CURRICULUM VITA
Yohn Jairo Parra Bautista
3550 Esplanade Way apt 8105
Tallahassee, Florida 32311
713.296.0081; yohn.parrabautista@famu.edu
PhD in Computational and Data Enabled Science and Engineering
[Big Data Science, Engineering and Analytics][SEP]
College of Science and Technology[SEP][CST]
Florida A&M University

GOAL: Undertake a research/ teaching career in an academic institution as a knowledge creation researcher using the tools of data science, engineering and analytics while focused on providing: 1. Decision makers the ability to resolve societal/ industrial problems with a better understanding of mitigating situations and their societal impacts; 2. This generation of students the significant opportunities in this new and exciting field through experiential learning and new technologies embracement.

ACADEMIC

 August 2019 -Present, Florida Agricultural and Mechanical University, Department of Computer and Information Sciences. Teaching and assisting in the development and implementation of curriculum in Data Science

EDUCATION

September 2014- December 2019, Jackson State University, Jackson, MS, College of Science, Engineering and Technology PhD Program- Computational and Data Enabled Science & Engineering (Big Data Science, Engineering and Analytics) [94 semester hours] Awarded: May 3, 2020

1. Dissertation; Sentiment Analysis of Consumer Complaints: An Examination of Lexicon-Based and Deep Learning Approaches: We utilize an unsupervised lexicons approach and supervised Naïve Bayes and Unsupervised Word2Vec to create a model to help decision makers understand better the relationships between financial services and predatory lending thematic. We interconnect three general-purpose lexicons and one domain specific lexicon with the idea to train complaint data with those lexicons. Using data exploration with sentiment analysis we found ambiguity between those unigram words related with Predatory Lending Thematic (PLT) into the whole dataset. Word2Vec model as a deep learning approach allow find semantic candidate interchangeable words that match the PLT to create a new corpus to train the model.
2. Research Interests: Text and Unstructured Data Mining and Analysis, Sentiment Analysis, Natural Language Processing, Artificial Intelligence [Machine and Deep Learning applications], Automated Decision Making with applications in Precision Medicine/ Health, Business Intelligence and Deep Learning applications; Global Supply Chain/Logistics
3. Research Tools: RStudio and Python (programming software), Endnote, Hadoop architecture, Spark cluster, SQL, Amazon Web Services (AWS), Tableau, XSEDE, HPC
4. Webinars and Podcasts Attended:
     a. Critical Thinking in Data Science (with Debbie Berebichez),
     b. Patterns for Successful Data Science Projects (Data Science Central, March 14, 2019),
     c. Clean Data & Accurate Machine Learning Models (Data Science Central, February 18, 2019),
     d. Creating Business Applications with R & Python (DSC, February 18, 2019),
     e. AI Models and Active Learning (DSC, December 4, 2018),
     f. State of the Art Deep Learning on Apache Spark (DSC, October 31, 2018)
5. Supplemental Instructions:
     a. Dataquest, provides hands-on training with data science projects

b. Data camp, provides interactive courses on different thematics related to Big Data
c. Edx, provides online courses with different universities that teach the fundamental knowledge of computer science, math and statistics
d. Coursera provides online courses with content across many fields.
e. Certifications in following [undertaken in collaborative learning seminars at Jackson State University under supervision of and examination by professor of record]
   i. Edx, Data Science: Productive Tools (Harvard University's PH125.5x)
   ii. Edx, Natural Language Processing, (Microsoft DEV288x)
   iii. Edx, Deep Learning with TensorFlow, (IBM: DLO120EN)

January 2014–August 2014, Jackson State University English Language Institute

January 2013 - December 2013, Master in Logistics Universidad EAFIT, Medellin, Colombia

September 2000- May 2006 Bachelor of Science in International Business (with honors) Universidad Santo Tomas, Bucaramanga, Colombia [210 semester hours]

INTERNSHIPS and GRADUATE ASSISTANTSHIPS
January 2017- December 2017: College of Science, Engineering and Technology, Jackson State University, Jackson State University [JSU]:
1. Internship: Big Data Analysis for Neurodegenerative Diseases with CSET faculty and Florida International University, Center for Advanced Technology and Education, funded by NSF SCOPE [Scalable Omnipresent Environment] grant [see below], Supervisors: Richard A. Aló PhD, JSU; Malek Adjouadi PhD, FIU
   a. Utilizing the Alzheimer ADNI [Alzheimer's Disease Neuro Imaging] data for:
      • Exploring the social determinants in patients with some antecedent in cognitive impairment using brain images and clinical data.
2. Internship: GIS Projects, funded by NSF SCOPE grant [see below]
   a. Using GIS tools to study at zip code level with tapestry data the social determinants that are high correlate with Alzheimer disease.
January 2016 – December 2016: College of Science, Engineering and Technology, Jackson State University:
3. Internship sponsored by INDRA Technologies, Barcelona, Spain, utilizing their Smart Cities designated Sofia Software Platform for:
   a. Developing a Smart Parking System on campus using GoPro Cameras to capture images to be displayed on a platform (mobile, app, website) that can be used with a language program that Sofia provides to gain insights/ create knowledge for decision makers.
   b. Sofia is a middleware that allows the interoperability of multiple systems and devices, offering a semantic platform to make real world information available to smart applications (Internet of Things)
4. Graduate Assistantship: National Science Foundation sponsored graduate assistantship to work on developing applications for:
   a. NSF SCOPE grant [SCalable OminPresent Environment] with California Institute of Telecommunications and Information Technology [CalIT2] Falko Kuenster PI, UC San Diego; Subcontract to R. A. Alo; Major Instrumentation Grant which constructs and utilizes a 3D visualization wall to provide visualization capabilities and velocity for Big Data Volume

