AMBASSADE
DE FRANCE
AUX ÉTATS-UNIS

Liberté
Égalité
Fraternité

Service pour la Science et la Technologie
Office for Science and Technology

Office for Science and Technology (OST)
for the United States South-East, Atlanta

Embassy of France in the United States of America

OST South-East Panorama Reports
No. 4, October 27, 2021

AI Research Perspectives Through the US South-East

Research Analyst	**Zachary Hopton**
Deputy Attachés	**Yann Méheust, Benjamin Boumard**
Attaché	**Rami Abi-Akl**

Acknowledgements
Yves Frénot, Counselor for Science and Technology (09/2018 – 08/2021)
Mireille Guyader, Counselor for Science and Technology (09/2021 – present)
Xavier Bressaud, Attaché for Science and Technology, Washington
The OST Team across the United States

Abstract

Research pertaining to artificial intelligence (AI) increasingly transcends boundaries between fields as scientists discover countless applications for intelligent machines everywhere from the operating room to the depths of space. As research progresses, there is a growing pressure to address the unique ethical and legal questions AI brings about. The present report aims to provide a starting point for exploring the implications of AI research by analyzing trends among more than 600 AI researchers in the South-East United States (US). With its diversity of higher-education institutes attracting talent across sectors, the US South-East is a global hub for academic research. Initial trends indicate that AI researchers in the region most commonly work in health and biology, basic AI concepts, and electrical engineering, with some variation between states and territories. The region is a thriving ecosystem overall for AI research, but universities in Georgia and Florida were found to be the primary contributors. The present data highlights the need for increased investment in AI research in diverse fields, encouragement of collaboration to foster the interdisciplinary nature of AI research, and a greater focus on closing the gap between developments in AI and the regulations of such technology.

I. Introduction

Artificial Intelligence

To understand the diverse applications of AI, and the ethical concerns that arise with them, it is important to understand what the term means. At its broadest, AI encompasses any computational system, not necessarily robotic, that acts intelligently. To be more precise, scientists often discuss two types of AI. General (or strong) AI is a theoretical system with all the abilities of a human. Technical (or narrow) AI is a system with one or more specific human abilities. Most researchers agree that general AI does not exist yet.

Prevalent among technical AI, machine learning (ML) models aim to carry out the human ability to learn new information based on experience. In supervised ML, a model is trained with *labeled* data, which it can subsequently use to classify new inputs or make predictions about the future. In unsupervised ML, a model identifies patterns and clusters in *unlabeled* data sets, lessening the need for human input. While these are powerful means of analyzing data, they require vast amounts of input for training, hence the rise in “Big Data” analytics. There are numerous types of models used to implement ML, neural networks being one popular approach. These networks are a type of ML algorithm that mimic the brain’s method of learning by using layers of interconnected nodes (symbolic of neurons) whose relationships change with new input.

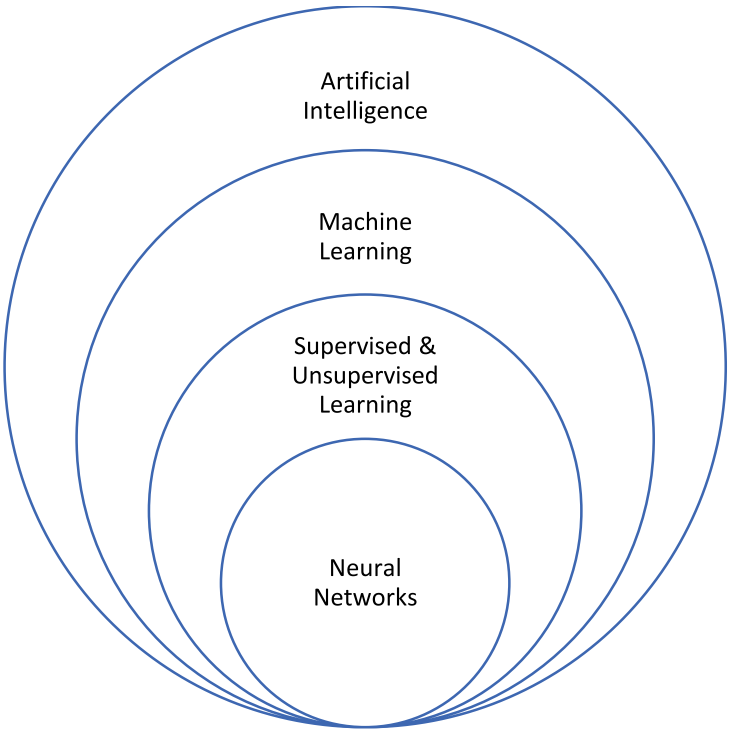


Figure 1. Conceptual relationship between common AI terminology.

Because of AI's unique ability to find patterns in data and make predictions with precision, researchers from virtually all sectors have made inroads in finding new applications for the technology. These applications have yielded cutting-edge solutions for problems in cybersecurity, finance, biotechnology, energy, and countless other fields. For example, researchers with Clemson University's Precision Agriculture team are using AI to develop programs capable of aiding in decisions on pesticide use, crop planting, and irrigation. These programs aim to give data-driven solutions unique to individual farms' environmental conditions, so farmers are able to lower production costs and farm more sustainably. In a vastly different setting, the Biomedical Image Analysis for Image Guided Interventions Laboratory at Vanderbilt University uses ML to automatically segment medical images. This leads to feats such as computer recognition of brain tumors in ultrasound images. The group also uses their automated imaging analyses to directly inform surgeries such as cochlear implant placements.

With the potential to help save lives, the environment, and money, the motivation to research AI is immense. However, as with any technology, it is important to consider AI's ethical and legal implications. For example, the question of legal accountability is far from straightforward when humans are injured due to autonomous or semi-autonomous robotics or vehicles. Economists are uncertain about the effects of AI with respect to potential displacement of human jobs, and there are greater calls to legally ensure data privacy, as exemplified by the European Union's 2018 legislation on the matter. Given that AI algorithms are trained on data from humans, it is also important to consider how they may reinforce biases that already exist in many societies. For example, in 2015, Amazon designed a ML program to automatically identify the applicants who were most likely to succeed in pools of potential hires based on their resumes. Because the model was trained using the resumes of previous Amazon hires, it's performance reflected Amazon's historic bias of hiring more males than females for software engineering positions. That is, the ML algorithm 'learned' to use the applicants' gender as a means of predicting their success at Amazon. This made it more difficult for females to be selected as potentially successful employees. While Amazon discontinued this algorithm, it highlights the potential for AI to magnify pre-existing biases in contexts such as hiring, credit approvals, and healthcare.

Crafting policy that addresses these ethical concerns while exploiting AI's benefits is impossible without perspective on the technology's capabilities. It is with this in mind that the present report seeks to provide a panorama of cutting-edge AI research by summarizing the work hundreds of academic researchers in the South-East US. With knowledge of research trends in AI, collaboration between policymakers and the actors at the forefront of AI research will be more easily accessible.

Academic Research in the South-East US

The South-East US consists of Alabama (AL), Florida (FL), Georgia (GA), Mississippi (MS), North Carolina (NC), Puerto Rico (PR), South Carolina (SC), and Tennessee (TN) (see Figure 2). The region is home to internationally competitive research universities such as Vanderbilt University, the Georgia Institute of Technology, and the University of North Carolina. With countless devoted research centers, initiatives, workshops, and conferences pertaining to artificial intelligence, the South-East US is an optimal starting point for accessing state-of-the-art AI research. The computing abilities of the South-East US are also well-suited for AI research, exemplified by powerful supercomputers such as the world's second fastest at Oak Ridge National Laboratory in Tennessee. Rather than focusing on any one field in isolation, research programs at South-East universities frequently emphasize interdisciplinary approaches to solving problems like sustainable energy production and education technology. This approach is exemplified by schools like University of Central Florida and University of Puerto Rico – Mayagüez, as well as by coordinated, private-public collaboration in units like the Research Triangle Park. Together, this ensures that AI research taking place will be applied to diverse ends.

