



## BU Spark! Justice Media Co-Lab Computational Journalism Practicum

XCC 433 Spring 2020  
Wednesday 1:00-3:45 pm  
WED 406  
[jump to class outline](#)

**Course requirements:** [online application](#)

**Class tools** ([links forthcoming](#))

- [Slack](#)
- [Github](#)
- [Google Drive folder](#)
- Readings and Supplementary Materials

### Course info

**Instructor:** Brooke Williams  
**Contact Info:** brooke@bu.edu; text/whatsapp/signal: 202-615-3551  
**Office Hours:** [Sign up on calendar](#); [Meet in Zoom office](#)

**Instructor:** Osama AlShaykh  
**Contact Info:** osama@bu.edu; PHONE: 858-361-9043  
**Office Hours:** [Sign up on calendar](#); [Zoom Link](#)

**Spark! Liaison:** Ziba Cranmer  
**Contact Info:** [zcranmer@bu.edu](#); 503-807-7815;  
**Office Hours:** TBD; ([zoom Link](#))

**Spark! Student PM/ TA:** TBD  
**Contact Info:** [TBD@bu.edu](mailto:TBD@bu.edu);  
**Office Hours:** TBD; Zoom Link

**Eng. in Residence:** Bandon Das/ Langdon White  
**Contact Info:** [bsd@redhat.com](mailto:bsd@redhat.com)/ [Langdon@redhat.com](mailto:Langdon@redhat.com)  
**Office Hours:** @ (virtual only) **Forthcoming**  
**Course Objectives**

Welcome to the Justice Media Computational co-Lab, an XCC BU Hub course and collaboration among Spark!, the College of Communication and the Center for Data and Computing Sciences. This course is a newsroom and a laboratory, where you will have an opportunity to work on interdisciplinary, student teams to co-produce a data-driven news investigation for one of our established media partners. Your projects will focus on issues surrounding justice.

Indeed, shedding light on inequities and injustices is among our highest callings in both journalism and civic technology. In this course, we will come together to dig deeper into issues surrounding justice using advanced computational tools and precision journalism. Each of you will have the opportunity to make a real difference in communities and gain a professional publication clip showing your skills. These are real world stakes with professional news organizations, but unlike a newsroom, you will have professors and mentors who specialize in and practice data science and journalism to guide and support you along the way.

In this course, you will learn how to think creatively and critically about government data in order to find the most important stories that shed light on inequities and injustices in your publication partner's coverage area. You'll learn how to request data under public records laws and how to think creatively about methods and solutions if the data you want don't yet exist. You will learn how to get access to data via APIs and web-scraping methods for a journalism project. You will learn to translate real life requirements and needs to software and tools to aid a journalist produce an important story, as well as how to work with data scientists and software engineers to express the needs and expectations.

You'll learn to work with peers who have different specialities and come up with creative ways to solve problems such as how to continuously, automatically scrape a website for data or how to link multiple datasets to find answers to journalistic questions in the pursuit of justice.

### **Learning outcomes - HUB**

This class has been approved for Spring 2021 as a Cross College Challenge (XC 433) which will meet the BU Hub Learning Outcomes for the following 4 Hub Areas:

#### **1. Creativity/Innovation**

*As teams work on their projects and address specific challenges throughout the semester, students:*

- Learn creativity as an iterative process of imagining new possibilities that involves risk-taking, use of multiple strategies, and reconceiving in response to feedback. They will be able to identify individual and institutional factors that promote and inhibit creativity.
- Engage in creative activity by conceiving and executing original work as part of their team.
- Embrace creativity in the process of conceptualizing a story vision and data sources. Students also will be given an opportunity to identify technology tools for development that would contribute to improvements in investigative journalism processes.
- Develop creative approaches to collect and update datasets for news investigations
- Strategies for overcoming government efforts to unfairly block public information.
- Think critically about your reporting plan and develop a creative methodology
- Explore ways to visualize data in unique, unsung and engaging ways for an investigative news report and create an interactive visualization for publication.
- Create multimedia elements for news investigations

## **2. Oral and/or Signed Communication**

*In communicating with team members, sponsors and other constituents of the project, and through their final team project presentations, students will:*

