



HDR DSC: The Metropolitan Chicago Data-science Corps (MCDC)



Lizhen Shi, Francisco Iacobelli, Suzan van der Lee, Rohan Attele, Michelle Birkett, Bennett Goldberg, Nadja Insel, Mark Potosnak, Eunice Santos, Matt Sperry, Ragaa Abdallah, Katherine Amato, Elsa Anderson, Denise Drane, Sungsoon Hwang, Arend Kuyper, Ting Liu, William Miller, Howard Rosing, Yoo-Seong Song, Rob Voigt, Sara Woods, Phil Yates, and members of the MCDC collaboration

Aims

- ❑ Aim 1: Establish a network across diverse academic institutions to discover and disseminate best practices in data science.
- ❑ Aim 2: Design cross-disciplinary training and learning opportunities that provide broad analytical and transferable skills to students
- ❑ Aim 3: Facilitate access to open data, real-world data science projects, and training materials for the common good.
- ❑ Aim 4: Empower the community and nation to transform data into actionable and inspiring knowledge for better decision-making.

Challenges

Diversity

- ❑ The level at which students engage with the MCDC course and practicum is extremely diverse. This is exacerbated in mixed student teams.
- ❑ The kinds of projects are also very diverse in terms of needs, scope and techniques

Communication

- ❑ Each partner follows a different path to MCDC and has different expectations. Consistent messaging and managing expectations is hard.

Coordination

- ❑ Universities have different academic calendars and community partners operate on yet different timelines. Matching projects with practica and DAU teams/faculty mentor requires broad knowledge.
- ❑ Incorporating new academic partners

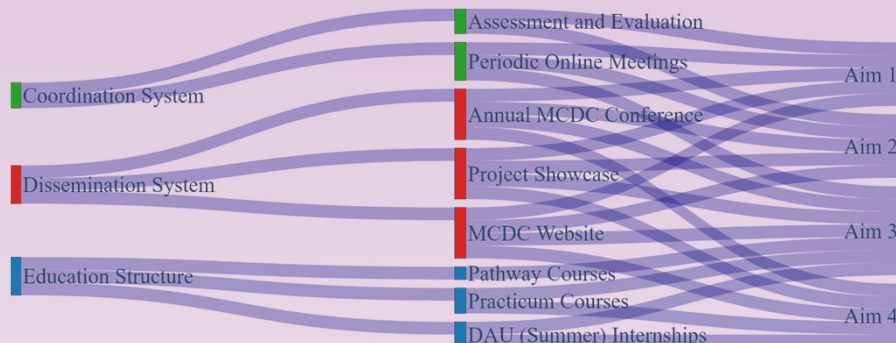


Figure 1. MCDC Structure Sankey Diagram

Methodology

The MCDC is structured into three main systems: Education Structure, Coordination System, and Dissemination System. Each system plays a crucial role in achieving the project aims through various pathways and activities. Fig. 1 provides a detailed overview of the relationships and contributions of each component to the project aims.

outlines the process of collecting RDS from the community to delivering valuable insights and actionable knowledge to community partners. Along this pipeline, students engage with a comprehensive range of topics that integrate mathematics, statistics, computer science, and domain-specific expertise, equipping them to analyze and interpret complex data with proficiency.

Accomplishments

By implementing the strategic

methodology and rigorous data science curriculum outlined above, MCDC has achieved the following outcomes:

- ❑ **Diverse Collaboration:** Universities, students, and faculty from various backgrounds collaborated to build data science capacity and strengthen community ties.
- ❑ **Data Science Capacity:** Developed data science skills at teaching-focused universities through impactful coursework and practicums.
- ❑ **Community Impact:** Engaged non-profit partners effectively use data science to support their missions.
- ❑ **Knowledge Sharing:** Cross-university collaboration fostered best practices, shared curricula, and practical, real-world experience for students.

Next Steps

Assemble Curriculum and Best Practices

- ❑ Develop techniques and incentives to engage a diverse body of students.

Empower Students

- ❑ Students in DAU teams bring vastly different experience, expertise, and perspectives. We aim to harness each individual's strengths.

Expand Academic Partner Engagement

- ❑ Expand data science education to community colleges in the region while offering ongoing support and professional development for their ds programs.

Data Science Curriculum

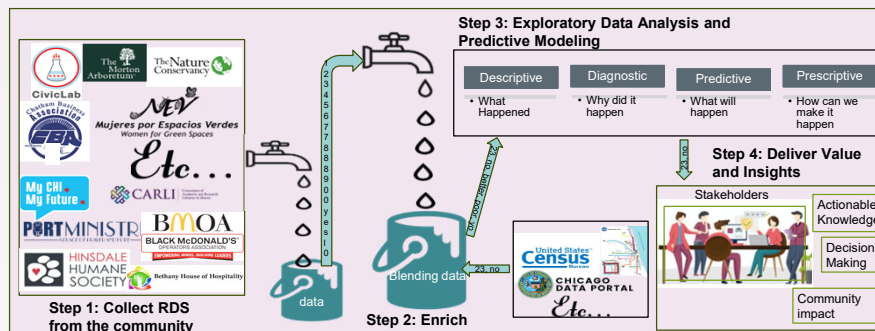


Figure 2. MCDC Data Science Workflow

The MCDC curriculum centers on multiple key areas of data acumen, integrated within the MCDC Data Science Pipeline (Fig. 2). This pipeline