   b. NSF SENSEI Grant [SENSor Environment Imaging] with University of Illinois Chicago, University of California San Diego, CalIT2, Scripps Institute of Technology and University of Hawaii; NSF Major Instrumentation Grant, which is developing spherical sensor camera for marine science and bio medical applications.

January 2015 – December 2015:
1. Internship: Toyota Haley Barbour Internship, Toyota Corporation America:
    a. Projects:
        i.   Developed an automated Geospatial tool that generates potential Toyota customer hot spot segments by integrating the ESRI Tapestry data to be used by Toyota.
        ii. Developed visualization tool (using Geospatial Technologies) focused on end-user commodity data for exports of goods by principal end-user category in USA.

INDUSTRIAL EXPERIENCE
January 2010- December 2012 Claro S.A, Bucaramanga, Colombia,
     Data Analytics in Telecommunications
June 2008- December 2009 LECOMEX- Bogota, Colombia,
     International- Business Intelligence and Analytics
June 2005 - May 2008 Kiddos S.A – Children's Wearing Apparel
     Business Intelligence and Assistant Analyst focus on export-imports

SKILLS
- Native language: Spanish-Reading, Writing, and Speaking Proficiency excellent
- Language: English: Reading, Writing, Speaking Proficiency-excellent
- Computer Tools: Microsoft Office, Data Science/Analytics Software (SAP, R, Python, Tableau, Matlab), GIS Technologies, their integration into Data Analytics, Visualization Wall [Scalable Omni-Presence Environment- a collaborative visualization instrument for exploration of Big Data enabling users to Visualize BD from Nano to Micro to Mega scale]; XSEDE, Hadoop and Spark architectures

MEMBERSHIPS
- IEEE Computer Society
- Association for Computing Machinery (ACM)

EXTRA-CURRICULAR INTERESTS
     Hiking, Gym, Meditation, Music, Physical Fitness, Poetry, Dance, Art, Writing, and Photography

PROFESSIONAL MEETINGS/ DEVELOPMENT/AWARDS
- Workshop: NSF San Diego Supercomputer Center Summer Institute, August 5th to August 9th 2019: week-long workshop focused on a broad spectrum of introductory-to-intermediate topics in High Performance Computing and Data Science for researchers in academia and industry, especially in domains not traditionally engaged in supercomputing, who have problems that cannot typically be solved using local computing resources. SDSC is providing resources to a larger number of modest-sized computational research projects that represent, in aggregate, a tremendous amount of scientific progress. The Summer Institute will give the attendees an overview of topics in High Performance Computing and Data Science and accelerate their learning process through highly interactive classes with hands-on tutorials on the Comet [PetaScale] Supercomputer. https://si19.sdsc.edu/

- Workshop: NSF Cyber Carpentry: Data Life-Cycle Training with DataNet Federation Consortium, July15 to July 26th, University of North Carolina at Chapel Hill. The overarching goal of the Cyber Training program is to prepare, nurture, and grow the national scientific workforce for creating, utilizing, and supporting advanced cyberinfrastructure (CI) that enables

cutting-edge science and engineering and contributes to the nation's overall economic competitiveness and security. http://cybercarpentry.web.unc.edu/

- Summit: Association of Computing Machinery-Institute of Mathematical Statistics Interdisciplinary Summit on the Foundations of Data Science, June 2019, San Francisco, CA-distinguished speakers and panelists addressing topics such as deep learning, reinforcement learning, fairness, and ethics, in addition to discussions about the future of data science and the role of ACM and IMS.
  https://www.acm.org/articles/bulletins/2019/march/acm-ims-data-science-summit-2019

- Award: NSF INCLUDES ASSIST Grant through Great Minds, in STEM, October 6-9, 2018 Early-Career Faculty Symposium delivering faculty development workshops on: the tenure process, NSF proposal and reviewer opportunities, effective mentorship, inclusive pedagogy, innovating practices from research and teaching, and best practices for engaging Underrepresented Minority Students; engagement with the larger HENAAC Conference audience and serve as judges for the Research Posters Competition.