- Craft and deliver responsible, considered and well-structured oral and/or signed arguments using media and modes of expression appropriate to the situation.
- Interact, attend and respond thoughtfully to others.
- Speak/sign effectively in situations ranging from the formal to the extemporaneous and interact comfortably with diverse audiences.
- Pitch ideas and receive interim feedback from peers, professors and media partners
- Present projects at the end-of-semester presentations
- Develop questions and conduct interviews with sources from victims to experts to government officials
- Engage with professional media partners
- Experience in communicating effectively with diverse audiences.
- Negotiate with government officials in charge of public records and data

## **3. Research and Information Literacy**

*By engaging in substantial research to complete their projects, and following an iterative process for creating a project proposal, completing interim assignments and a final report, students will:*

- Search for, select, and use a range of publicly available and discipline-specific information sources ethically and strategically to address research questions.

- Understand the overall research process and its component parts, formulate good research questions or hypotheses, gather and analyze information, and critique, interpret, and communicate findings.
- Advanced online research skills for finding and exploring data and data options
- The latest online tools and computational methods for investigating those in power
- For some teams, the fruits of research will yield new approaches to enduring questions, or new artistic expressions, or fresh arguments.
- This experience is first and foremost an exercise in research through diverse and interdisciplinary methods including data science and journalism using primary and secondary sources as well as relying on both data and interviews with a range of individuals from experts to those impacted to those in power.

#### **4. Teamwork/Collaboration**

*Students are assigned to a team and receive explicit training in teamwork. Through this sustained collaboration on the project, and through the completion of a team contract and project plan, students will:*

- Learn the characteristics of a well-functioning team.
- Develop the tools and strategies for working successfully on a diverse team, including assigning roles and responsibilities, giving and receiving feedback, and engaging in meaningful group reflection that inspires collective ownership of results.
- Establish team agreements for productive collaboration.
- Provide mid-term and end-of semester feedback to their teammates.
- Collaborate on computational and journalistic solutions and make decisions together while giving and receiving feedback.
- Develop an interdisciplinary methodology to complete a data-driven news investigation
- As a team, achieve publication of a collaborative, data-driven investigation that informs the public, holds powerful accountable and otherwise supports justice.

#### **Course format**

This course is a laboratory and a newsroom. It features faculty lectures and guest speakers such as editors for media partners and leaders in civic tech. You will learn skills, practice them and collaboratively apply what you learn to your team project — a data-driven investigation to be published with a media partner. We will have regular team meetings, and we have reserved some class time for working on projects together with faculty guidance.

This course has neither a mid-term nor a final exam.

#### **There will be four phases of the course:**

- I. Team agreements, scope and methodology

- II. Research, data collection, data analysis, visualization for reporting
- III. Reporting, interviewing, writing, visualization for publication
- IV. Editor engagement, fact checking, data integrity checks and otherwise preparing for publication

### **Team resources:**

Students are also able to access computing services e.g. AWS credits, etc. They must be requested at this link and we will assess the request and get back to your team as soon as possible. <https://tinyurl.com/computing-intake>

Each team is allowed up to \$200 for incidental expenses directly associated with implementing their project (i.e. not pizza, etc.). You can submit receipts or procurement requests through this form: [<https://bu.campusgroups.com/>] For questions, please email Korinne Dizon at [kdizon@bu.edu](mailto:kdizon@bu.edu).(please copy [buspark@bu.edu](mailto:buspark@bu.edu))

### **Required materials:**

- A functioning laptop with a power cord, about 800 MB of free space and administrative privileges to install software. Bring your laptop to each and every class. Please do not hesitate to see us privately if you have any concerns about this requirement. We're happy to help.

### **Strongly recommended reading and engagement:**

- Data Literacy: A User's Guide, David Herzog (online or bookstore)
- Numbers in the Newsroom (Available for \$10 here: <https://store.ire.org/products/numbers-in-the-newsroom-using-math-and-statistics-in-news-second-edition-e-version>)
- Student membership for \$25 to Investigative Reporters and Editors. NICAR-L is a great way to engage with the computational journalism community.
- The news! Please regularly read the news stories your media partner is publishing and also focus on data journalism published in newspapers, on the radio and TV, as well as via nonprofits such as the [Center for Public Integrity](#), [Reveal](#), the [GBH Center for Investigative Reporting](#) and [ProPublica](#). Twitter feeds of data journalists who inspire you.

