

Co-Designing Learning Materials to Empower Laypersons to Better Understand Big Data and Big Data Methods

Lisa M. Schilling¹

Department of Medicine, University of Colorado Anschutz
Medical Campus

lisa.schilling@ucdenver.edu

Janet Corral¹

Department of Medicine, University of Colorado Anschutz
Medical Campus

janet.corral@ucdenver.edu

Bethany Kwan¹

Department of Family Medicine, University of Colorado
Anschutz Medical Campus

bethany.kwan@ucdenver.edu

Griselda Pena-Jackson¹

2040 Partners For Health

gpenajackson@2040Partnersforhealth.org

Julie Ressalam¹

Data Science to Patient Value Program,

University of Colorado Anschutz Medical

Campus

julie.ressalam@ucdenver.edu

Seth Russell¹

seth.russell@ucdenver.edu

Data Science to Patient Value Program,

University of Colorado Anschutz Medical

Campus

¹University of Colorado Anschutz Medical

Campus

13199 E Montview Blvd

Aurora, CO 80045

ABSTRACT

University of Colorado Anschutz Medical Campus' Data Science to Patient Value Program and 2040 Partners for Health sought to create open learning materials for engaged citizens and community leaders regarding big data and big data methods to support their collaboration in patient-centered and participatory-based community research and evaluation. 2040 is a local non-profit organization that cultivates partnerships in Aurora, Colorado neighborhoods to tackle critical health needs. Our goal was to co-design and co-create a series of big data learning modules accessible to community laypeople, so they might better understand big data topics and be empowered more actively engage in health research and evaluation that uses big data methods.

Keywords

Big data, community engagement, education, co-design.

1. PROJECT DESCRIPTION

'Big data', a recently popularized term, is often a scary black box that only an elite few truly understand. Our goal was to elucidate this complicated subject for a layperson audience to support their engagement as collaborators in projects using big data and big data methods. This project is joint effort of the University of Colorado Anschutz Medical Campus' (CU) Data Science to Patient Value (D2V) Program, the Colorado Clinical and Translational Sciences Institute's (CCTSI) Community Research Liaisons (CRLs), and 2040 Partners for Health (2040), a community-based nonprofit focused on health needs of the local population. Key to this collaborative project was the thoughtful engagement of both campus and community members in co-creating content that was accessible to community members both in terms of open access and in terms of content suitability for their learning needs [1][2][3]. Our team consisted of experts in data science (Lisa Schilling, MD, MSPH, Seth Russell, MS), community engagement (Bethany Kwan, PhD, MSPH, Griselda Peña-Jackson, BS, BA), medical ethics (Matt Wynia, MD, MPH), educational design (Janet Corral, PhD, Catherine Battaglia, PhD, RN) and community members.

After establishing the team's roles and responsibilities, we separated into working groups to determine the optimal methods for designing and implementing our curriculum and for co-designing the curriculum with our invited community members (henceforth 'laypersons'). We reviewed the relevant scientific and grey literature to identify other community layperson co-designed curriculum, spoke with the authors/producers of similar work [1], and reviewed literature on best practices for explainer videos, as this was the eventual education medium preferred by our laypersons.

1.1 Guiding Principles

Each workgroup identified a set of guiding principles.

The community engagement workgroup agreed upon:

- Use of the principles of co-design [4][5]
- Goal of supporting person-centered involvement in translational research and evaluation [6][7][8]
- Emphasis on community-relevant examples and applications of data science as a foundation of the educational approach

The educational design workgroup agreed upon:

- Use of instructional methods that support open access, sharing and reuse of the learning materials, consistent with the principles of the Open Educational Resource (OER) movement [9][10]
- Use of evidence-based methods to create the content and assessments, as well as instructional videos and multimedia learning material [11][12][13]

1.2 Co-creating with the Community

We held our first 'Community Data Discussion' in June of 2018 and opened it broadly to the Aurora community. The Aurora community is one of the most diverse and largest underserved populations in the United States [14][15], and is the neighborhood of our growing CU Anschutz Medical Campus. 2040 was an essential trusted point of contact and outreach with this community. Seventeen persons, not including the project

team, attended this event ranging from laypersons, to community advocates, to leaders at various health and wellness organizations. From this event we gained a better understanding of the level of knowledge regarding big data, data analytics, and data science among the general community. Community members identified a variety of topics and issues which were important to them, and this was used to inform topics that later served as examples for the curriculum. We also recognized the need to focus our layperson recruitment on the intended audience for the educational modules - community members who would potentially engage as collaborators in patient-centered outcome research. To identify layperson representatives, we invited community members previously engaged in health and healthcare research at CU, including the CCTSI's CRLs, to co-design the curriculum and learning modules. CRLs are employees of CU and the CCTSI and are members of the community they are representing. By establishing relationships and systems of communication between the community and researchers, CRLs ensure community needs and perspectives are represented in research priorities. For this project, CRLs were also able to identify other community members with an interest in collaborating on research projects involving big data.

We had an initial meeting jointly with the invited layperson to describe the project's goals and decide upon agreeable organizational details, such as meeting times and processes. At this initial meeting, we worked together on the first module so all attendees would understand the nature of the work, and so we could get input on the best processes to employ during future co-design sessions. We also previewed an example explainer video and decided upon the preferred media of the learning materials. The options discussed included written-only, video recorded lectures (with or without a person in the recording), and animated explainer videos. The laypersons felt that the animated explainer videos were compelling and approachable. We also agreed that 3-5 minutes per video was an appropriate length for laypersons. This length would also positively challenge our subject matter experts to focus on core concepts with layperson-friendly explanations.