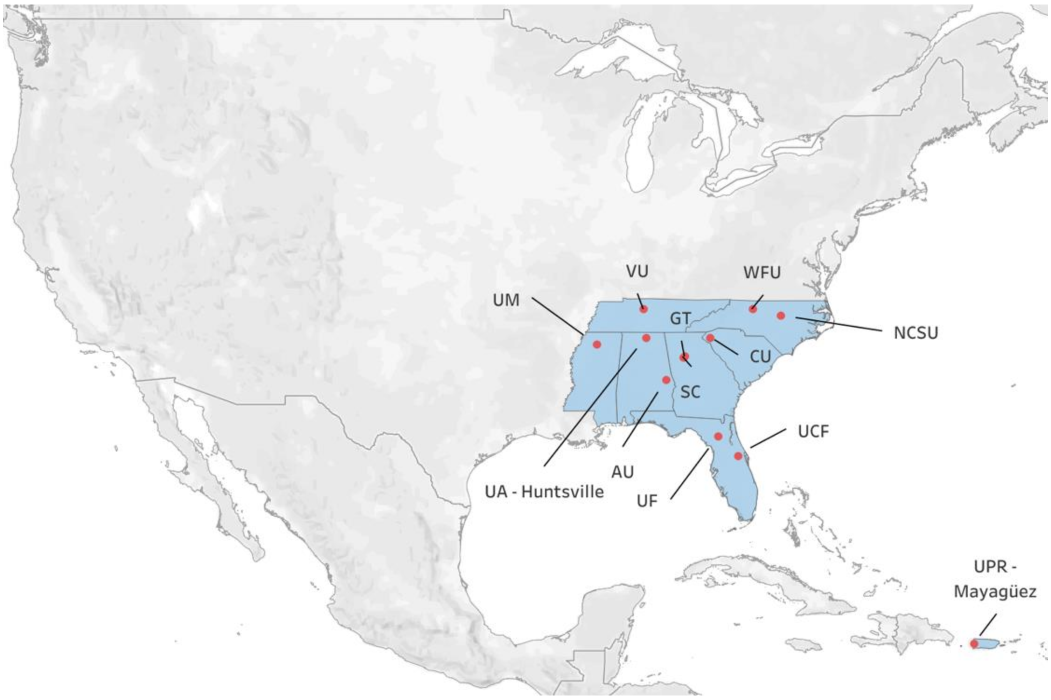


Figure 2. The South-East institutions present in the data analysis. AU = Auburn University, CU = Clemson University, GT = Georgia Institute of Technology, NCSU = North Carolina State University, SC = Spelman College, UA = University of Alabama, UCF = University of Central Florida, UF = University of Florida, UM = University of Mississippi, UPR = Universidad de Puerto Rico, VU = Vanderbilt University, WFU = Wake Forest University.

II. Methodology

The primary goal of the present panorama was to characterize the work being done by AI researchers in the US South-East. The search for AI researchers was carried out within 12 higher-education institutions (See Figure 2). This sample of institutions was chosen to be representative of the broader spectrum of universities in the South-East US, including both large, public state schools (e.g., University of Mississippi) and smaller private schools (e.g., Spelman College and Wake Forest University). Other considerations included geographic distribution (i.e., sampling at least one institute in every state or territory in the region), a balance of institutions with and without medical schools, and a mix of urban and rural settings.

For each institution, individual faculty's biographies and research group profiles (when applicable) were assessed. Criteria for inclusion in the analyses were intentionally broad in order to capture the wide variety of AI researchers. In short, a researcher was included if they had a clear, stated research interest in AI, ML, or neural networks. When published works or research statements indicated an application to AI, researchers focused on algorithms, Bayesian statistics, and precision/"smart" technology were also included.

Upon selection for inclusion, several data points were collected. In an effort to characterize AI research in the South-East US, each researcher was given at least one, but up to three tags describing the field(s) their research encompassed. There were four categories of tags: **Natural Sciences (NS)**, **Engineering Disciplines (ED)**, **Industry Applications (IA)**, and **Computer Science Topics (CST)**. The Natural Sciences refer to physical and biological sciences, as well as quantitative social sciences, while Engineering tags consist of specific engineering sub-fields. Industry Applications encompass domains like agriculture or finance, where research has directly marketable applications. Computer Science Topics refer to a wide variety of technologies such as computer vision and robotics rooted in coding. For a complete list of the 35 tags used, their broader categories, and brief descriptions with example researchers, see Table 1.

To complement data on the number of AI researchers in different states and fields, a measure of research impact was also collected with each researcher. Specifically, each researcher's *h* index was recorded. The *h* index captures the impact and the productivity of a researcher, in that a high *h* index requires both publishing large amounts of papers and publishing papers that are frequently cited. For example, an *h* index of 20 indicates a researcher has published 20 papers that have each been cited at least 20 times. *h* index was obtained from researchers' Google Scholar profiles or from the search engine Scopus. No *h* index could be found for 14 faculty.

Tag (Tag Category)	Example Problems	Sample Lab
Agriculture (IA)	Automatic Quality Control; Predictive Crop Care	The Advanced Ag Lab
Architecture (IA)	Automated Design; Modeling	Clemson University Institute for Intelligent Materials
Atmospheric Sciences (NS)	Climate Forecasting; Air Quality Monitoring	The Luthy Lab
Autonomous Vehicles (CST)	Self-Driving / Self-Flying Vehicles	The AutoAI Lab
Basic (CST)	Knowledge Theory; Optimization; Algorithms	Statistical Machine Learning Lab
Chemistry (NS)	Reaction Prediction; Molecular Simulation	The McDaniel Research Group
Communications (IA)	Signal Processing; Multi-Object Network Coordination	Communication & Wireless Network Lab
Computer Vision (CST)	Visual Perception; Pattern Recognition	The Eye-Team
Cybersecurity (CST)	Cyber-Physical System Security; Fraud Detection	The Embedded Systems Lab
Data Analytics (CST)	Big Data Analytics; ML for Analysis of Diverse Data Sets	The Vanderbilt Data Science Institute
Defense (IA)	Military Training; Hazard Detection	The Mixed Emerging Technology Integration Lab
Digital Humanities (NS)	Digital Archives; Historic Document Analyses	The Center for Digital Humanities at Vanderbilt
Education (IA)	Online Learning; Personalized Curricula	The Help through Intelligent Support Lab
Energy (IA)	Power Management; Sustainable Energy Development	UF Power Lab
Engineering, Aerospace (ED)	Automated Space Exploration; Automated Landings	Dynamics & Control Systems Lab
Engineering, Chemical (ED)	Drug Discovery; Chemical Manufacturing Optimization	Data-Driven Process Systems Engineering Lab
Engineering, Civil (ED)	Automated Construction; Modeling Infrastructure	The Network Dynamics Lab
Engineering, Electrical (ED)	Smart Energy Grids; Internet of Things	The GREEN Lab
Engineering, General (ED)	Highly Interdisciplinary Engineering Problems	The GETCOT Lab
Engineering, Industrial (ED)	Supply Chain Management; Manufacturing Systems	The AIMS Lab
Engineering, Materials (ED)	Materials Informatics; Corrosion	The Russo Lab
Engineering, Mechanical (ED)	Nanoengineering; Engineering Design	MINDS Lab
Engineering, Nuclear (ED)	Radiation Detection; Medical Physics	LANNS Lab
Entertainment / Gaming (IA)	Social Media Analytics; Virtual Character Design	The Arnav Lab
Environmental (NS)	Quantitative Geology; Environmental Monitoring	Bio-Optical Oceanography Lab
Ethics (CST)	Biased Algorithms; Data Privacy	The AI Things Lab
Finance (IA)	Stock Market Analytics; Loan & Credit approval	Financial Services Innovation Lab
Human-Computer Interaction (CST)	User Experience; Assistive Technology	Mobile & Invasive Computing Laboratory

Health / Biology (NS)	Imaging Analyses; Surgical Robotics	The Computational Biology Lab
Math (NS)	Simulation & Modeling; Applied Statistics	Research Group in Mathematical Neuroscience
Natural Language Processing (CST)	Speech Recognition; Text Parsing; Search Engines	Social & Language Technologies Lab
Physics / Astronomy (NS)	Particle Discovery; Planetary Detection	Nonlinear AI Lab
Robotics (CST)	Autonomous Agents; Motion Planning	The Assistive Robotics Lab
Transportation (IA)	Travel Demand Modeling; Traffic Networks	The Garrow Group
Urban Analytics (IA)	Sustainable Cities; Large-Scale Networks; Urban Mobility	The Scope Lab

Table 1. Tag descriptions and sample researchers. IA = Industry Applications, NS = Natural Sciences, CST = Computer Science Topics, ED = Engineering Disciplines.

III. Trends in AI Research through the South-East US

A total of 648 AI researchers were included for analysis. A testimony to the highly interdisciplinary nature of AI research in the South-East US, almost twice as many tags (1,189) were used to classify these researchers. Using the data from the 12 selected universities, it is possible to visualize general ideas about the overall state of AI research in the South-East US, as well as state-to-state trends.