Please do not hesitate to reach out privately if you have any concerns about obtaining the required or recommended equipment and materials so we can try to help.

### **Grading**

The course grade will be based on the following:

40%: project memos, take-home assignments, teamwork and collaboration, class participation

60 %: Final draft of data-driven investigation with at least one accompanying data visualization and at least one multi-media element such as photographs or audio.

### **Mentoring and office hours**

Office hours with instructors are listed above. We are also available on Slack or email for scheduling and questions.

### **Class and university policies**

**Academic conduct, plagiarism and fabrication:** You may discuss homework assignments with classmates, but you are solely responsible for what you turn in. Collaboration in the form of discussion is allowed and encouraged. We understand that there may be teams working on projects together and they must document which team member completed different tasks.

Any use of 3rd party software, APIs, or algorithms in your product creation should be through third party licensing terms, open source licensing terms, or references articles. It should be clearly documented. This includes using snippets of software or the complete software.

All journalism must be factual, original and clearly cited and attributed in your story. You may not take information from another news story or website and put it in your story without proper attribution. You may not make up or invent any portion of your news stories. Journalists lose their jobs and sometimes their careers when they plagiarize and fabricate.

“Plagiarism is the act of representing someone else’s creative and/or academic work as your own, in full or in part. It can be an act of commission, in which one intentionally appropriates the words, pictures or ideas of another, or it can be an act of omission, in which one fails to acknowledge/document/give credit to the source, creative and/or copyright owner of those works, pictures or ideas. Any fabrication of materials, quotes or sources, other than that created in a work of fiction, is also plagiarism. Plagiarism is the most serious academic offense that you can commit and can result in probation, suspension or expulsion.”

We – both teaching staff and students – are expected to abide by the guidelines and rules of the Academic Code of Conduct.

**Disability Accommodations:** If you are a student with a disability or believe you might have a disability that requires accommodations, please contact the [Office for Disability Services](#).

**Sexual Misconduct:** Boston University is committed to fostering a safe, productive learning environment. Title IX and our school policy prohibit discrimination on the basis of sex, which regards sexual misconduct — including harassment, domestic and dating violence, sexual assault, and stalking. We understand that sexual violence can undermine students’ academic success, and we encourage students who have experienced some form of sexual misconduct to talk to someone about their experience, so they can get the support they need. Confidential support and academic advocacy resources can be found with the Center for Sexual Assault Response & Prevention.

**Equal Opportunity:** This course celebrates diversity and welcomes all. BU has strict guidelines on classroom behavior and practices when it comes to treatment of students and guests on the basis of race, color, religion, sex, gender identity, sexual orientation, age, mental or physical disability, genetic information, military service, national origin, or due to marital, parental, or veteran status. Discrimination for any of these reasons is prohibited. Please refer to the [Equal Opportunity/Affirmative Action Policy](#) for more details.

**Positive Classroom Experience:** At your discretion, please alert one or both of your instructors to anything related to preferred pronouns, preferred name or nickname, and/or any extenuating circumstances or triggers that might affect your classroom experience. We want to make sure you have the most positive experience in the classroom as possible.

This course affirms people of all gender expressions and identities. If you prefer to be called a different name than what is on the class roster, please let us know. Feel free to correct us on your preferred gender pronoun. If you have any other questions or concerns, please do not hesitate to let us know.

**Social Climate:** Any student who has difficulty affording groceries or accessing sufficient food to eat every day, or who lacks a safe and stable place to live, and believes this may affect their performance in the course, is urged to contact the Dean of Students for support. Furthermore, please notify one or both of your professors, if you are comfortable in doing so. This will enable us to provide any resources we may possess.

It is not unusual for students to feel stress, and about 15% of students experience depression, anxiety or other mental health concerns. Please know that we are here to help you find the resources to help you get through this stressful time.

If work shown in this class, professional or student-generated, offends you in any way, please mention it in class or talk to us privately about it so that we can all learn from each other. This is not to say we will ever restrict freedom of speech or water down an aggressive or edgy idea, but we want to discuss anything that someone deems troublesome or offensive.

Finally, there are many [resources available](#) to students.

