Understanding the Impact of Data Privacy Regulations on Software and Its Stakeholders

Commonwealth Cyber Initiative (CCI) Fall 2022 Research Call:
Securing Interactions between Humans and Machines

Chris Brown (PI)
Computer Science
Virginia Tech
dcbrown@vt.edu

Aaron F. Brantly (Co-PI)
Political Science
Virginia Tech
abrantly@edu

1 Introduction

As societal dependence on technology grows, software is increasingly collecting data on users. This data is collected and sold by firms to provide potential benefits to user experiences, such as hyper-personalized marketing [29] and behavioral analysis for UX and UI optimization [21]. However, the exploitation of user data constitutes a fundamental human security challenge. For example, Facebook providing private intelligence company Cambridge Analytica unauthorized access to data harvested from personally identifiable information for millions of unknowing users [24]. Data resulting from users' interactions in an increasingly digitized world, and data privacy policies, regulations, and laws are critical for safeguarding the humans behind the data.

Data privacy regulations are written contractual or legal frameworks that aim to protect users' data from unethical, unauthorized, or illegal usage by tech companies, third-party firms, and governments. Data privacy issues have been of concern since the early days of computers, but have grown in importance as new technologies enable novel means to collect and process data generated by users. Policies pertaining to data privacy are increasing across jurisdictions, such as the *General Data Protection Regulation* (GDPR) in the EU, the *California Consumer Privacy Act* (CCPA) and *Consumer Data Privacy Act* (CDPA) in the US, and *Lei Geral de Proteção de Dados* in Brazil. Prior work has explored the GDPR and its effect on data organization [26] and requirements [6, 25], as well as automated compliance [27, 34] and benchmarking [32]. However, little is known about the impact of these laws on software and its stakeholders. To that end, the proposed work will explore how adherence to data privacy regulations affects code, developers, and users.

1.1 Project Objectives

The principal objective of this project is to understand the impact of implementing data privacy regulations and laws on software products, developers, and end users. The proposed work consists of a series of related studies that examine the impact of implementing regulations that affect data privacy in open source software at the code level, within developer communities, and among the end users of products. To study the impact of data privacy standards on code, developers, and users, we will analyze code modifications and stakeholder (both developer and user) perceptions related to the GDPR and CCPA. Thus, the proposed work aims to answer the following research questions:

RQ1: What impact do data privacy regulations have on software?

Software is a set of instructions used to operate a computer and execute tasks to manipulate data [30]. To understand the impact of privacy regulations on software, we will longitudinally examine changes to code bases. This provides us with a before and after snapshot of programs at the software level, and gives a proxy measure of the amount of effort required to implement changes for privacy regulations. We will analyze code changes by mining repositories on GitHub [1], an online platform containing millions of software projects with code contributions from developers [23]. We anticipate changes incorporating data privacy will be nonintuitive and require substantial edits to code.

RQ2: What impact do data privacy regulations have on software developers?
While RQ1 addresses the impact of privacy regulations at the code level, that code is created and maintained by software developers–who are responsible for making adjustments to code to implement data privacy regulations. Thus, RQ2 seeks to understand the concerns and challenges developers face. To investigate developer perceptions, we will analyze comments made in online programming communities such as GitHub and StackOverflow [2], a popular Q&A website for programmers. These comments constitute both qualitative and quantitative (iterated sentiment or solutions) data that provide insights into how privacy regulations affect software developers. We hypothesize developers face difficulties seeking out solutions to the complex challenges presented by data privacy.

RQ3: What impact do data privacy regulations have on software users?
Data privacy regulations also modify user experiences in software applications. Consumers may not understand technical changes, but can recognize differences in their user experience–such as increased cookie acceptance pop-ups during web browsing or emails for agreeing to updates to terms and conditions. To understand how these changes impact user experiences and are perceived by users, we will use the Twitter archive [4] to uncover relevant social media posts related to data privacy regulations. As with changes in sentiment and solutions for developers, Twitter data will provide a longitudinal look at perceptions of the implementation of privacy laws from the user's perspective. We speculate users will have mixed perceptions of data privacy regulations, praising increased security of their information but complaining about changes to user experiences.

1.2 Intellectual Merit

The proposed research will advance knowledge in computing research areas related to the development, maintenance, deployment, and privatization of software. This project will also advance knowledge within the social sciences, in particular, public policy studies pertaining to the creation and implementation of regulations and laws. The research outcomes will enhance our understanding of the following:

- the impact of introducing new data privacy requirements on software developers, code bases, and development processes;
- the challenges software developers face and questions they ask when implementing changed data privacy standards; and
- the impact of these modifications in software applications on user experiences.

Moreover, we will provide implications and guidelines for improving how policies related to data privacy are instituted by characterizing responses from constituents to facilitate easier transitions for regulation adherence for software developers and users.

Relevance to CCI. This project will advance the state of the art in cybersecurity by conducting cross-disciplinary research to understand how privacy regulations for securing users' data impact software and the stakeholders involved. Cybersecurity incorporates "resources, processes, and structures" to protect and defend digital spaces from misaligned and unethical activities [22]. Data privacy laws are formal structures intended to enhance human interactions with machines through software by protecting users' information, preventing unauthorized usage, and providing them control over their data.