Results

High in both impact (as indicated by *h* index) and volume of research (indicated by the number of AI researchers found in each state), the universities sampled in Georgia and Florida are leading the region in AI research (see Figure 3). Notably, North Carolina has a relatively high volume of AI research, and Tennessee's AI research is exceptionally high impact. Mississippi and Puerto Rico had both the lowest volume of researchers and the lowest overall impact. While controlling for the number of universities sampled in each state renders their *h* indices more comparable, they still rank low with respect to volume. (see Figure 3b).

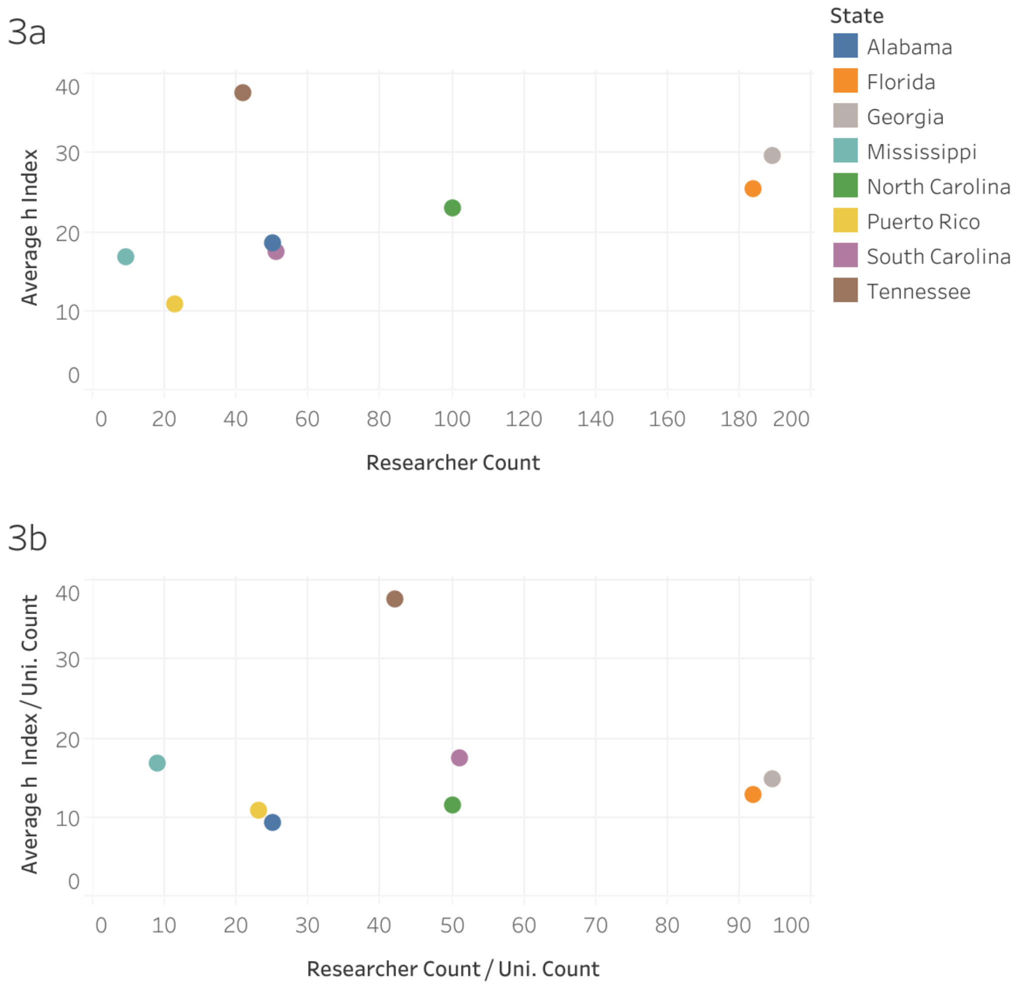


Figure 3. States and territories plotted with respect to the impact and volume of their AI research. 3a represents raw data. Values in 3b control for the number of universities sampled in each state.

Across the South-East US, the data indicated that there are vast amounts of researchers applying artificial intelligence to problems in health and biology, and exploring the basic science underlying AI in order to optimize it's use (see Figure 4). Among the Engineering Disciplines, the data indicate that the greatest volume of AI research is being conducted in the context of electrical engineering. Energy research was the most frequently used tag among the Industrial Applications class, with 34 groups labeled accordingly. AI research was highly prevalent in multidisciplinary Computer Science Topics such as data analytics, autonomous vehicles, and computer vision. In general, very few researchers were using AI for digital humanities, nuclear engineering, and defense. It should be noted that this does not indicate a complete lack of research in these sectors through the South-East US. Rather, it means that only among the sampled academic institutes, few researchers are using artificial intelligence to solve problems in these fields.

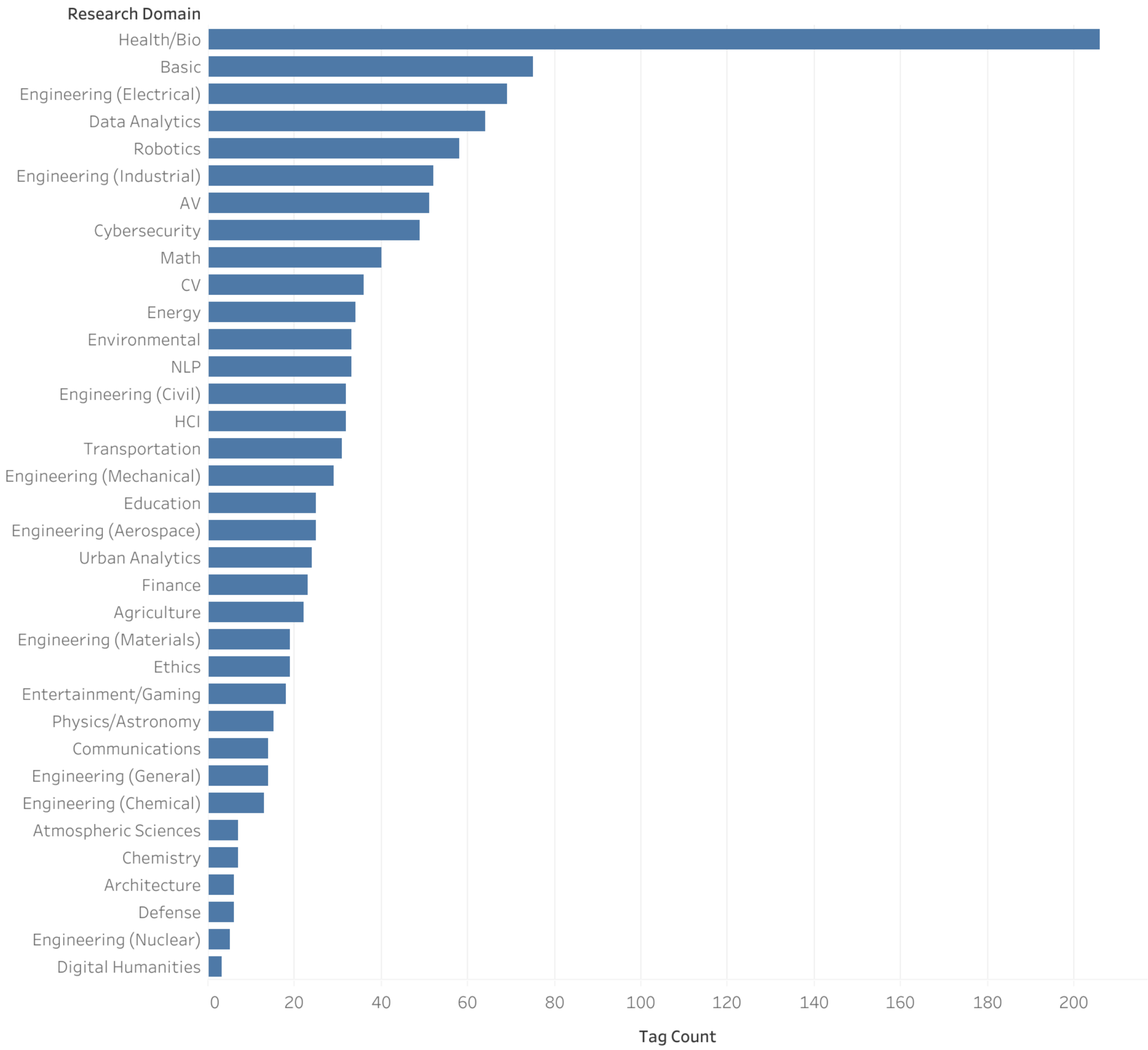


Figure 4. Bar length indicates the number of times each tag was used. Values are representative of all eight states and territories. AV = Autonomous Vehicles; CV = Computer Vision; NLP = Natural Language Processing; HCI = Human-Computer Interaction.