The College of Communication has a [Diversity, Equity and Inclusion](#) Committee as well as a DEI student group and [Facebook group](#). All are welcome.

The College of Engineering has an [Inclusion and Outreach program and team](#).

**Student athletics:** All student-athletes should be provided with a sheet from Student-Athlete Support Services regarding absences throughout the semester. These sheets should be handed in as soon as possible to avoid potential conflicts and so arrangements can be made to provide for missed lecture notes, classwork, or discussion.

**Recording of classes:** Due to the majority of classes being offered in the Learn from Anywhere format, please expect that each class session will be recorded. It is important to note that recordings on Zoom might capture the chat during the class, including private chats. If you have questions or concerns regarding recording of this class, please don't hesitate to ask.

### **About your instructors:**

Brooke Williams is an investigative journalist and associate professor of the practice of computational journalism at Boston University. She specializes in data-driven reporting and storytelling to hold the powerful accountable and inform people. Her work has contributed to a Pulitzer Prize for Investigative Reporting, a George Polk Award and a Gerald Loeb Award, among many others. Most recently, as a contributor to [The New York Times](#), she co-authored the front-page, data-driven investigative series [Think Tanks Inc.](#) Currently, she is working on a nationwide, data-driven investigation of the Department of Justice, focusing on federal prosecutorial misconduct and accountability, for [The Intercept](#). Previously, she was an [investigative journalism fellow](#) at Harvard University. Williams grew up in California and graduated from the [Missouri School of Journalism](#). While in J-school, she worked as a photojournalist for The Maneater, a student newspaper (favorite shot [here](#)) and as a beat reporter at a daily newspaper, the Columbia Missourian. She began her career as a writer at the [Center for Public Integrity](#) in January 2002, where she co-authored [Harmful Error: Investigating](#)



[America's Local Prosecutors](#), [The Buying of the President 2004](#) and [Windfalls of War](#), an investigation into defense contracts that won a George Polk Award. In 2004, Williams joined the watchdog team at the San Diego Union-Tribune, a newspaper where she was a finalist for the Livingston Award for Young Journalists for [stories](#) examining how the city mishandled public land. In 2007, after wildfires ravaged San Diego County, her prize-winning [investigation](#) into contractors the city hired to haul away debris resulted in a federal criminal [probe](#) and taxpayer [settlement](#). In 2009, Williams left the Union-Tribune to join [inewssource](#). In 2012, she continued writing occasionally for inewssource while launching her career as a freelance journalist. Her freelance investigations have appeared in local, national and international publications including on the web, radio, [television](#) and in print. Williams also joined the [Lab@Edmond J. Safra Center for Ethics](#) at Harvard University as an investigative journalism fellow and later joined the Project on Public Narrative at Harvard University. Check out some of her work [here](#).

[Osama Alshaykh](#) is a faculty member in the electrical and computer engineering department at Boston University. Osama is also the CEO of nxtec corporation and CTO and board member of Talon. He served as an Associate Editor for IEEE Transactions on CSVT and an area Editor for Signal Processing: Image Communication, Theory, Techniques & Applications Journal. Osama was part of the founding team of PacketVideo Corporation and was its CTO. He has served on technology and development boards including AT&T, Verizon, Android, Euclid, Talon and LogicBlox.

**Class outline**

Date/ Session	Class Topics	Assignments (due following class)
<b>January 27</b>	Overview of the class  Partner pitches  Teaming activity  Team formation	Verbal pitch with 1-3 slides for a data-driven investigation. You will choose from partner stories or, if approved, a student pitch. We will enforce a 5-minute time limit per presentation. Please be sure to include — in a clear, concise and engaging way — the premise of the investigation: How will it further justice, and why is it important to the public?

