

## **Bridging Past and Present: Modernizing Du Bois' Visualizations for Contemporary Data Science Education**

**Abstract (250 words)** - Narrative overview of the proposed project for a non-specialist audience and suitable for general publication, such as the AUC Data Science Initiative website. Clearly state why the work is important in data science and what will be accomplished.

Towards the end of the 19<sup>th</sup> century, Dr. W.E.B. Du Bois and an interdisciplinary team of students and alumni from across the South envisioned and created an innovative collection of hand-drawn data visualizations depicting the state of Black America. The revolutionary data visualizations were part of an exhibit at the 1900 Paris Exposition held in Paris, France. Modern researchers have extracted the data from the Du Bois visualizations and recreated them using technologies such as Tableau. However, many of the social issues visualized in the Du Bois exhibit still exist in 2024. Although some of the charts in the Du Bois exhibit are no longer measurable or applicable in modern America, the need for visualizing the state of Black America still exists. This project aims to recreate and adapt the Du Bois data visualizations using current data (2000-2024). The vision is to create a carefully curated repository of charts designed and implemented by students and faculty at institutions around the United States. The data visualizations will be created using various visualization tools like Tableau and Python. Creators will contribute completed charts, instruction manuals and/or videos and data via a user portal for public access. The objectives of the project are to recognize the historic contributions of Du Bois, provide a public repository for students and faculty to display the results of their data visualization research, provide instruction manuals that can be used as modules in visualization courses, and provide visualizations that researchers can reference in African-American research.

## Proposal Description (6 pages max) - Include sections:

Proposing Principal Investigator(s) (up to 5)

**PI: Sajid Hussain PhD** ([sajid.hussain@mmc.edu](mailto:sajid.hussain@mmc.edu)), Professor/Chair, Dept of Computer Science and Data Science, Meharry Medical College

**Consultant: Dr. Nadine Shillingford PhD** ([nshillingford@truezerotech.com](mailto:nshillingford@truezerotech.com)), Data Analytics Engineer, True Zero Technologies

**Collaborator 1:** Graham West, PhD ([gwest@mmc.edu](mailto:gwest@mmc.edu)), Assistant Professor, Dept of Computer Science and Data Science, Meharry Medical College

**Collaborator 2:** Lei Qian, PhD ([lqian@mmc.edu](mailto:lqian@mmc.edu)), Associate Professor, Dept of Computer Science and Data Science, Meharry Medical College

**Collaborator 3:** Thema Monroe-White, PhD, Associate Professor of Artificial Intelligence (AI) and Public Policy, George Mason University

## Objectives and Significance of the Project

**Develop and Enhance Data Visualization Curriculum:** Create comprehensive instructional materials, including manuals and videos, that can be integrated into graduate-level data science and analytics courses to enhance the learning experience and improve curriculum design.

**Promote Historical and Modern Data Visualization Techniques:** Recognize and celebrate the historical contributions of Dr. W.E.B. Du Bois by recreating his data visualizations with modern data and tools. This will not only honor historical achievements but also bridge them with contemporary data visualization practices using technologies like Tableau and Python.

**Establish a Collaborative and Public Repository:** Develop a publicly accessible repository where students and faculty from various institutions can contribute, share, and display their data visualizations. This repository will serve as a resource for both instructional purposes and as a reference for ongoing research related to the African-diasporic experience and data science education.

## Plan and Methodology

The following structured plan and methodology will ensure the project meets its objectives, creates high-quality educational resources, and provides a valuable repository for ongoing research and education in data visualization.

### **Phase 1: Project Planning and Team Formation**

*Project Planning:* Define the project scope, objectives, and deliverables and develop a detailed project timeline with milestones and deadlines. [Completed]

*Team Formation:* Assemble an interdisciplinary team of faculty, graduate students, and alumni and assign roles and responsibilities based on expertise (data collection, visualization, curriculum development, etc.).

### **Phase 2: Data Collection and Preparation**

*Historical Data Review:* Review the original Dr. W.E.B. Du Bois (Battle 2018) charts and identify data points that are still relevant and those that need to be updated with modern data.

*Modern Data Collection:* Collect relevant data from reliable sources such as the US Census Bureau and Data.Gov to update the visualizations. Ensure data is clean, well-documented, and formatted for analysis.

### **Phase 3: Visualization Creation**

*Tool Selection and Setup:* Select visualization tools such as Tableau and Python and ensure all team members have access to the necessary software and training.