Trends within the states and territories of the South-East US largely follow those of the overall region, with some notable exceptions (see Figure 5). For example, Georgia shows particularly strong performance with respect to AI research in industrial, mechanical, and aerospace engineering, as well as robotics. With eight researchers, North Carolina appears to have a relatively robust volume of agricultural AI research for the region. Although few researchers throughout the whole US South-East were applying AI to education, Florida and North Carolina excelled in this sector. The top three sectors for each state are as follows: Alabama (Basic, Data Analytics, Electrical Engineering); Florida (Health/Biology, Basic, Electrical Engineering); Georgia (Health/Biology, Industrial Engineering, Robotics); Mississippi (Data Analytics, Environmental, Health/Biology); North Carolina (Health/Biology, Basic, Math); Puerto

Rico (Electrical Engineering, six-way tie for the next positions; see Figure 5); South Carolina (Health/Biology, Autonomous Vehicles, Electrical Engineering); Tennessee (Health/Biology, Robotics, five-way tie for next positions; see Figure 5). In order to emphasize absolute performance within each state and territory, Figure 5 contains raw data that does not control for the number of universities sampled in each state.

	State							
Research Domain	AL	FL	GA	MS	NC	PR	SC	TN
Agriculture	3	6	0	0	8	3	2	0
Architecture	0	1	4	0	0	0	1	0
Atmospheric Sciences	3	0	1	0	2	0	1	0
AV	6	17	11	0	8	0	8	1
Basic	11	27	11	1	12	3	7	3
Chemistry	0	1	2	0	2	0	0	2
Communications	1	8	3	0	1	0	1	0
CV	4	12	11	1	2	0	3	3
Cybersecurity	4	15	17	0	7	1	5	0
Data Analytics	8	21	17	3	10	1	1	3
Defense	1	2	1	0	1	0	0	1
Digital Humanities	0	1	0	0	0	0	1	1
Education	2	9	3	0	8	1	1	1
Energy	0	11	14	1	3	0	2	3
Engineering (Aerospace)	2	5	16	0	2	0	0	0
Engineering (Chemical)	4	1	8	0	0	0	0	0
Engineering (Civil)	0	10	9	1	4	2	4	2
Engineering (Electrical)	10	27	11	0	7	4	8	2
Engineering (General)	0	5	4	0	2	1	1	1
Engineering (Industrial)	1	6	37	0	5	3	0	0
Engineering (Materials)	0	5	9	0	2	3	0	0
Engineering (Mechanical)	1	4	20	1	1	1	0	1
Engineering (Nuclear)	0	1	4	0	0	0	0	0
Entertainment/Gaming	0	3	6	0	4	1	4	0
Environmental	1	8	6	2	10	3	1	2
Ethics	3	4	6	0	3	0	3	0
Finance	1	6	13	0	3	0	0	0
HCI	4	15	4	0	6	0	3	0
Health/Bio	8	57	60	2	31	3	16	29
Math	3	6	14	1	12	2	1	1
NLP	4	11	9	0	6	0	1	2
Physics/Astronomy	1	6	3	0	1	3	1	0
Robotics	3	16	21	1	6	1	6	4
Transportation	1	7	9	1	2	1	7	3
Urban Analytics	1	11	7	0	2	0	2	1
Grand Total	91	345	371	15	173	37	91	66

Tag Count
0 60

Figure 5. Values in the table represent the number of times each tag was used with respect to each state and territory in the South-East US. Darker red tabs indicate higher values, grey represents zero.

Research Impact

The average *h* indices for each tag yield somewhat contrasting results relative to analyses of research volume (see Figure 6; 14 researchers with no available *h* index were excluded from this analysis). The three highest impact fields, for example, were found to be defense, transportation, and physics/astronomy.

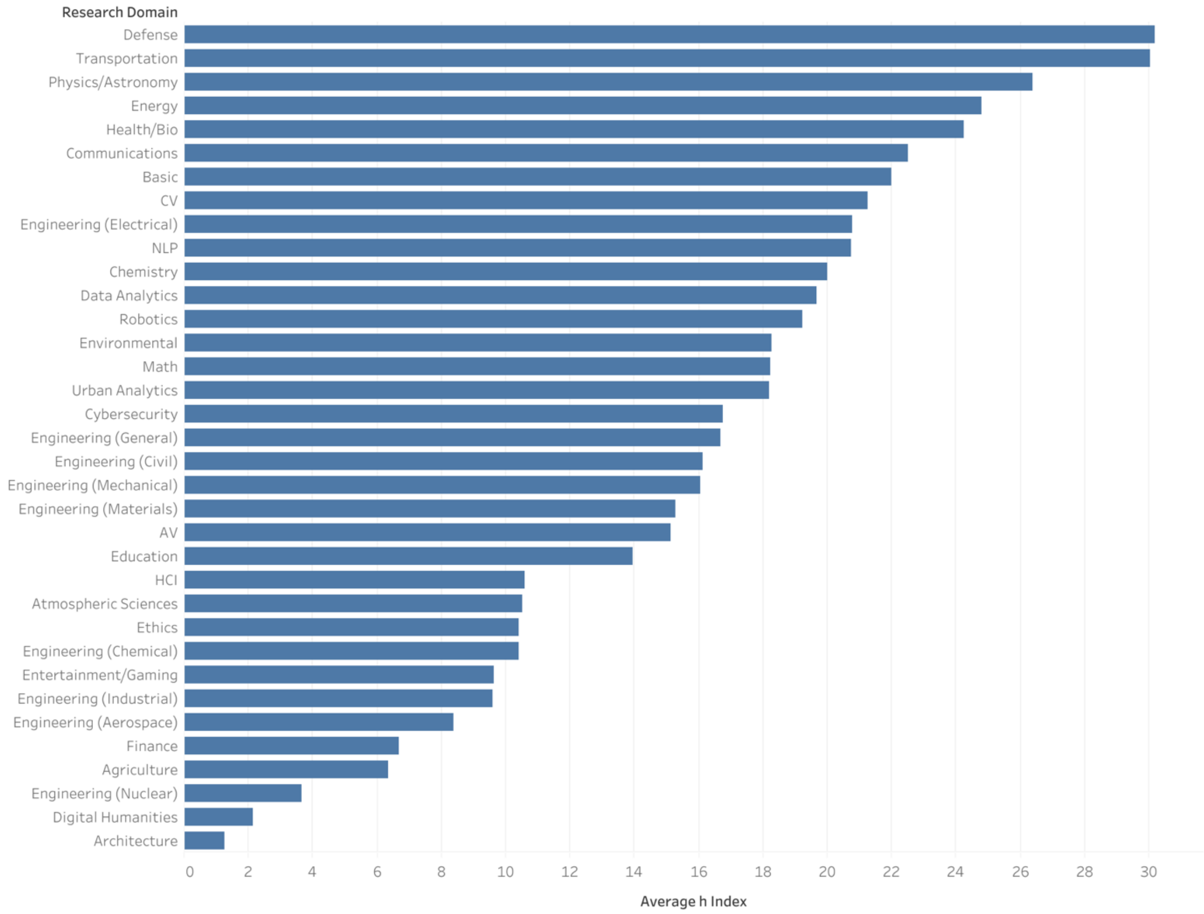


Figure 6. Bar length indicates the *h* index for all researchers in each tag. Values represent the average across all eight states and territories.

As stated previously, some of the fields with the greatest volume of AI research in the South-East US are health/biology, electrical engineering, and data analytics. The visualizations in Figure 7, which overlay research impact and volume metrics, show that states' average *h* indices for these fields were often between 20 and 30. Comparing these two metrics side-by-side adds more nuance to the trends in Figure 6. For example, Figure 7 clearly demonstrates that while AI research in defense was relatively high impact, few researchers in the sampled universities focus on this field. The same can be seen for AI research in physics/astronomy. Thus, Figure 7 highlights the importance of considering that exceptionally high *h* indices may be the result of relatively small samples of researchers that are particularly vulnerable to outliers.

The selected maps in Figure 7 were chosen for brevity, and only consist of those referred to in the text. However, maps for all 35 tags, broken into the categories of Industrial Applications, Natural Sciences, Engineering Disciplines, and Computer Science Topics, can be found in Appendix 1.