<p><b>February 3</b></p>	<p>Lightning Lectures:</p> <p>Data-driven news investigations, justice, why it's important, examples</p> <p>Computational methods: Breaking down the alphabet soup of CS terminology</p> <p>Student pitches presented, discussion</p> <p>Finalize projects</p> <p>Team building and agreement</p>	<p>Team Agreements (<i>more to come on what needs to be included in the agreement</i>)</p>
<p><b>February 10</b></p>	<p>Data collection techniques: advanced web scraping, point and click web scraping tools, APIs,</p> <p>Lightning lectures:</p> <p>Obtaining public data using open records requests.</p> <p>Advanced web searching to find data (and documents) online.</p> <p>Building a database for the data for easy access m retrieval and linking.</p> <p>Creative thinking about linking datasets to find insights and uncover inequities. Examples</p>	<p>Focus on data collection</p> <p>If your team has a dataset in hand, please start exploring what's in it and start making a list of potential ways to look at it for your investigation and/or combine it with other data. Additionally, start a list of potential problems with the data and what's necessary for data clean up and organization/standardization.</p>

	<p>of common data linkages (names, locations, etc.)</p> <p>Team check ins</p>	
<b>February 17</b>	<p>Data preprocessing Data linking Data quality Data transforming</p> <p>Journalistic examples</p> <p>Team check ins</p>	<p>Data cleaning and standardizing, preparing for analysis. Outline due next week on your quality checks, data linking and methodology.</p>
<b>February 24</b>	<p>Data analysis methods: NLP, topic modeling, entity extraction, knowledge graphs, bag of words, conditional probability, sentiment, etc.</p> <p>SQLite for data-driven investigative journalism</p>	<p>SQL Exercise — answer questions and complete analysis in SQLite to find a potential news story. We will provide a handout and dataset.</p>
<b>March 3</b>	<p>Data bias: definition, origins, and mitigation</p> <p>Data bias in journalism: What have we learned?</p> <p>Team check ins</p>	<p>Data ready for analysis</p> <p>Readings on data bias to be provided</p>
<b>March 10</b>	<p>Workshop</p> <p>Teams will work in breakout rooms on data driven investigation with guidance from the professors and mentors.</p>	<p>Present with 1-3 slides on an investigative news story based on your data. We will enforce a 5-minute time limit per presentation. Please be sure to include — in a clear, concise and engaging way the premise of the investigation: How will it further justice, and why is it important to the public?</p>

<p><b>March 17</b></p>	<p>Team presentations of your plan for an investigative story stemming from your data.</p> <p>Discussion, feedback and team check ins</p> <p>Interviewing techniques: How to loosen lips during an investigation.</p>	
<p><b>March 24</b></p>	<p>Lightning lecture: Writing, structuring and presenting a complex, data-driven investigation</p> <p>Workshop</p> <p>Teams will work in breakout rooms on data driven investigation with guidance from the professors and mentors.</p>	<p>First draft due next week</p>
<p><b>March 31</b></p>	<p>Data integrity checks — peer review, reproduce peer outcomes and analysis using the same methods. Give feedback.</p> <p>Facts and news. Overview of journalistic practices and fact checking standards for a data-driven investigation.</p> <p>Student lightning talks followed by discussion: What did you find? Elevator explanation of your findings.</p>	

April 7	<p>Workshop</p> <p>Teams will work in breakout rooms on data driven investigation with guidance from the professors and mentors.</p> <p>Student lightning talks followed by discussion: Lessons and learned and ideas moving forward</p>	Second draft due next week.
<b>April 14</b>	<p>Visual exploration: What are the most effective ways to visualize your story?</p> <p>Tableau, timeline.js, QGIS</p>	<p>Second draft due to partner</p> <p>Get creative and visualize your data. Be prepared to present in the next class about how you did it and the choices you made — what is each visual exploration trying to convey?</p>
<b>April 21 (no classes - Wednesday)</b>	<p><b>optional workshop</b></p> <p>If desired, teams will work in breakout rooms on data driven investigation with guidance from the professors and mentors.</p>	Please be prepared for us to send your draft, including visuals, to your media partner this week.
<b>April 28 (Last class)</b>	<p>Ideation session: technology tools that will facilitate improvements in the investigative journalism process (I wish I had...)</p> <p>Workshop</p>	
<b>April 30-May 3 (Study Period)</b>	<p>Workshop</p> <p>Teams will work in breakout rooms to meet partner</p>	

	expectations on their data-driven investigation with guidance from the professors and mentors.	
<b>May 4-8</b>	<p>Workshop</p> <p>Teams will work in breakout rooms to meet partner expectations on their data-driven investigation with guidance from the professors and mentors.</p>	no final exam
<b>May 6 Justice Chronicles: Media Co-Lab Showcase</b>	Present your story. Include slides with the lede, visualizations and any multi-media.	