The proposed work will contribute empirical analyses seeking to understand the impact of data privacy laws on code, developers, and users. Further, we will motivate innovation in the workforce by providing recommendations to improve the rollout of data privacy-related policies and insight into tools and resources to ease regulation compliance.

Cross-Disciplinary Aspects. The cross-disciplinary aspects of this work include: *software engineering* (SE), to analyze code and developer effects; *human-computer interaction* (HCI), to study user experiences; *cybersecurity* (CS) and *privacy*, to safeguard user data; and *political science* and *policy*, to understand data privacy regulations and legislation.

1.3 PIs' Expertise

PI Brown has researched ways to improve developer behavior and decision-making by building and evaluating automated bots and tools [15, 19, 33], conducting user studies [7, 14, 16], exploring interdisciplinary concepts, such as *nudge theory* [17, 20], and performing empirical analyses of tools and processes on GitHub [18, 35, 36]. PI Brown will lead the efforts to investigate the impact of data privacy regulations on SE by analyzing code modifications (RQ1) and developer perceptions (RQ2) on GitHub and Stack Overflow.

Co-PI Brantly is the director of the Tech4Humanity Lab at Virginia Tech and has written extensively on cybersecurity [31, 10, 9, 13, 12]. Brantly's research has focused on the impact of intrusions into user privacy with an emphasis on violations of user privacy among human rights and democracy activist communities [8]. Brantly has experience with natural language processing and big data collection and analysis on social media platforms [11]. Brantly also has extensive experience as a senior program officer for ICT Innovation at the National Democratic Institute focused on developing technical solutions for privacy and cybersecurity in developing democracies. Co-PI Brantly will provide insight into policies and supervise the analysis of user perceptions on Twitter (RQ3).

2 Research Plan

2.1 Methods

The proposed work will consist of three main phases to answer our research questions.

2.1.1 RQ1: Software

Method and Evaluation. To answer our first research question, we will analyze code changes associated with the implementation of data privacy regulations. To identify code changes required to implement regulated privacy standards, we will analyze pull requests on open source GitHub repositories. A pull request (PR) is a mechanism for developers to propose changes to a code base that elicits feedback from project maintainers before being accepted, or merged into the project. We will search GitHub for PRs containing relevant keywords, such as `"GDPR"` and `"CCPA"`.[1] Then, we will inspect the code changes and descriptions to determine if the contributions implement changes based on data privacy laws. The GitHub API will also be used to derive metrics to further examine changes, such as lines of code modified, number of commits, and total contributors.

[1] For example, `https://github.com/apache/hive/pull/3629`

2.1.2 RQ2: Developers

Method and Evaluation. To answer our second research question, we will analyze comments from software developers on online programming communities. We will use a web crawler to collect Stack Overflow posts containing relevant keywords the post title and question. To analyze questions developers ask, we will analyze question posts to derive themes and use metrics such as the number of upvotes from other users and answer responses to characterize questions. Further, we will determine developer opinions on data privacy regulations by collecting responses on StackOverflow questions[2] and comments on GitHub PRs. Comments will be analyzed using natural language processing (NLP) techniques, such as sentiment analysis, and qualitative analysis to determine the developer emotions and perceptions with regard to implementing data privacy regulations.

2.1.3 RQ3: Users

Method and Evaluation. We will answer our third research question by investigating comments made by software users about data privacy regulations on Twitter. To collect comments from users, we will mine tweets using a Twitter academic license[3] to analyze relevant comments on the introduction of data privacy regulations. We will determine relevant quotes by searching for hashtags `"#GDPR"` and `"#CCPA"`,[4] in addition to searching for related terms describing regulation changes that impact user experiences, such as `"cookie acceptance"`. We will use sentiment analysis to understand users' perspectives on data privacy regulations. Additionally, we will employ surveys to collect data describing how user experiences changed and general opinions on data privacy laws.

2.2 Deliverables

Our findings will be submitted to peer-reviewed conferences and workshops related to SE, HCI, and Cybersecurity. Potential venues include Foundations of Software Engineering (ESEC/FSE), Human Factors in Computing Systems (CHI), Mining Software Repositories for Privacy and Security (MSR4P&S), and the Symposium on Usable Privacy and Security (SOUPS). Beyond technical venues, we will also target papers for the Journal of Cybersecurity and the Journal of Cyber Policy, as well as engage with colleagues in the Hague Program for Cyber Norms, the Internet Observatory, and the CCI.

We will provide additional deliverables to extend and broaden the reach of this work. We will create a database containing the data collected from this project for researchers to analyze to answer new research questions. Additionally, we will publish relevant findings on the Tech4Humanity blog [3] and other industry-focused resources to share our results with practitioners and the general public. Upon completing our analysis, we plan to give a public lecture at Virginia Tech in which we will present and discuss our findings. Finally, we will maintain a project website with updates and data repositories related to this work for the purpose of replication within the broader scientific and policy communities.