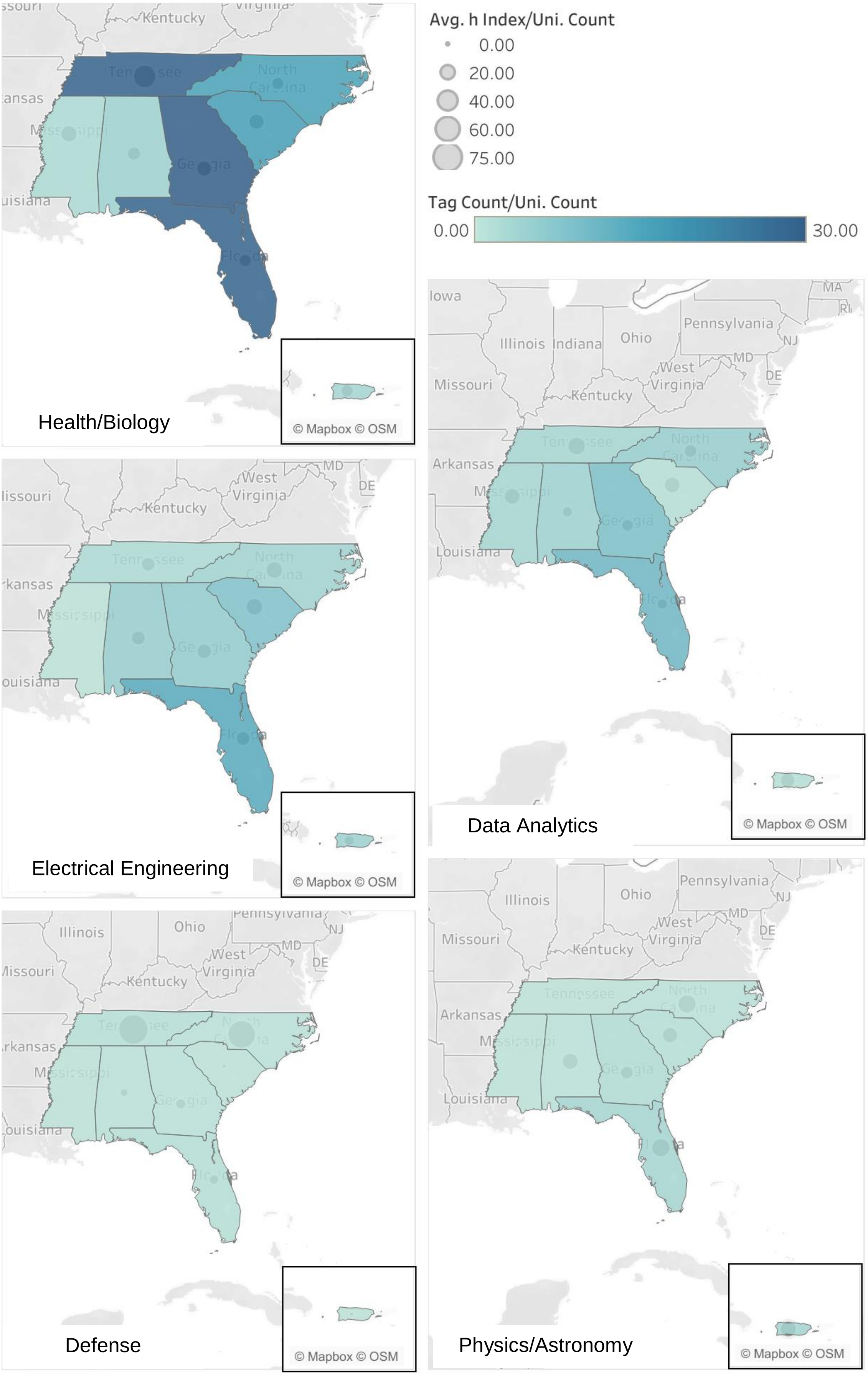

Avg. h Index/Uni. Count
0.00
20.00
40.00
60.00
75.00
Tag Count/Uni. Count
0.00
30.00
Health/Biology
Data Analytics
Electrical Engineering
Defense
Physics/Astronomy
© Mapbox © OSM

Figure 7. Selected research impact and density maps. The states' colors indicate the research volume (Tag Count/University Count) in each state for a given tag, where light blue is lower volume and dark blue is higher. The size of the light grey circles represents the research impact (Average *h* Index/University Count) for a given tag in each state. Here, larger circles indicate higher *h* indices. All values control for differences in the number of universities sampled in each state.

IV. Conclusion

The immense volume of AI research taking place in the US South-East is a testament to its efficacy in a wide array of contexts. Thus, the present report highlights the need for greater investments in training and funding AI researchers in academia. The South-East's three most widely researched fields in AI provide logical starting points. Considering the ongoing virulence of COVID-19, there is an unprecedented need for investments in the data-driven conclusions AI can bring to medicine, such as public health decision aid, precision medicine, and drug discovery. Moreover, the conclusions from the present report emphasize the importance of basic AI research. Given its role in optimizing technology rooted in AI, as well as in novel discoveries, basic research can be considered a bottleneck in the research and development process. More investments made in basic research lead to a wider bottleneck, opening up entirely new means of efficiently applying AI to diverse problems. The third most widely researched field of AI in the South-East US, electrical engineering, is the basis for countless developments in sustainable energy, automated vehicles, robotics, and cybersecurity. Investing in the use of AI in electrical engineering therefore has profound implications in the invention of new technologies in the future.

With over 600 AI researchers documented in just 12 of the many universities of the South-East US, the region has ample potential for collaboration with both domestic and foreign entities. While the researchers sampled here were from academia, their websites and biographies frequently detailed patents, start-ups, and consulting roles for private and public entities. With this in mind, there is room for collaboration with organizations of all kinds, including universities, but also technology companies, entrepreneurs, think-tanks, and policymakers. Generally, the results here indicate that most of the region is a friendly environment for AI research. However, universities in Georgia and Florida stand out as particularly equipped for collaboration in this domain. Of course, this necessarily depends on the entities' specific goals and focuses, but it is our hope that the trends presented here will provide a meaningful starting point in searching for optimal partnerships. As AI research is so often interdisciplinary, collaboration provides a rich means for exchanging perspectives between researchers, technology developers, and policymakers.

Although a snapshot of just one region of the world, the present report is a clear indicator that AI research is a robust, accelerating field. With labs already focusing on neural-computer interfaces, beyond-6G wireless networks, and advanced speech and facial recognition software, the technology harnessing AI appears to be outpacing the policy and laws monitoring it. This puts a spotlight on the need to create public and private regulations addressing potential ethical and legal concerns arising from AI. It is critical that policymakers connect with academic researchers at the forefront of AI development in order to better understand the implications of their research. In turn, such connections allow for more up-to-date, consistent regulations and laws that maximize the good of AI while minimizing any harm it is capable of doing.