- Award: NSF INCLUDES ASSIST Grant through Great Minds, in STEM October 8-11, 2017-agenda same as above.

- Conference: ACM Richard Tapia, Celebration of Diversity in Computing Conference, September 2017

- Conference: 15th LACCEI International Multi-Conference for Engineering, Education, and Technology, July 2017

- Award: NSF INCLUDES ASSIST Grant through Great Minds, in STEM, October 6-9, 2016 agenda same as above.

- Workshop Award: Global Environment for Network Innovations [GENI] Regional Workshop and GENI Engineering Conference (GEC) 25, March 13-17, 2017, Florida International University: introduction to GENI testbed and keynote speech plus two tutorials: My First GENI Experiment [SDN using OpenFlow], and an advanced track (Wireless, Distributed Computing using Hadoop)..

- Award: Texas Advanced Computational Center Summer Supercomputing Institute: High Performance Computing Training, Austin, Texas, August 1-5, 2016: introduced researchers on high performance computing, data analytics, and scientific visualization; how to effectively use advanced TACC computing resources and technologies for big data analysis and scientific visualization of their research.

- Conference: Supercomputing 2015, November 2015, Austin, Texas

- Conference: Great Minds in STEM October 2015 Pasadena

- Workshop: NSF XSEDE Workshop, New Orleans, September 8-9, 2015: training as Graduate Student Champion for NSF HPC resources

- Conference: 14th international Symposium on Recent Advances in Environmental Health Research, September 2016, Jackson State University.

- Workshop: Logistics in Supply Chain Management, SAP software Logistics, SAP University Annual Conference, Las Vegas, May 2015

PROFESSIONAL PUCLICATIONS / PRESENTATIONS
- Yohn Jairo Parra Bautista Sudha Yerramilli, "Assessing Geographical Inaccessibility to Health Care: Using GIS Network Based Methods", URISA's 2013 GIS in Public Health Conference June 17- 20, 2013 Hyatt Regency - Miami, Florida, 2013
- Yohn Jairo Parra; David Olson; Richard Alo: Comparison of Lexicon Performance on Unstructured Behavioral Data. 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS).
- Yohn Jairo Parra; Moving from Data Analysis to Practice [MAP]: A Study of Health Determinants of COVID-19, GOOGLE LLC COVID-19 AI and Data Science for Social Good Fund Award, September 2020, CPI
- Yohn Jairo Parra Bautista, Richard Aló, and NingNing Wang: "Deep Learning, Cloud Computing for Credit/Debit Industry Analysis of Consumer Behavior," 2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom).
- Parra B Yohn Jairo, Richard A. Aló, and David Olson: "Comparison of Lexicon Performances on Unstructured Behavioral Data," 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS).
- Jinwei Liu, Richard A. Aló, and Yohn Jairo Parra Bautista: "DeepTrace: Improving US Pandemic Health Care through Health Disparity Identification and Determinant Tracing," 2020 International Conference on Intelligent Data Science Technologies and Applications (IDSTA).
- Jinwei Liu, Richard A Aló, Yohn Jairo Parra Bautista: DeepTrace: Improving Pandemic Health Care by Identifying Disparities and Determinants, "BigComp Conference 2021"
- Yohn Jairo Parra Bautista, Richard A Aló, Jinwei Liu: Behavior Analysis of Pandemic Source Media Communications. "BigComp Conference 2021"

SYNERGISTIC ACTIVITIES
1. Involved in developing an automated Geospatial tool that generates potential Toyota customer hot spot segments by integrating the ESRI Tapestry data that can be used by Toyota.
2. Worked on a project involving visualization of End-use commodity data about Exports of Goods by principal End-use Category in USA category using Geospatial technologies

COLLABORATORS & OTHER AFFILIATIONS: Toyota Haley-Barbour Internship, Spatial Analysis to Model Market Segmentation Using GIS tools, June 2015

GRADUATE ADVISORS AND DOCTORAL STUDIES SPONSORS.
1. Richard Aló, PhD, Dean, College of Science and Technology, Florida A &M University.
2. Sudha Yerramilli, PhD, Interim Director, National Center for Biodefense Communications, Jackson State University[1]SEP
3. David Olson, PhD, James & H.K. Stuart Chancellor's Distinguished Chair Department of Supply Chain Management and Analytics College of Business, University of Nebraska-Lincoln
4. NingNing Wang, PhD, Assistant Professor, Math and Statistical Sciences, Jackson State University
5. Sungbum Hong, PhD, Associate Professor, Computer Science, Jackson State University
6. Kamal Ali, PhD, Chair of Computer Engineering, CSET, Jackson State University