[2] For examples, see `https://stackoverflow.com/search?q=ccpa`

[3] `https://developer.twitter.com/en/products/twitter-api/academic-research`

[4] For examples, see `https://twitter.com/hashtag/GDPR` and `https://twitter.com/hashtag/CCPA`

Plan for Additional Funding

Funding from the CCI Securing Interactions between Humans and Machines research call will support the research team in initiating the proposed work, in particular starting the GitHub and Stack Overflow analyses for RQ1 and RQ2. Our preliminary results will help us pursue additional funding opportunities. We aim to submit the the proposed research to the National Science Foundation (NSF). Specifically, we will target calls for proposals within the NSF Computer and Information Science and Engineering (CISE) and Technology, Innovation and Partnerships (TIP) directorates as well as the Designing Accountable Software Systems (DASS) program. Other potential privacy-related funding opportunities are available via the Open Technology Fund and the Hewlett Foundation.

Commercialization and Economic Impact

There are potential opportunities for commercialization and economic impact from the outcomes of this work. Our results for RQ1 aim to provide insight into the impact of data privacy regulations on software. To that end, the answers to this question can motivate the design of intelligent systems to support developers as they refactor code to incorporate data privacy standards. For example, software bots to support automatic compliance for data privacy regulations in source code. Additionally, RQ2 seeks to understand questions and challenges faced by software developers. These results can motivate tools, such as integrated development environment plugins or code analysis tools, to help programmers adhere to data privacy standards within development processes.

The proposed work can also have an impact on the economy. The software industry is one of the most prominent, projected to amass nearly $600 billion dollars worldwide in 2022 [5]. Further, reworking code based on changed requirements after programs are deployed to users costs 10 to 200 times more change and leads to projects being over budget, behind schedule, and lower quality [28]. By understanding the impact of implementing data privacy regulations on software and its stakeholders, this work can motivate the design of tools and guidelines to minimize financial losses related to these changes.

Impact on the Commonwealth

In March 2021, Virginia became just the second state in the United States to introduce state-wide legislation regarding the protection of users' data online by enacting the Virginia Consumer Data Protection Act (CDPA). This law derives concepts from the GDPR in the European Union and the CPRA in California to protect the user data of Virginians and regulate companies who conduct business in the state or offer services in the Commonwealth. The proposed work can help inform policy makers of the impact of implementing new data privacy regulations and motivate tools and processes to make it easier for software developers and users to transition to new regulations. If amendments to the VA CDPA are introduced, the proposed work can help facilitate smooth transitions by understanding the amount of effort for code changes, the information required for developers, and the disruption to user experiences needed to implement new policies.

NSF BIOGRAPHICAL SKETCH

NAME: Brown, Chris

NSF ID: 000866061@nsf.gov
ORCID: 0000-0002-6036-4733

POSITION TITLE & INSTITUTION: Assistant Professor, Virginia Tech

(a) PROFESSIONAL PREPARATION -(see PAPPG Chapter II.C.2.f.(a))

INSTITUTION	LOCATION	MAJOR / AREA OF STUDY	DEGREE (if applicable)	YEAR YYYY
North Carolina State University	Raleigh, NC	Computer Science	PHD	2021
North Carolina State University	Raleigh, NC	Computer Science	MS	2017
Duke University	Durham, NC	Computer Science	BS	2013

(b) APPOINTMENTS -(see PAPPG Chapter II.C.2.f.(b))

2021 - present Assistant Professor, Virginia Tech, College of Engineering, Blacksburg, VA
2020 - 2020 Instructor of Record, North Carolina State University, College of Engineering, Raleigh, NC
2018 - 2018 Quality Engineering Intern, Red Hat, Raleigh, NC
2017 - 2017 Quality Engineering Intern, Red Hat, Raleigh, NC
2016 - 2021 Graduate Research Assistant, North Carolina State University, College of Engineering, Raleigh, NC
2016 - 2016 Software Quality Engineer Intern, Blackbaud, Charleston, SC
2015 - 2016 Graduate Teaching Assistant, North Carolina State University, College of Engineering, Raleigh, NC
2013 - 2015 Python Developer, Bank of America, Charlotte, NC

(c) PRODUCTS -(see PAPPG Chapter II.C.2.f.(c))

Products Most Closely Related to the Proposed Project

1. Brown C, Parnin C. Understanding the impact of GitHub suggested changes on recommendations between developers. Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. (ESEC/FSE); 2020; Virtual Event USA. New York, NY, USA: ACM; c2020. Available from: https://dl.acm.org/doi/10.1145/3368089.3409722 DOI: 10.1145/3368089.3409722
2. Wang P, Brown C, Jennings J, Stolee K. Demystifying regular expression bugs. Empirical Software Engineering. 2021 November 05; 27(1):-. Available from: https://link.springer.com/10.1007/s10664-021-10033-1 DOI: 10.1007/s10664-021-10033-1
3. Brown C, Parnin C. Sorry to Bother You: Designing Bots for Effective Recommendations. 2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE). 2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE); ; Montreal, QC, Canada. IEEE; c2019. Available from: https://ieeexplore.ieee.org/document/8823645/ DOI: 10.1109/BotSE.2019.00021
4. Wang P, Brown C, Jennings J, Stolee K. An Empirical Study on Regular Expression Bugs. Proceedings of the 17th International Conference on Mining Software Repositories. (MSR);