Limitations & Future Directions

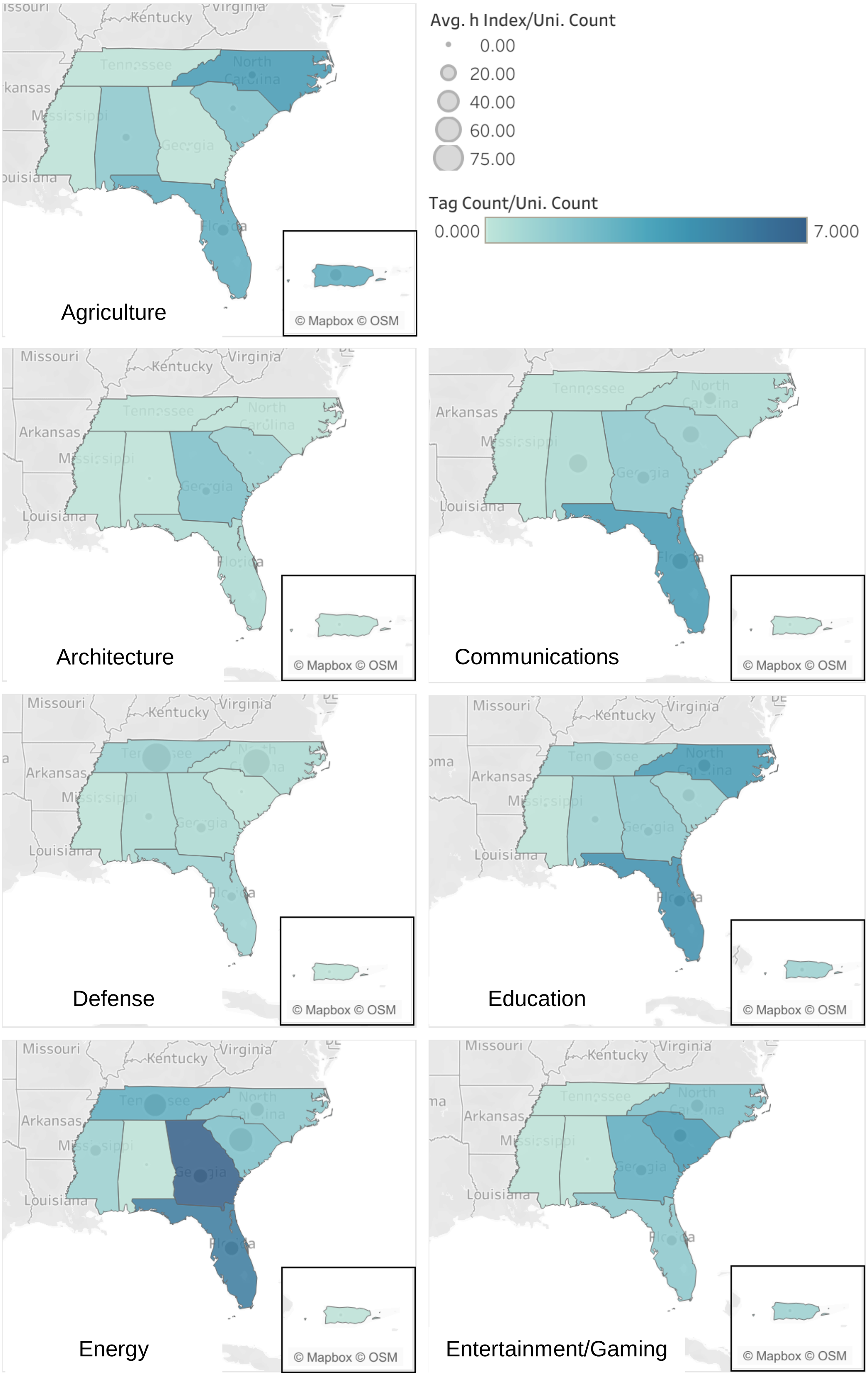
One of the principal limitations of this work comes from generalizability. The conclusions here are an important initial step in understanding the AI ecosystem of the region, but several major universities (some examples being University of Tennessee – Knoxville, Emory University, Duke University, University of North Carolina – Chapel Hill, and University of South Carolina) should be added to the analyses in the future. Considering the hospitals, national research labs, and urban environments associated with many of these universities, their inclusion could stand to change the trends seen here, and certainly warrants further work. That being said, the diverse array of universities included (see Section II, Method) was meant to provide more generalizable data than a selection of only private or liberal arts schools, for example.

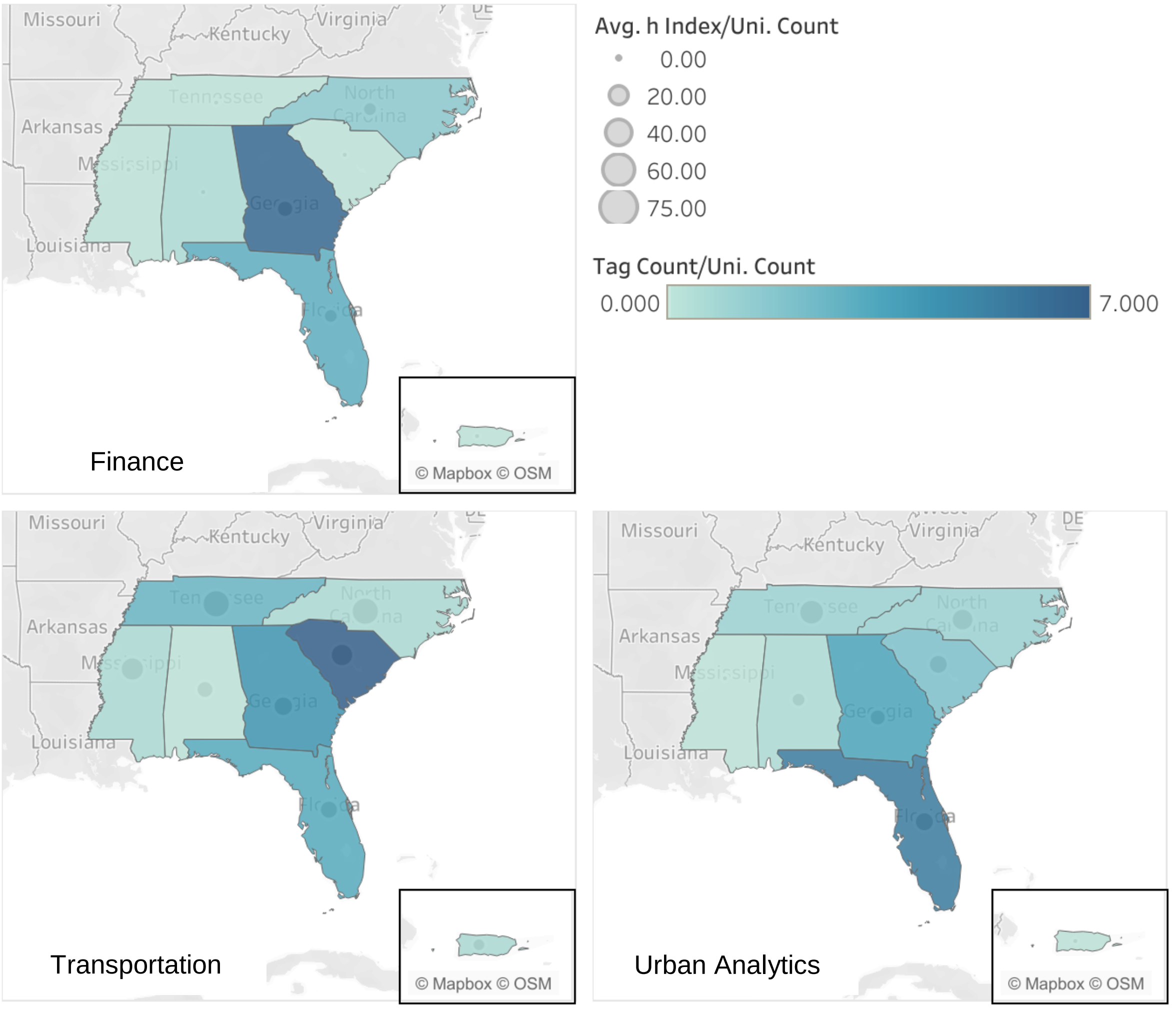
Another limitation was in the process of categorizing the researchers' fields. Many of the tags overlapped to high degrees (e.g., robotics and mechanical engineering; energy and electrical engineering; basic research and math). While researchers' work was categorized based on the emphases in their own biographies or published works, one could still question the efficacy of classifying research as interdisciplinary as that seen here. Allowing for up to three tags partially addressed this concern. Another possible remedy would be creating more specific tags. Although adding more tags may make data less easily interpretable, in some cases it may be warranted. For example, given the vast number of health and biology researchers, an interesting route for future work would be to make health/biology a category of tags, similar to Computer Science Topics or Engineering Disciplines. The 200+ AI researchers that are currently tagged health/biology could then be classified with more specific tags (e.g., medical imaging, medical robotics/prosthetics, molecular biology, ecology, etc.). By any means, a finer analysis of AI research in certain domains may help better characterize the work being done in the South-East US.

Finally, interviewing AI researchers in the South-East US would provide a qualitative complement to the data presented here. Potential lines of questioning include perspectives on international collaboration, the nature of the relationship between academia and the tech industry, and acquiring first-hand perspective on the ethical considerations in AI.