*Visualization Design:* Recreate historical visualizations with modern data. Design new visualizations that reflect contemporary social issues.

*Review and Iteration:* Conduct regular review sessions to ensure accuracy and clarity. Iterate on designs based on feedback from the team and external reviewers.

### **Phase 4: Instructional Material Development**

*Manuals and Videos:* Develop comprehensive instructional manuals detailing the creation process of each visualization and/or create instructional videos that demonstrate how to use the visualization tools and interpret the data.

*Module Integration:* Design course modules that can be integrated into existing data science curricula and test modules in pilot courses and refine based on student feedback.

### **Phase 5: Repository Development and Launch**

*Repository Design:* Design a user-friendly online portal for the repository and ensure the repository supports easy uploading, sharing, and access to visualizations.

*Content Upload:* Populate the repository with completed visualizations, manuals, videos, datasets, metadata, and documentation for each entry.

*Public Launch:* Launch the repository to the public and promote the repository through academic channels, social media, and relevant conferences.

### **Phase 6: Evaluation and Improvement**

*Feedback Collection:* Gather feedback from repository users, educators, and researchers using surveys, and usage analytics to assess the impact and effectiveness.

*Continuous Improvement:* Regularly update the repository with new visualizations and instructional materials and address any issues or suggestions for improvement based on feedback.

*Reporting and Dissemination:* Prepare a final report summarizing the project's outcomes, impact, and lessons learned. Share findings at a workshop or conference such as the 2025 South Big Data Hub Partnership Forum.

## **Project Outcomes, Deliverables, Timeline**

### **Project Outcomes**

*Enhanced Data Science Curriculum:* Integration of new instructional materials into graduate-level courses, leading to improved teaching methodologies and student engagement in data visualization.

*Recognition of Historical Contributions:* Increased awareness and appreciation of Dr. W.E.B. Du Bois' work in data visualization among students and faculty, fostering a deeper understanding of the historical context and its relevance today.

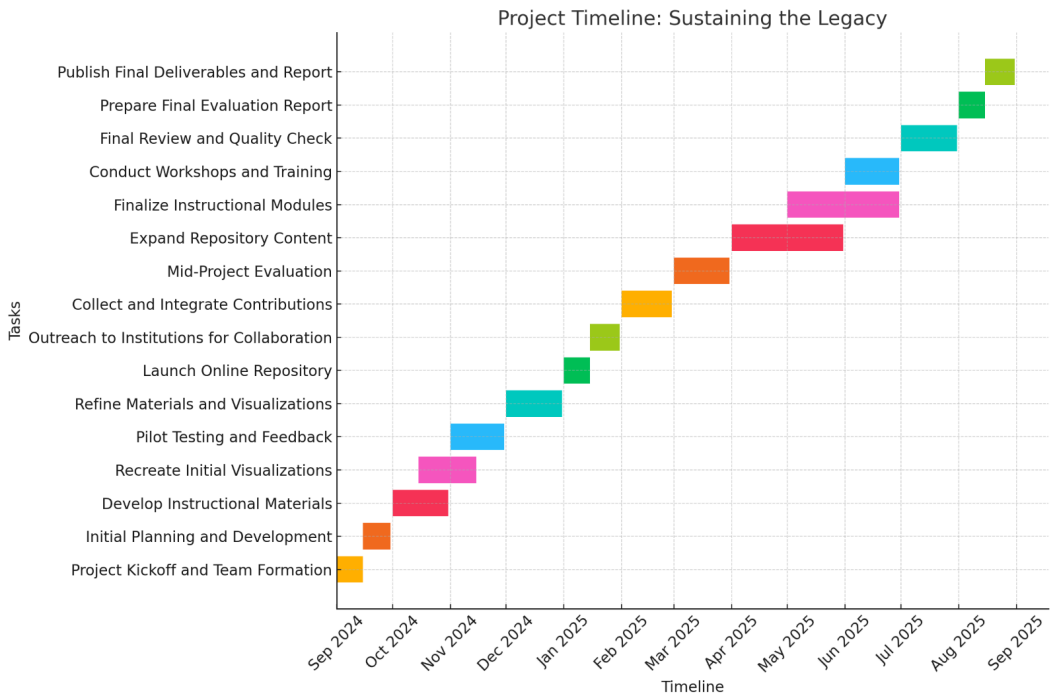
*Comprehensive Public Repository:* Establishment of a publicly accessible online repository of modernized data visualizations, instruction manuals, and videos, providing a valuable resource for educators, students, and researchers.