2020; Seoul Republic of Korea. New York, NY, USA: ACM; c2020. Available from: https://dl.acm.org/doi/10.1145/3379597.3387464 DOI: 10.1145/3379597.3387464

5. Brown C, Parnin C. Comparing Different Developer Behavior Recommendation Styles. Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops. ICSE '20: 42nd International Conference on Software Engineering; 27 0 20; Seoul Republic of Korea. New York, NY, USA: ACM; c2020. Available from: https://dl.acm.org/doi/10.1145/3387940.3391481 DOI: 10.1145/3387940.3391481

Other Significant Products, Whether or Not Related to the Proposed Project

1. Behroozi , Parnin , Brown . Asynchronous technical interviews: Reducing the effect of supervised think-aloud on communication ability. Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE); 2022; ACM; c2022.
2. Brown C, Middleton J, Sharma E, Murphy-Hill E. How software users recommend tools to each other. 2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). 2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC); ; Raleigh, NC. IEEE; c2017. Available from: http://ieeexplore.ieee.org/document/8103460/ DOI: 10.1109/VLHCC.2017.8103460
3. Chris Brown, Chris Parnin. Sorry to Bother You Again: Developer Recommendation Choice Architectures for Designing Effective Bots. 2020 IEEE/ACM 2nd International Workshop on Bots in Software Engineering (BotSE); 2020; ACM; c2020. Available from: http://dx.doi.org/10.1145/3387940.3391506 isbn: 9781450379632
4. Brown C, Parnin C. Nudging Students Toward Better Software Engineering Behaviors. 2021 IEEE/ACM Third International Workshop on Bots in Software Engineering (BotSE); 2021; Madrid, Spain. IEEE; c2021. Available from: https://ieeexplore.ieee.org/document/9474399/ DOI: 10.1109/BotSE52550.2021.00010
5. Sun P, Brown C, Beschastnikh I, Stolee K. Mining Specifications from Documentation using a Crowd. 2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering. 2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER); 2019; Hangzhou, China. IEEE; c2019. Available from: https://ieeexplore.ieee.org/document/8668025/ DOI: 10.1109/SANER.2019.8668025

(d) SYNERGISTIC ACTIVITIES -(see PAPPG Chapter II.C.2.f.(d))

1. Program committee member and reviewer for software engineering and CS education-related academic workshops and conferences (BotSE, CHASE, SIGCSE BOF track)
2. Reviewer for software engineering-related academic journals (EMSE, TSE, PeerJ Computer Science, IEEE Software, IEEE Software Special Issue: Bots in Software Engineering)
3. As a Session Chair for the CHASE 2022 workshop, I coordinated the research presentations for the Mixed Session and facilitated a discussion and Q&A between the audienace and authors after the talks.
4. Invited talk at the It Will Never Work in Theory session at the Strange Loop 2022 industry-focused conference
5. Mentoring graduate and undergraduate students in the Department of Computer Science at

Virginia Tech. Students learn research methods and concepts to investigate ways improve the behavior, productivity, and decision-making of software engineers.

https://www.overleaf.com/project/631b9d580ee167548f2706bf

PI/co-PI/Senior Personnel: Brown, Chris

PROJECT/PROPOSAL PENDING SUPPORT

1. Project/Proposal Title: (This proposal) CCI: Understanding the Impact of Data Privacy Regulations on Software and Its Stakeholders

   Proposal/Award Number (if available):

   Source of Support: CCI

   Primary Place of Performance: Virginia Tech

   Project/Proposal Support Start Date (if available): 12/2022

   Project/Proposal Support End Date (if available): 12/2023

   Total Award Amount (including Indirect Costs): $60,000

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2023	1

   Overall Objectives: Understand the impact of data privacy regulations on software code, developers, and users.

   Statement of Potential Overlap: N/A

2. Project/Proposal Title: Supporting Under-Resourced Software Engineering Job Seekers through Facilitating Online Collaboration in Technical Interview Preparation

   Proposal/Award Number (if available):

   Source of Support: Google

   Primary Place of Performance: Virginia Tech

   Project/Proposal Support Start Date (if available): 11/2022

   Project/Proposal Support End Date (if available): 12/2023

   Total Award Amount (including Indirect Costs): $60,000

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed

Year	Person-months per year committed
2023	1

Overall Objectives: Develop an online platform (FASIT) to incorporate social interactivity into technical interview preparation for under-resourced job seekers.