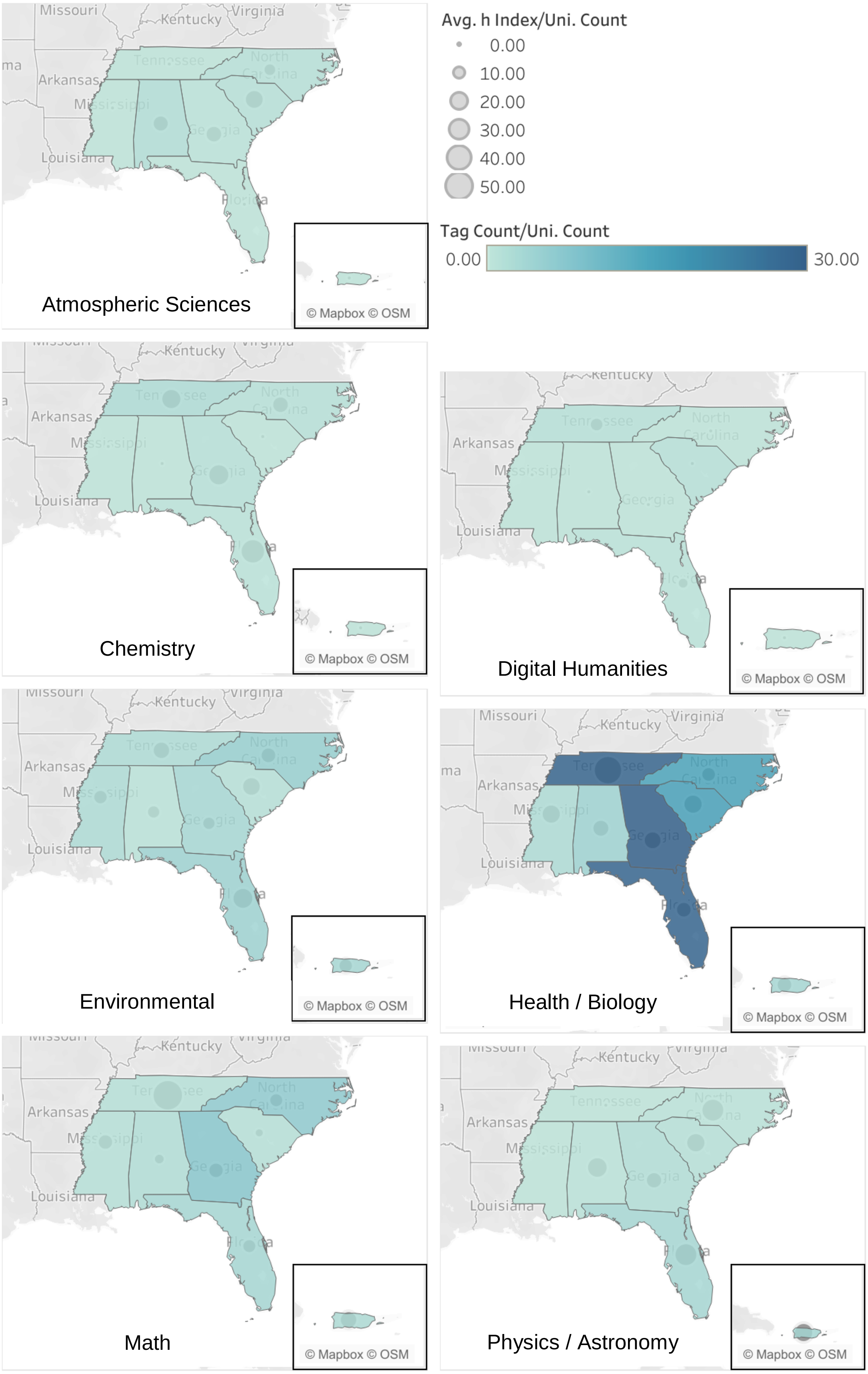
Appendix 1

Industry Applications

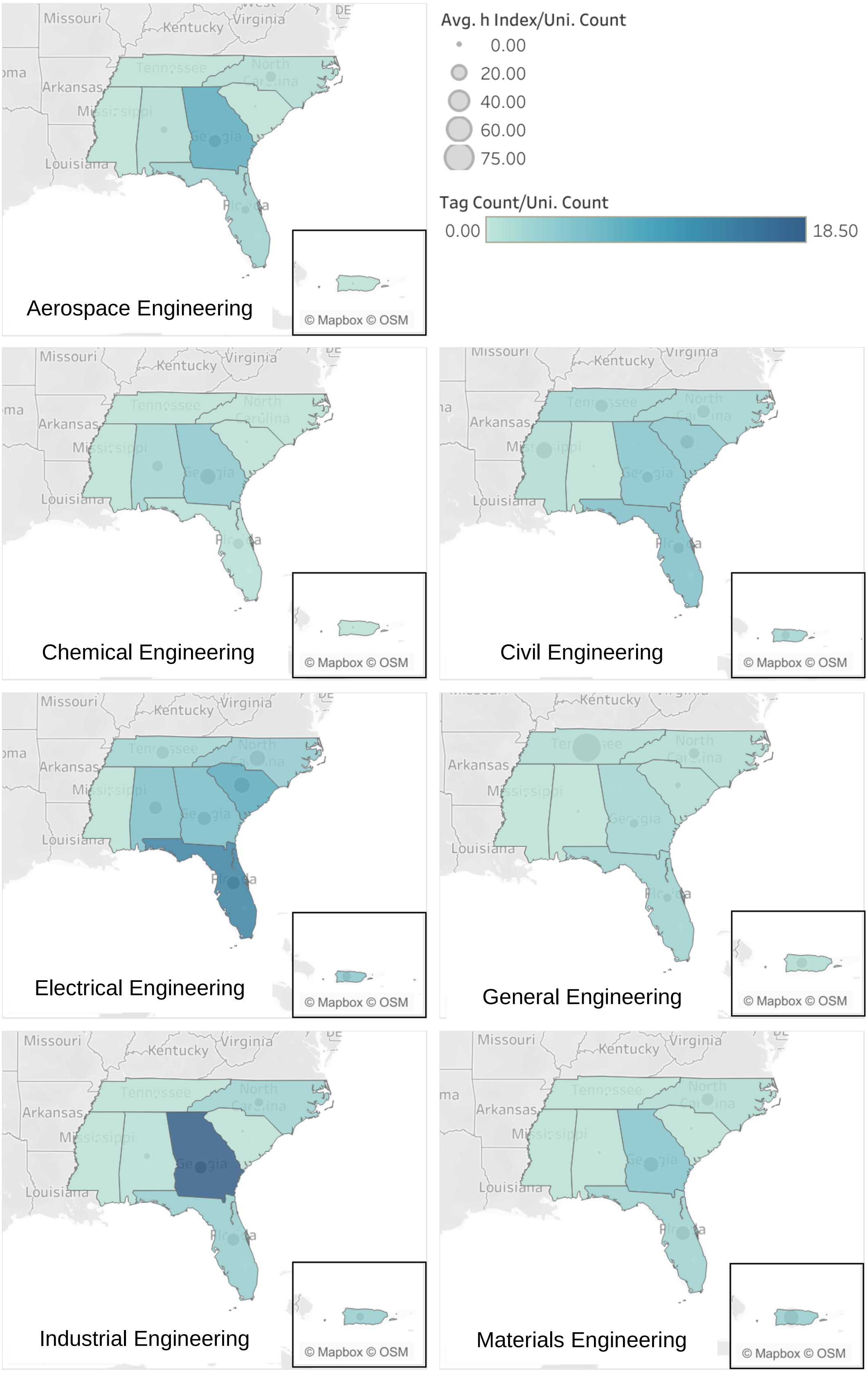


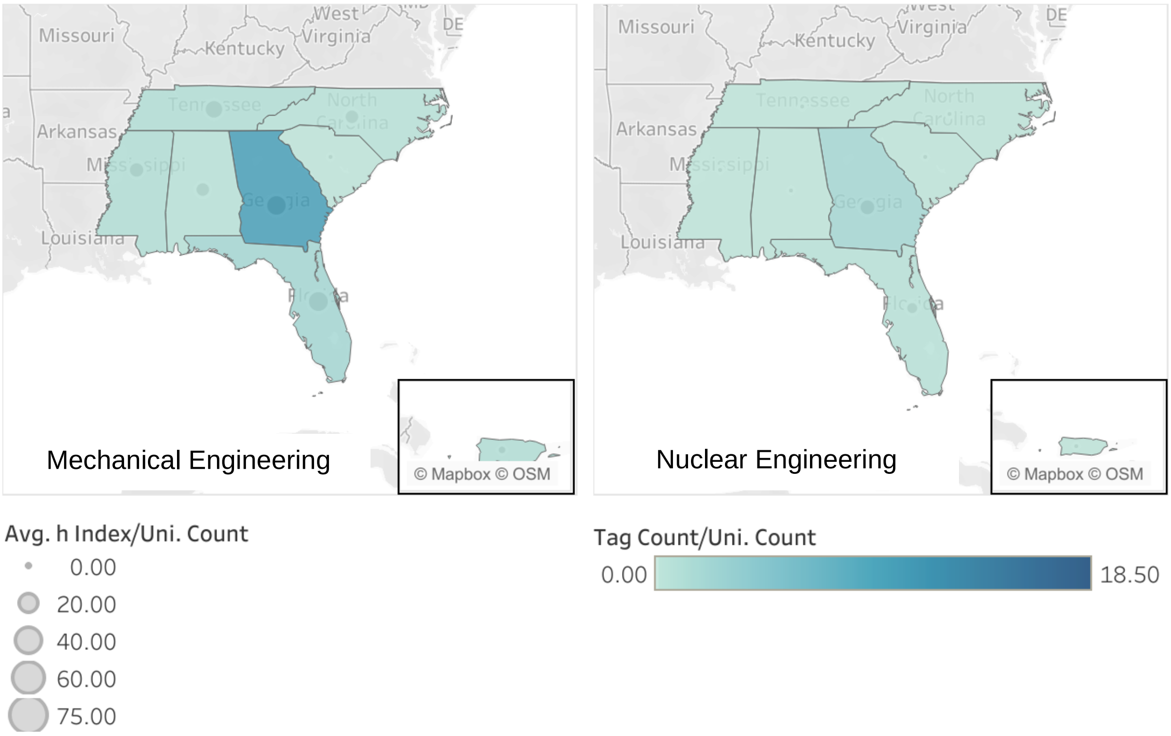

Avg. h Index/Uni. Count
0.00
20.00
40.00
60.00
75.00
Tag Count/Uni. Count
0.000
7.000
Finance
Transportation
Urban Analytics
Missouri
Kentucky
Virginia
Tennessee
North Carolina
Arkansas
Mississippi
Georgia
Louisiana
Florida
© Mapbox © OSM

Natural Sciences