*Collaborative Research and Learning Platform:* Creation of a collaborative platform that facilitates the sharing of data visualization projects, promoting interdisciplinary collaboration and continuous learning.

**Project Deliverables**

Phases	Deliverables
Phase 2: Data Collection and Preparation	Datasets
Phase 3: Visualization Creation	Modernized data visualizations
Phase 4: Instructional Material Development	Manuals, guides, and video tutorials
Phase 5: Repository Development and Launch	User-friendly portal giving access to searchable and categorized instructional materials, visualizations, and supporting data
Phase 6: Evaluation and Improvement	Evaluation and impact reports

**Timeline**



**Anticipated Impacts**

**Enhanced Data Science Education:** Improved curriculum including contemporary and historical perspectives on data visualization and increased student engagement and understanding of complex data science concepts.

**Increased Awareness and Historical Understanding:** Recognition of the historical significance of Du Bois' Contributions and its relevance to current social issues. Students and

faculty will gain a deeper appreciation for the historical context of data visualization and the ongoing impact of social issues visualized by Du Bois.

**Resource for Research and Education:** The creation of a comprehensive, publicly accessible repository will provide a valuable resource for educators, students, and researchers. This repository will facilitate the sharing and dissemination of knowledge and best practices in data visualization and serve as a reference point for researchers in African-American studies and other fields.

**Collaborative Learning and Community Building:** The project will foster collaboration among students and faculty from various institutions, promoting interdisciplinary learning and the sharing of diverse perspectives. The project will also engage a wider community in discussions about data visualization and social issues.

**Professional Development and Skill Building:** Participants will develop and refine their skills in data visualization tools such as Tableau and Python, enhancing their professional competencies resulting in hands-on experience with real-world data and visualization techniques.

**Impact on Social Awareness and Policy:** The visualizations will bring attention to ongoing social issues affecting Black America, potentially influencing public discourse and policy decisions. They can be used in various educational outreach initiatives, raising awareness and fostering informed discussions among broader audiences.

## **Dissemination Plan**

Objective: To effectively share the project's outcomes, deliverables, and impacts with a broad audience, including educators, students, researchers, and the general public by:

**Presenting Papers and Posters at Academic Conferences and Workshops** targeting educators, researchers, and students in the fields of data science, history, and African-American studies.

**Publishing Papers in Peer-Reviewed Journals** targeting academic and professional communities.

**Publishing Blog Posts and Feature Articles** for institutional websites and online platforms targeting the general public, educators, students, and researchers.

**Creating the Online Repository and Accompanying Website** including forums and discussion boards (to facilitate collaboration and feedback) targeting the general public, educators, students, and researchers.

**Creating Dedicated Social Media and Online Platforms** to share updates, milestones, evaluation reports, feedback, and new content with the general public, educators, students, and professionals.

**Conducting Institutional and Community Outreach Programs** targeting educational institutions, community organizations, and local libraries. Partner with universities and colleges to integrate project materials into their curricula.

**Hosting Webinars and Virtual Info Sessions** to introduce the project to a wider audience.

## **Project Assessment**

**Formative Assessment** to monitor progress and make ongoing improvements during the project lifecycle. Proposed activities include:

1. Bi-monthly meetings with collaborators to review progress, address challenges, and adjust plans as needed.
2. Implementing pilot tests of instructional materials in select courses and gathering feedback from students and instructors.
3. Arranging peer reviews of visualizations and instructional content by experts in data science and African-American studies to ensure accuracy and relevance.

**Summative Assessment** to evaluate the overall success and impact of the project upon its completion. Proposed activities include:

1. Distributing surveys to students, educators, and repository users to gather their perceptions of the project's effectiveness, usability, and impact on learning.
2. Analyze data from the online repository to measure the number of visits, downloads, and user interactions. Track the geographic distribution of users to assess the project's reach.

**Impact Assessment** to measure the broader impact of the project on data science education, historical awareness, and community engagement. Proposed activities include:

1. Assessing improvements in students' understanding of data visualization concepts and historical context through pre- and post-course assessments.
2. Evaluating the extent to which the new instructional materials have been integrated into graduate-level data science courses and their impact on teaching practices.
3. Tracking citations and references to the repository's visualizations and materials in academic publications and research projects.
4. Collecting testimonials and case studies from community organizations and educators who have used the project's resources.