Statement of Potential Overlap: N/A

https://www.overleaf.com/project/631b9d580ee167548f2706bf

NSF BIOGRAPHICAL SKETCH

NAME: Brantly, Aaron

POSITION TITLE & INSTITUTION: Associate Professor, Department of Political Science, Virginia Tech

(a) PROFESSIONAL PREPARATION -(see PAPPG Chapter II.C.2.f.(a))

INSTITUTION	LOCATION	MAJOR / AREA OF STUDY	DEGREE (if applicable)	YEAR YYYY
Queens University of Charlotte	Charlotte, NC	English Literature	BA	2004
The American University	Washington, D.C.	Public Policy	MOTH	2008
The University of Georgia	Athens, GA	International Relations/Comparative Politics	PHD	2012

(b) APPOINTMENTS -(see PAPPG Chapter II.C.2.f.(b))

2021 - present Associate Professor, Department of Political Science, Virginia Tech, Blacksburg, VA

2020 - present Director, Tech4Humanity Lab, Blacksburg, VA

2017 - 2021 Assistant Professor, Department of Political Science, Virginia Polytechnic Institute and State University, Blacksburg, VA

2014 - 2020 Senior Research Scientist, United States Army Cyber Institute, West Point, NY

2014 - 2017 Assistant Professor, Department of Social Sciences, United States Military Academy, West Point, NY

2014 - 2017 Cyber Policy Fellow, Combating Terrorism Center, West Point, NY

2013 - 2017 Adjunct Professor, Defense Strategic Studies Program, National Security Studies Institute, University of Texas, El Paso, TX

(c) PRODUCTS -(see PAPPG Chapter II.C.2.f.(c))

Products Most Closely Related to the Proposed Project

1. Brantly, Aaron F. and Damien Van Puyvelde. 2019. Cybersecurity: Politics, Governance and Conflict in Cyberspace, Cambridge, UK: Polity Press (978-1509528097). Cybersecurity : politics, governance and conflict in cyberspace. 2019 09.
2. Brantly, Aaron F. and Damien Van Puyvelde. 2017. US National Cyber Security: International Politics, Concepts and Organization Abingdon, UK: Routledge Press (978-0415787994). US National Cybersecurity : International Politics, Concepts and Organization. 2017 09.
3. Brantly Aaron Franklin. The Decision to Attack : Military and Intelligence Cyber Decision-Making. University of Georgia Press; 2016. isbn: 978-0820349206
4. Brantly A. Risk and uncertainty can be analyzed in cyberspace. Journal of Cybersecurity. 2021 January 01; 7(1):-. Available from: https://academic.oup.com/cybersecurity/article/doi/10.1093/cybsec/tyab001/6146840 DOI: 10.1093/cybsec/tyab001
5. Brantly A. From Cyberspace to Independence Square: Understanding the Impact of Social Media on Physical Protest Mobilization During Ukraine's Euromaidan Revolution. Journal of

Information Technology & Politics. 2019 August 21; 16(4):360-378. Available from: https://www.tandfonline.com/doi/full/10.1080/19331681.2019.1657047 DOI: 10.1080/19331681.2019.1657047

Other Significant Products, Whether or Not Related to the Proposed Project

1. Brantly A. The Cyber Losers. Democracy and Security. 2014 May 22; 10(2):132-155. Available from: http://www.tandfonline.com/doi/abs/10.1080/17419166.2014.890520 DOI: 10.1080/17419166.2014.890520
2. Brantly A, Brantly N. Patient-centric cybersecurity. Journal of Cyber Policy. 2020 December 09; 5(3):372-391. Available from: https://www.tandfonline.com/doi/full/10.1080/23738871.2020.1856902 DOI: 10.1080/23738871.2020.1856902
3. Brantly A. The Most Governed Ungoverned Space: Legal and Policy Constraints on Military Operations in Cyberspace. SAIS Review of International Affairs. 2016; 36(2):29-39. Available from: https://muse.jhu.edu/article/641158 DOI: 10.1353/sais.2016.0018
4. Brantly A. Entanglement in Cyberspace: Minding the Deterrence Gap. Democracy and Security. 2020 June 20; 16(3):210-233. Available from: https://www.tandfonline.com/doi/full/10.1080/17419166.2020.1773807 DOI: 10.1080/17419166.2020.1773807
5. Brantly A. The cyber deterrence problem. 2018 10th International Conference on Cyber Conflict (CyCon). 2018 10th International Conference on Cyber Conflict (CyCon); ; Tallinn. IEEE; c2018. Available from: https://ieeexplore.ieee.org/document/8405009/ DOI: 10.23919/CYCON.2018.8405009

(d) SYNERGISTIC ACTIVITIES -(see PAPPG Chapter II.C.2.f.(d))

1. Senior Expert (2019-2020) United States Cyberspace Solarium Commission. Washington, D.C.
2. Expert (2017) National Intelligence Estimate on Future of AI/ML in Warfare, Washington, D.C.
3. Academic Minor Program Advisor (2019-) Integrated Security, Virginia Tech
4. Co-PI DARPA Compass Program – AGRIPPA Team (2018-2020)
5. Conference Organizer, 2016 CyCon U.S. Conference on International Cyber Conflict in Collaboration Between the United States Army Cyber Institute and NATO CCDCOE. Washington, DC.

