbourne, FL

<u>Assistant Professor</u>, *Information Systems (joint w/ Computer Science)* *August, 2014 - May, 2015*

- *Research Lead for Novel Process Mining Project*: Led the research, design, and development of a Bayesian Model Merging approach to learning business process models from event data.
- *Research Co-Lead for Information Extraction in Reseller Markets Project*: Co-led the research and development of an information extraction engine used to discover, and predict bargains for a-prior unknown entities in a reseller market.

IBM T. J. Watson Research Center, Yorktown Heights, NY

<u>Research Staff Member</u>, *Customer Experience Lab (CXLab)* *August, 2012 - August, 2014*

- *Research Lead on Predictive Models for Financial Customer Profiles Project*: I led the research

and prototype development of a machine learning system that leveraged structured and unstructured data, as well as dependency models to predict the financial state of a customer. The result of this work was used to inform customer service representatives about how to best interact with customers.

- *Research Lead on Unstructured Data Analysis for Financial Services Project*: I lead the develop of an information extraction system that discovers financial entities from unstructured dialog transcripts. This work was used to power customer profile discovery as well as other financial services predictive models.

Research Staff Member, *Business Service Informatics Lab*   *August, 2008 - August, 2012*

- *Research Lead on the Knowledge-based Business Process Insights Project*: I conceived of, designed, and prototyped a novel process discovery algorithm that leverages the combination of prior knowledge and process execution data. The results of this project were used to drive innovation in a number of areas including: security log analysis, on-line shopping behavior analysis, customer journey modeling, business process modernization, and program understanding.
- *Research Co-Lead on the Business Process Clustering and Retrieval Project*: We developed a novel approach to organizing and retrieving business process assets in a repository by combining language based topic models built over a corpus of procedure documents along with the graph similarity models built using the structure of business process models. This work was inspired by my previous work on the Process Model Anonymization and Transformation project.
- *Researcher on the Interactive Blueprinting Project.* I developed an approach to interact with and annotate enterprise models for the purpose of collaborating on enterprise transformations projects.
- *Researcher on the Automatic Business Process Model Anonymization and Transformation Project*: I enhanced an automatic process diagram correction algorithm by replacing the hand-crafted rules it used with a machine learning approach that leveraged the textual and geometric features of process diagrams.

University of Colorado, Boulder, CO

Research Assistant, *Collaboration Technology Research Group*   *August, 2003 - August, 2008*

- *Research Lead on Control-flow Discovery*: This work addressed the problem of automatically inferring the control-flow of an organization's business processes from a history of process execution data. Previous solutions to this problem were either not scalable; did not discover points of decision, concurrency, and synchronization; did not discover latent activities (activities that don't explicitly appear in the execution log, but are necessary to model a business process), did not discover cyclic control-flow structures, or used ad-hoc mechanisms to deal with noise and exceptions. I developed a control-flow discovery algorithm that addressed all of the above issues.
- *Research Lead on Activity Discovery*: In control-flow discovery, it is frequently a hidden assumption that the process execution history is at the appropriate level for discovering intelligible control-flow models. However, frequently process execution histories contain very low-level data. Given this, I developed an activity discovery algorithm that incorporates methods from statistical machine learning to find high-level activity classes from low-level process execution histories.
- *Research Lead on Information-Flow Discovery*: Most process discovery research is narrowly focused on learning the control-flow of an organization's processes. This, however, neglects the established concept that a business process is more than control-flow. In this work, I extended the scope of process discovery to automatically discover the information-flow of an organization's business processes.

EDUCATION	**University of Colorado**, Boulder, CO
	Doctor of Philosophy, Computer Science *August 2008*

EDUCATION

University of Colorado, Boulder, CO

Doctor of Philosophy, Computer Science   *August 2008*

- Dissertation: "Automatic Discovery of Workflow Models"

Florida Agricultural & Mechanical University (FAMU), Tallahassee, FL

Master of Science, Computer Science   *May, 2002*

- Thesis: "ISDS: A Hybrid Approach to Software Requirement Specifications"

Bachelor of Science, Computer & Information Sciences - Magna Cum Laude   *May, 2000*

- Minor in Business Administration

HONORS & AWARDS

- IBM Research Eminence and Excellence Award, 2012
- IBM Micro-MBA, 2012
- IBM Patent Plateaus, 2009, 2010, 2012, 2013
- IBM Outstanding Research Accomplishment
- NSF AGEP Fellow, 2003-2008
- Outstanding Teaching Assistant Award, 2003
- Chancellor's Teaching Fellowship, 2003
- NSF CREST Fellow, 2000-2002
- Ronald E. McNair Fellow, 1999
- FAMU FGLAMP Scholar, 1997-1999

SELECTED PUBLICATIONS

- **Rembert A.J.**, Omokpo A., Mazzoleni, P., Goodwin R.T., (2013). Process Discovery with Prior Knowledge. In *Proceedings of the International Conference on Service Oriented Computing (ICSOC)*, December, 2013. Berlin, Germany (12% Acceptance Rate).
- Qiao M., Akkiraju R., **Rembert A.J.**, (2011). Towards Efficient Business Process Clustering and Retrieval: Combining Language Modeling and Structure Matching. In *Proceedings of the International Conference on Business Process Modeling (BPM)*. August, 2011. Clermont-Ferrand, France (13 % Acceptance Rate).
- Mukherjee D., Dhoolia P., Sinha S., **Rembert A.J.**, Nanda, M. (2010). From Informal Process Diagrams to Formal Process Models. In *Proceedings of the Conference on Business Process Modeling (BPM)*. September, 2010. Hoboken, NJ USA (14 % Acceptance Rate)
- **Rembert A.J.**, Ellis C.A. (2009). Learning the Behavioral Perspective of a Business Process using ICN Control-flow Models. In *Proceedings of the International Conference on Service Oriented Computing (ICSOC)*. 2009. November, 2009. Stockholm, Sweden (19 % Acceptance Rate)

COMMUNITY INVOLVEMENT

- Engineer's Week Volunteer Science Teacher (taught Alice Programming), 2010, 2011, 2012, 2013
- Introducing Girls to Engineering Workshop (taught Robot Programming), 2016, 2017, 2018