**Performance Metrics** to quantify specific aspects of the project's performance. Proposed metrics include:

1. Number of Contributed Visualizations
2. Repository Engagement (number of unique visitors, page views, and time spent on the repository).
3. Educational Materials Utilization (number of downloads and uses of instructional manuals and video tutorials).
4. Collaboration and Partnerships (number of collaborating institutions and the diversity of their contributions).

**Feedback Mechanisms** to create channels for continuous improvement based on user feedback. Proposed activities include:

1. Online Feedback Forms on the repository website
2. Follow-Up Surveys to users who have accessed the repository or participated in workshops to gather additional insights on their experiences.
3. Annual Review Meetings: Hold annual review meetings with collaborators to discuss feedback, evaluate progress, and plan future initiatives.

## **Sustainability Plan**

**Financial Sustainability** by continuously seeking out grant opportunities from educational and research funding bodies, including NSF.

**Ongoing Engagement and Collaboration** by fostering a community of practice by creating networking opportunities, such as regular webinars, conferences, and workshops, for educators,

researchers, and students and encouraging active participation and contributions to the repository from a diverse group of institutions and individuals.

**Resource Maintenance and Updates** by establishing a schedule for regularly updating the repository with new data visualizations, instructional materials, and case studies encouraging contributors to submit updated versions of their visualizations and materials as new data and tools become available, Ensure the repository's technical infrastructure is robust and scalable to handle increasing traffic and data storage needs. Develop an archival strategy to preserve historical versions of visualizations and materials, ensuring that they remain accessible for future reference.

**Institutional Partnerships and Integration** by forming long-term partnerships with academic institutions to embed the project's materials into their curricula and research programs. Offer co-branded courses, workshops, and certifications in collaboration with partner institutions.

**Educational Outreach** by partnering with K-12 schools and community organizations to extend the project's reach and introduce younger audiences to data visualization and historical research.

**Evaluation and Reporting** using regular assessments and publishing annual reports detailing the project's progress, financial status, and future plans.



## References Cited (WIP)

<https://www.tableau.com/blog/how-web-du-bois-used-data-visualization-confront-prejudice-early-20th-century>

<https://towardsdatascience.com/present-to-persuade-lessons-from-the-data-visualizations-of-w-e-b-du-bois-a50775f0fa90>

[https://www.thinknum.com/resources/webinar/present-to-persuade?utm\\_source=abdus&utm\\_medium=webinar&utm\\_campaign=dubois](https://www.thinknum.com/resources/webinar/present-to-persuade?utm_source=abdus&utm_medium=webinar&utm_campaign=dubois)

<https://medium.com/nightingale/w-e-b-du-bois-staggering-data-visualizations-are-as-powerful-to-day-as-they-were-in-1900-64752c472ae4>

<https://www.dignityanddebt.org/projects/du-boisian-resources/>

Cunningham, Andrew, Alyxander Burns, and Narges Mahyar. "Looking to the Past to Visualize the Present: Revisiting WEB Du Bois' Abolitionist Visualizations."

Battle-Baptiste, Witney, and Britt Rusert, eds. *WEB Du Bois's data portraits: Visualizing black America*. Chronicle Books, 2018.

Katherine Fusco, Lynda C. Olman, Techniques of Justice: W. E. B. Du Bois's Data Portraits and the Problem of Visualizing the Race, *MELUS*, Volume 46, Issue 3, Fall 2021, Pages 159–187, <https://doi.org/10.1093/melus/mlab031>

Tackett, M., & Çetinkaya-Rundel, M. (2023). Analyzing and Recreating Data Visualizations of W.E.B. Du Bois. *CHANCE*, 36(1), 40–47. <https://doi.org/10.1080/09332480.2023.2179279>

[https://public.tableau.com/app/profile/chimdi.nwosu/viz/W\\_E\\_B\\_DuboisPortraitGallery/MainGallery](https://public.tableau.com/app/profile/chimdi.nwosu/viz/W_E_B_DuboisPortraitGallery/MainGallery)



**Budget (1 page max)** - Itemized statement of expected costs and a subtotal of these costs, along with a subtotal for the reassigned time if requested. Note that funds must be spent by the project end date.

**Budget Justification (2 pages max)** - A narrative explanation of the budget where each cost should be listed, briefly described, and justified. Include why each item is needed, its purpose, and how the dollar total was determined.

**Statement on Research Funding (1 page max)**