PI/co-PI/Senior Personnel: Brantly, Aaron

PROJECT/PROPOSAL SUBMISSION PLANNED

1. Project/Proposal Title: Integrated Security Destination Area 2.0

   Proposal/Award Number (if available):

   Source of Support: Virginia Tech

   Primary Place of Performance: Virginia Tech

   Project/Proposal Support Start Date (if available): 01/2023

   Project/Proposal Support End Date (if available): 01/2024

   Total Award Amount (including Indirect Costs): $40,884

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2023	1

   Overall Objectives: The creation of an academic hub for integrated security, privacy, and trust lead by the Integrated Security Education and Research Center (ISERC).

   Statement of Potential Overlap: N/A

2. Project/Proposal Title: Understanding the Impact of Data Privacy Regulations on Software and Its Stakeholders

   Proposal/Award Number (if available):

   Source of Support: Commonwealth Cyber Initiative

   Primary Place of Performance: Virginia Tech

   Project/Proposal Support Start Date (if available): 11/2022

   Project/Proposal Support End Date (if available): 11/2023

   Total Award Amount (including Indirect Costs): $60,000

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed

Year	Person-months per year committed
2022	1
2023	1

Overall Objectives: The proposed work will explore how adherence to new data privacy regulations affects software and its stakeholders.

Statement of Potential Overlap: None

FILENAME: #VALUE!
PRIN. INVESTIGATOR: Chris Brown
BUDGET PERIOD: 12/1/2022 through 12/31/2023
DUE DATE: 9/30/2022

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

			12/01/22-11/30/23 Period 1			TOTAL
	NAME	% EFFORT	REQUESTED SALARY	CAL MONTHS	FRINGES	REQUESTED
1	Chris Brown, Assist. Prof. (SMR)	0%	$0	0.000		$0
2	GRA - Step 13 (AY) CS	100%	$22,286	9.000	$2,068	$22,286
3	GRA - Step 13 (SMR) CS	100%	$7,287	3.000	$675	$7,287
	TOTAL PERSONNEL SALARIES		$29,573			$29,573
	FRINGE BENEFITS:					
	See rates in table below		$2,743			$2,743
	TOTAL SALARIES AND FRINGES		$32,316			$32,316
	EQUIPMENT (greater than or equal to $2000 per item)		$7,783			$7,783
	server	$7,783				
	TRAVEL		$1,500			$1,500
	Domestic	$1,500				
	International	$0				
	TUITION COSTS		$16,901			$16,901
	TOTAL DIRECT COSTS		$58,500			$58,500
	F&A COSTS		$0			$0
	F&A Not Allowed By Sponsor 0.00%	$0				
	TOTAL COSTS		$58,500			$58,500
	Base for F&A Costs		*$33,816*			$33,816
	Base for F&A Costs		*$33,816*			$33,816
						$0
						$0

FRINGE RATES	Through 6/30/23	On/After 7/1/23
Regular Faculty	32.44%	33.91%
Special Research Faculty	35.41%	37.64%
Part Time Faculty	29.07%	26.82%
SMR Faculty/Wage Employee	7.64%	7.14%
GRA	9.32%	9.21%
Classified Staff	49.36%	50.67%

FILENAME: #VALUE!
PRIN. INVESTIGATOR: Chris Brown
BUDGET PERIOD: 12/1/2022 through 12/31/2023
DUE DATE: 9/30/2022

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

			12/01/22-11/30/23 Period 1				TOTAL
	NAME		% EFFORT	REQUESTED SALARY	CAL MONTHS	FRINGES	REQUESTED
4	Aaron Brantly, Political Science		0%	$0	0.000		$0
5	GRA - Step 13 (AY)		0%	$0	0.000		$0
6	GRA - Step 13 (SMR)		0%	$0	0.000		$0
7			0%	$0	0.000		$0
	TOTAL PERSONNEL SALARIES			$0			$0
	FRINGE BENEFITS:						
	See rates in table below			$0			$0
	TOTAL SALARIES AND FRINGES			$0			$0
	TRAVEL			$1,500			$1,500
	Domestic		$1,500				
	International		$0				
	TUITION COSTS			$0			$0
	TOTAL DIRECT COSTS			$1,500			$1,500
	F&A COSTS			$0			$0
	F&A Not Allowed By Sponsor	0.00%	$0				
	TOTAL COSTS			$1,500			$1,500
	Base for F&A Costs			*$1,500*			$1,500
	Base for F&A Costs			*$1,500*			$1,500
							$0
							$0

FRINGE RATES	Through 6/30/23	On/After 7/1/23
Regular Faculty	32.44%	33.91%
Special Research Faculty	35.41%	37.64%
Part Time Faculty	29.07%	26.82%
SMR Faculty/Wage Employee	7.64%	7.14%
GRA	9.32%	9.21%
Classified Staff	49.36%	50.67%

