

# Nelson Abdiel Colón Vargas, Ph.D. (Nel Abdiel)

211 River Landing Dr., Apt. 401  
Daniel Island, South Carolina  
29492, USA

+1 (787) 550-8328  
[nel.abdiel@gmail.com](mailto:nel.abdiel@gmail.com)  
<http://www.nelabdiel.com>

EDUCATION	<p><b>George Washington University</b>, Washington, D.C. M.B.A. in Information Systems &amp; Technology Management (<i>Online</i>) In Progress</p> <p><b>The Data Incubator</b>, Washington, D.C. Certificate in Data Science 2016 • <i>Selected among the less than 4% of +1900 applicants.</i></p> <p><b>University of Iowa</b>, Iowa City, Iowa Ph.D. in Mathematics 2016 M.Sc. in Mathematics 2014</p> <p><b>University of Northern Iowa</b>, Cedar Falls, Iowa M.A. in Mathematics 2010</p> <p><b>University of Puerto Rico, Rio Piedras Campus</b>, San Juan, Puerto Rico B.Sc. in Mathematics 2009</p>
EXPERIENCE	<p><b>Soteria - Security Consulting &amp; Data Analytics</b>, Charleston, S.C. Data Scientist - Contract Jun 2016–Present</p> <ul style="list-style-type: none"><li>• Developed predictive models to automate the detection of cyber threats like Phishing, Typosquatting and Network Intrusion with the used of SciKit-Learn, Pandas and NumPy in Python.</li><li>• Built multiple dashboards for internal use using Python, Flask, Django, HTML5/CSS3, Bootstrap, Bokeh and D3.js for analysts to interact with data from OSSEC and other services in a more visual matter.</li><li>• Created software for feature extraction from unstructured data from +30000 websites a day. Both structured and unstructured data is then automatically saved into a Cassandra database.</li><li>• Devised Python pipeline to automate the classification of +30000 websites received per day.</li><li>• Was involved in the decision making and engineering of technologies and infrastructure for ETL.</li><li>• Contributed to the writing of the technical section of an NSF SBIR/STTR grant proposal.</li><li>• Installed and configured OSSEC on servers, and designed and implemented new OSSEC rules to improve efficiency in data collection.</li><li>• Developed probabilistic model to predict the outcome of ping pong matches in the office.</li></ul> <p><b>K2 Data Science</b>, Remote Data Science Curriculum Advisor - Contract (Part-time) Dec 2016–Present</p> <ul style="list-style-type: none"><li>• Designed and developed assignments with detailed solutions for the the following subjects: Data Cleaning and Preprocessing, Exploratory Data Analysis, Regression, Classification, Clustering, Natural Language Processing, Image Processing and Time Series.</li><li>• Helped plan and develop comprehensive learning material to aid students with assignments and to prepare them for their future roles as Data Scientists.</li></ul> <p><b>The Data Incubator</b>, Washington, D.C. Data Scientist - Fellow Jan 2016–Mar 2016</p> <p><b>Selected Projects:</b></p> <ul style="list-style-type: none"><li>• <b>Capstone:</b> <a href="http://iokilos.herokuapp.com">http://iokilos.herokuapp.com</a>. Scraped over a thousand pages from various websites to collect data. Conducted a thorough analysis of the distribution of world records in Weightlifting. Developed an app for strength athletes' training cycles.</li><li>• Conducted open-ended analysis of user behaviors on +9GB of StackOverflow data using Scala and Spark.</li></ul>
SKILLS	<ul style="list-style-type: none"><li>• <i>Stack:</i> Python, PostgreSQL, Spark, Hadoop, D3.js, Theano, Jupyter, Zeppelin, Ambari, Git, L<sup>A</sup>T<sub>E</sub>X.</li><li>• <i>Knowledge of:</i> Machine Learning, Data Visualization, Object-Oriented Programming, Neural Nets.</li><li>• <i>Tools:</i> Pandas, NumPy/SciPy, Scikit-Learn/XGBoost, Keras, PyMC, Bokeh, Flask, HTML5/CSS3.</li><li>• <i>Familiar with:</i> MapReduce, Cassandra, HAWQ, SQLite3, CockroachDB, MongoDB, R, C++, OSSEC.</li><li>• <i>Dominance of:</i> English and Spanish.</li></ul>