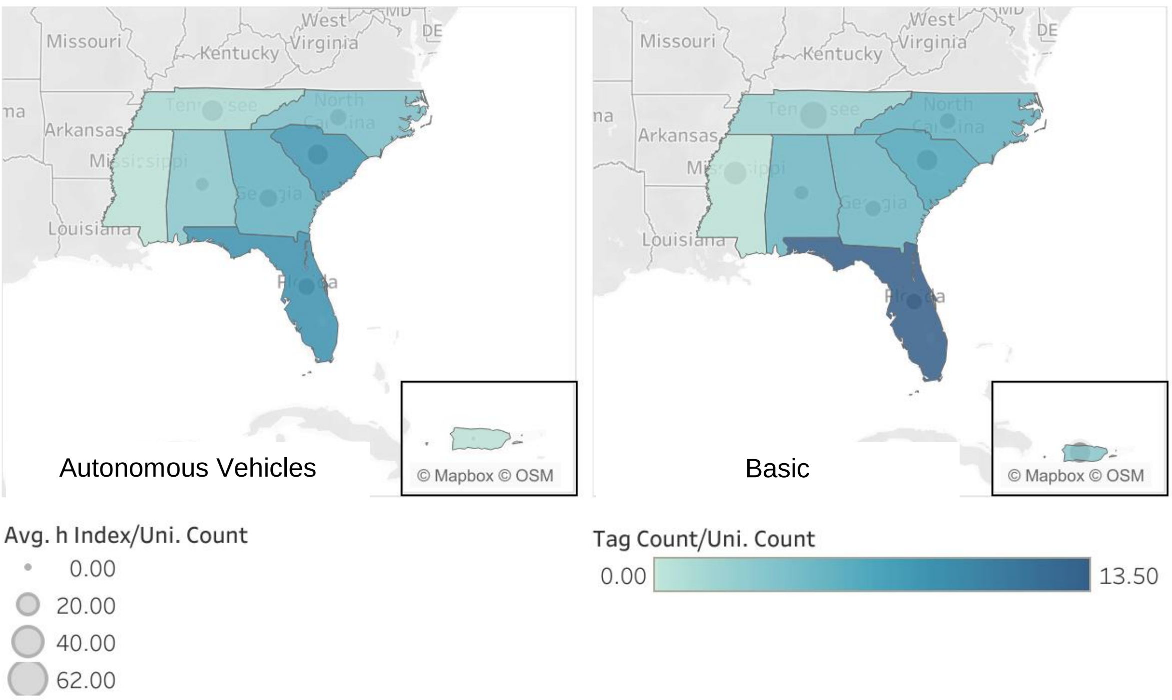


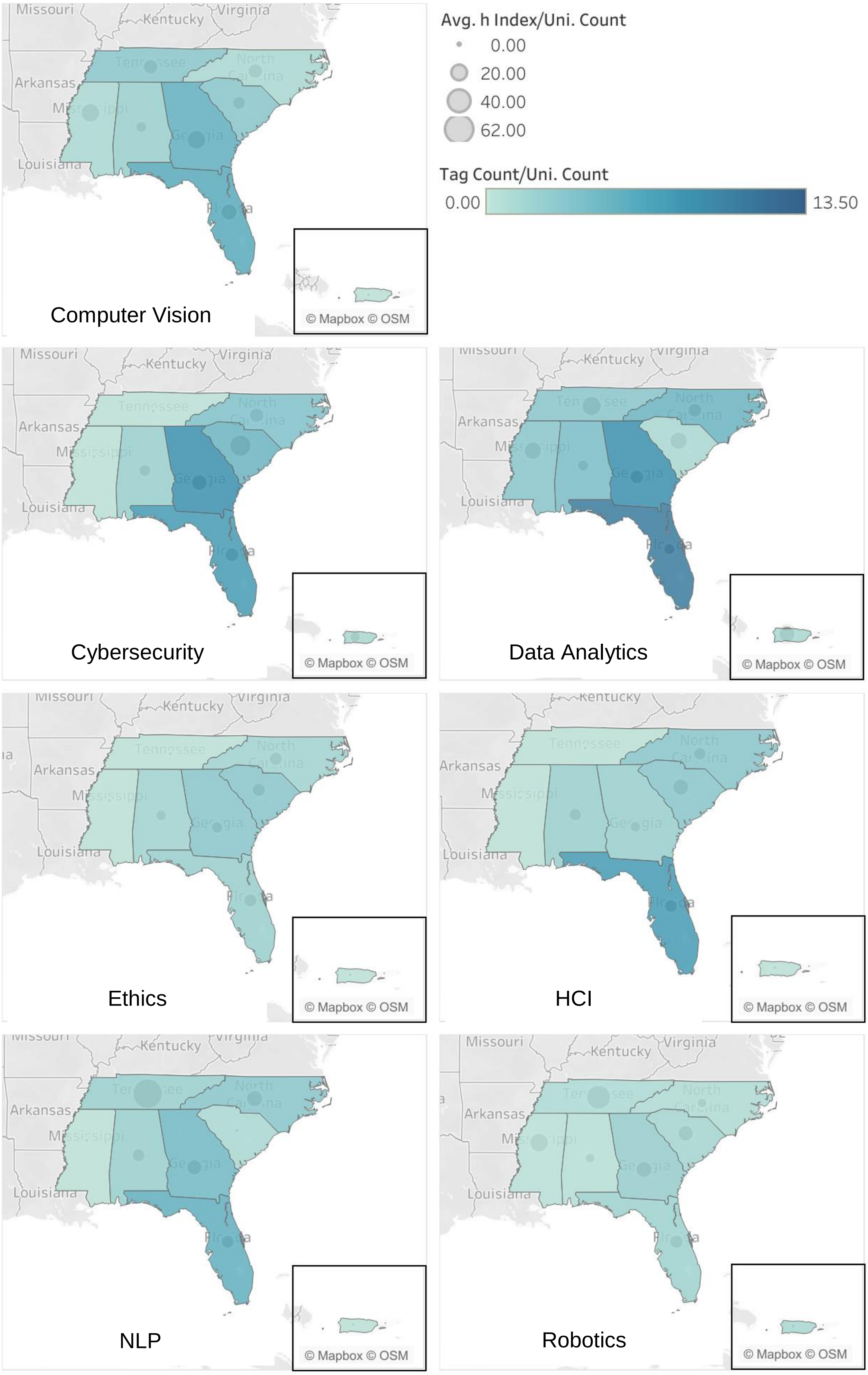
Engineering Disciplines





Computer Science Topics




Avg. h Index/Uni. Count
0.00
20.00
40.00
62.00
Tag Count/Uni. Count
0.00
13.50
Computer Vision
Cybersecurity
Data Analytics
Ethics
HCI
NLP
Robotics
Missouri
Kentucky
Virginia
Arkansas
Tennessee
North Carolina
Mississippi
Georgia
Louisiana
Florida
© Mapbox © OSM