FILENAME: #VALUE!
PRIN. INVESTIGATOR: Chris Brown
BUDGET PERIOD: 12/1/2022 through 12/31/2023
DUE DATE: 9/30/2022

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

		12/01/22-11/30/23 Period 1				TOTAL
	NAME	% EFFORT	REQUESTED SALARY	CAL MONTHS	FRINGES	REQUESTED
1	Chris Brown, Assist. Prof. (SMR)	0%	$0	0.000	$0	$0
2	GRA - Step 13 (AY) CS	100%	$22,286	9.000	$2,068	$22,286
3	GRA - Step 13 (SMR) CS	100%	$7,287	3.000	$675	$7,287
	TOTAL PERSONNEL SALARIES		$29,573			$29,573
	FRINGE BENEFITS:					
	See rates below		$2,743			$2,743
	TOTAL SALARIES AND FRINGES		$32,316			$32,316
	EQUIPMENT		$7,783			$7,783
	server	$7,783				
	TRAVEL					$3,000
	Domestic		$3,000			
	International		$0			
	TUITION COSTS		$16,901			$16,901
			#NAME?			#NAME?
	TOTAL DIRECT COSTS		$60,000			$60,000
	F&A COSTS		$0			$0
	F&A Not Allowed By Sponsor	0.00% $0				
	TOTAL COSTS		$60,000			$60,000
	Base for F&A Costs		*$35,316*			*$35,316*
	Base for F&A Costs		*$35,316*			*$35,316*
			$0			*$0*
			$0			*$0*

Requested VT F&A as a % of Total Funds Requested:	0.00%	0.00%
Requested VT + Sub F&A as a % of Total Funds Requested:	0.00%	0.00%

FRINGE RATES	Through 6/30/23	On/After 7/1/23
Regular Faculty	32.44%	33.91%
Special Research Faculty	35.41%	37.64%
Part Time Faculty	29.07%	26.82%
SMR Faculty/Wage Employee	7.64%	7.14%
GRA	9.32%	9.21%
Classified Staff	49.36%	50.67%

Budget Justification

The project period is for a duration of 12 months, 12/01/2022 – 12/31/2023.

Other Personnel
One calendar year graduate student from the Department of Computer Science will work on the project. The monthly stipend is $2,429, per the University's approved monthly stipend table, and based on a 20-hour workweek. The total student stipend is $29,573. An escalation factor of 5% is included and occurs August 15th each year.

Fringe Benefits

FRINGE RATES	Through 6/30/23	On/After 7/1/23
GRA	9.32%	9.21%

The fringe benefits total $2,743.

Annual negotiations with the Office of Naval Research (ONR) result in fixed rates for Employee Benefits covering the period July 1st- June 30th. Benefits include: Fee Waivers, Workman's Compensation, Retirement, Unemployment, FICA, Life Insurance, Hospitalization and Educational Leave. A copy of Virginia Tech's federally negotiated fringe rate agreement is available at: http://osp.vt.edu/resources/rates.html

Equipment
$7,783 is budgeted to allow for the purchase of a server for hosting data repositories (such as https://data.gov using a local https://getdkan.org - DKAN instance on the server.

Travel
Domestic travel in the amount of $3,000 is budgeted to allow the research team to travel to the annual CCI symposium.

The University follows the Commonwealth of Virginia travel policy and procedures which provide for reimbursement of "reasonable" cost in connection with official travel. As a State agency the University is obliged to reimburse travel costs in conformance with State policy. Reimbursement in compliance with this policy is consistent with the requirements of Federal Acquisition Regulation (FAR) 31.3.

Other Direct Costs
Tuition of $16,901 is budgeted for the entire project period, and based on the 2022- 2023 tuition fees for an engineering student at the Blacksburg, Virginia campus. All sponsored program proposals that include graduate student stipends in the budget must also include tuition and technology and library fees for the same time frame (AY) that the students(s) will be on GRA stipends. The amount includes a 3% escalation factor each year, which occurs on August 15th.

Total Direct Costs
$60,000

References

[1] Github. https://github.com.

[2] Stack overflow. https://stackoverflow.com.

[3] Tech4Humanity lab blog. https://tech4humanitylab.org/blog.

[4] Twitter. https://twitter.com.

[5] Technology markets: Software – worldwide. Statista, 2022. https://www.statista.com/outlook/tmo/software/worldwide.

[6] A. Alhazmi and N. A. G. Arachchilage. I'm all ears! listening to software developers on putting gdpr principles into software development practice. *Personal and Ubiquitous Computing*, 25(5):879–892, 2021.

[7] M. Behroozi, C. Parnin, and C. Brown. Asynchronous technical interviews: Reducing the effect of supervised think-aloud on communication ability. In *Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*. ACM, 2022.

[8] A. Brantly. The Cyber Losers. *Democracy and Security*, 10(2):132–155, 2014.

[9] A. Brantly. *The Decision to Attack: Military and Intelligence Cyber Decision-Making*. University of Georgia Press, 2016.

[10] A. Brantly. The cyber deterrence problem. *International Conference on Cyber Conflict, CYCON*, 2018-May:31–53, 2018.

[11] A. Brantly. From Cyberspace to Independence Square: Understanding the Impact of Social Media on Physical Protest Mobilization During Ukraine's Euromaidan Revolution. *Journal of Information Technology & Politics*, 16(4):1–19, 2019.

[12] A. Brantly. Risk and uncertainty can be analyzed in cyberspace. *Journal of Cybersecurity*, 7(1), 2021.

[13] A. F. Brantly. Public health and Epidemiological approaches to national cybersecurity: A baseline comparison. US National Cybersecurity International Politics, Concepts and Organization. Taylor and Francis, 2017.

[14] C. Brown, J. Middleton, E. Sharma, and E. Murphy-Hill. How software users recommend tools to each other. In *Visual Languages and Human-Centric Computing*, 2017.

[15] C. Brown and C. Parnin. Sorry to bother you: designing bots for effective recommendations. In *Proceedings of the 1st International Workshop on Bots in Software Engineering*, pages 54–58. IEEE Press, 2019.

[16] C. Brown and C. Parnin. Comparing different developer behavior recommendation styles. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, ICSEW'20, page 78–85, New York, NY, USA, 2020. Association for Computing Machinery.

[17] C. Brown and C. Parnin. "Sorry to bother you again: Developer recommendation choice architectures for designing effective bots". In *Proceedings of the 2nd International Workshop on Bots in Software Engineering*, BotSE 2020, Seoul, South Korea, 2020. ACM.

[18] C. Brown and C. Parnin. Understanding the impact of github suggested changes on recommendations between developers. In *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, page 1065–1076, New York, NY, USA, 2020. Association for Computing Machinery.

[19] C. Brown and C. Parnin. Nudging students toward better software engineering behaviors. In *3rd IEEE/ACM International Workshop on Bots in Software Engineering, BotSE@ICSE 2021, Madrid, Spain, June 4, 2021*, pages 11–15. IEEE, 2021.

[20] D. C. Brown, Jr. *Digital Nudges for Encouraging Developer Behaviors*. PhD thesis, North Carolina State University, Raleigh, NC, USA, 2021.

[21] R. E. Bucklin and C. Sismeiro. Click here for internet insight: Advances in clickstream data analysis in marketing. *Journal of Interactive marketing*, 23(1):35–48, 2009.

[22] D. Craigen, N. Diakun-Thibault, and R. Purse. Defining cybersecurity. *Technology Innovation Management Review*, 4:13–21, 10/2014 2014.

[23] GitHub. The state of the octoverse, 2019. `https://octoverse.github.com/`.

[24] J. Isaak and M. J. Hanna. User data privacy: Facebook, cambridge analytica, and privacy protection. *Computer*, 51(8):56–59, 2018.

[25] Z. S. Li, C. Werner, and N. Ernst. Continuous requirements: An example using gdpr. In *2019 IEEE 27th International Requirements Engineering Conference Workshops (REW)*, pages 144–149, 2019.

[26] J. C. Machado and P. R. Amora. The impact of privacy regulations on db systems. *Journal of Information and Data Management*, 12(5), 2021.

[27] A. Mahindrakar and K. P. Joshi. Automating gdpr compliance using policy integrated blockchain. In *2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, pages 86–93, 2020.

[28] N. R. Mead and T. Stehney. Security quality requirements engineering (square) methodology. *ACM SIGSOFT Software Engineering Notes*, 30(4):1–7, 2005.

[29] J. M. V. Mendia and J. J. A. Flores-Cuautle. Toward customer hyper-personalization experience — a data-driven approach. *Cogent Business & Management*, 9(1):2041384, 2022.

[30] R. Pressman. *Software Engineering: A Practitioner's Approach*. McGraw-Hill, Inc., USA, 7 edition, 2009.

[31] D. V. Puyvelde and A. F. Brantly. *Cybersecurity : politics, governance and conflict in cyberspace*. Cambridge, UK, 09 2019.

[32] S. Shastri, V. Banakar, M. Wasserman, A. Kumar, and V. Chidambaram. Understanding and benchmarking the impact of gdpr on database systems. *Proc. VLDB Endow.*, 13(7):1064–1077, mar 2020.

[33] J. Smith, C. Brown, and E. Murphy-Hill. Flower: Navigating program flow in the ide. In *2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 19–23, Oct 2017.

[34] D. Torre, G. Soltana, M. Sabetzadeh, L. C. Briand, Y. Auffinger, and P. Goes. Using models to enable compliance checking against the gdpr: An experience report. In *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems (MODELS)*, pages 1–11, 2019.

[35] P. Wang, C. Brown, J. A. Jennings, and K. T. Stolee. An empirical study on regular expression bugs. In *International Conference on Mining Software Repositories*, MSR 2020, Seoul, South Korea, 2020. IEEE Press.

[36] P. Wang, C. Brown, J. A. Jennings, and K. T. Stolee. "Demystifying regular expression bugs: A comprehensive study on regular expression bug causes, fixes, and testing". *Empirical Software Engineering*, 2021.