During co-design meetings, laypersons were asked to represent themselves as well as consider what others in their community might need to better understand this content. The scientific team prepared materials in advance of the co-design sessions. At the initial session, the learning module was delivered as a set of digital slides with a lecturer using notes (See Figure 1). This



Figure 1. Screenshot from Module 2's digital slides used for co-design.

process was modified slightly as we progressed, as we discovered that it was very important to get feedback on the exact words and examples used. Thus, for future sessions we spent more time upfront writing scripts and doing our best to remove technical and scientific jargon. We also relied more on the use of hand drawn illustrations to convey key concepts, as finding or creating digital representations for the co-design meetings was extremely time consuming and often not what we ultimately used (See Figure 2). The scripts and visual aids were drafted by two team members (Russell, Schilling) and revisions were done mainly by them with additional assistance from other team members. During the co-design sessions all members of the team brainstormed about the best examples to use to illustrate complex concepts. For example, when discussing the storage size of big data, the laypersons felt that using an example of digital photos storage on a phone would help other laypersons understand the concept of data storage size. Additionally, the mobile phone example helped to convey the use of databases to store information, with the example of a contacts database on mobile phones being a layperson-friendly way to convey this concept.

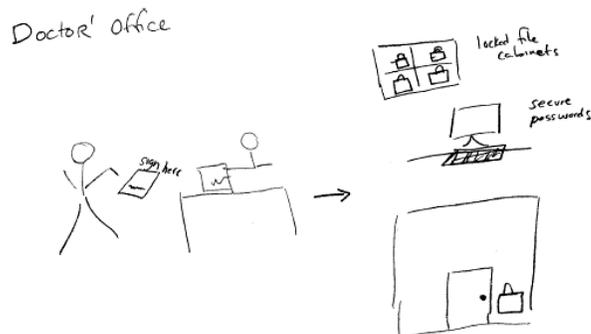


Figure 2. A visual aid design by our team to help illustrate data privacy.

After each session, team members incorporated feedback from the laypersons into the scripts by reducing jargon, simplifying language, and explaining data science concepts using examples that were identified during the co-design sessions. Once scripts and any corresponding visuals were fully drafted and approved by the team, we collaborated with a graphic design consultant to develop animated explainer videos for each module.

1.3 Co-Design Lessons Learned When Engaging the Community

We found that during a two-hour co-design session with the community, we could complete two learning modules. Each co-design session included 2-3 laypersons, a facilitator, the content presenter, and two note takers (See Figure 3).

We learned that efficient facilitation is a key component of the co-design process. Our experienced facilitators (Kwan, Pena-Jackson) stopped the presentation after every 2-3 slides to allow laypersons to share their immediate reactions to the content, identify any terms that they considered to be jargon or technical terminology (e.g. ETL, cloud computing, database) or unfamiliar vocabulary, despite common use in scientific communities (e.g. integration, algorithm).

The analogy about grocery store
shopping helps a lot, but the
pieces need to be connected better.
ie. is your personal budget the
data warehouse?

Figure 1. Notes from a layperson at a co-design session.

Group size is important. Based on the first co-design meeting, which had 7 layperson attendees, we realized that 2-3 laypersons were the optimal number of layperson attendees to allow meaningful engagement in co-design. Co-design is a detailed process, that requires deep engagement with the content and each other to correctly hear and represent participants' ideas. All team members had to be willing to listen and question one another's levels of understanding. A smaller group allowed for more productive and meaningful engagement.

2. GENERAL DESIGN AND COMMUNITY ENGAGEMENT LESSONS LEARNED

Our first community engagement event, referred to as a Community Data Discussion, was promoted by our CCTSI CRL partners and 2040 to those known to be interested in issues regarding big data. We found it difficult to describe to the Aurora community what their role would be and what we were trying to do due to the enigmatic nature of the phrase big data. The attendees ranged from representatives from community organizations that used data for program planning, evaluation and resource allocation, to experts in public health informatics interested in engaging with university projects. Although the event provided much useful information regarding the issues important to community members, including which data can inform community health resource needs and priorities, we quickly learned most attendees were not interested in participating in the co-design of learning materials. This helped us focus on engaging community laypersons that were already involved in collaborative research and more likely to see the value of this work and of their participation.

Even among our own scientific team, we frequently had to remind ourselves of, and stay focused on the project's goals. Our goal was to create learning modules that provide a basic understanding of the concepts and the vocabulary of big data and big data methods so that laypersons partnering in the design and dissemination of research and evaluation have more knowledge to support their engagement. Our goal was not to train people to be data scientists. For example, we currently have a funded project in which community members are engaged in research on developing and using health data record linkage methods. These educational modules will be important tools for increasing the capacity of these community members to engage meaningfully in this research project.

The process of creating learning modules that are attractive and understandable for a complex topic is very time consuming. The scientific teams had to be willing to revise content and focus the depth of the content on key concepts necessary for community

members' understanding. Having clear learning objectives and creating content maps of the learning material is essential to allow for modules that build off one another and limit any unnecessary content redundancies.

3. Outcomes

The current video series consists of 9 modules. We are currently finalizing the following video learning modules and their titles:

1. Big Data: What, Where, Why
2. Big Data Collection
3. Big Data Storage
4. Turning Data into Information
5. Big Data Privacy & Identity Protection
6. Accountable Data Use
7. Basic Concepts of Statistics
8. Making Predictions
9. AI and Machine Learning

They will be hosted on a public-facing, non-password protected sites of CU and 2040, as well as on YouTube, to support open access. While these formats do not guarantee equal access for all due to barriers associated with internet access and cellular data via mobile phones, they will be available to most persons.

4. EXTERNAL RESOURCES

Our materials are being finalized; please go to www.CUD2V.org or www.2040partnersforhealth.org/ for updates.

5. ACKNOWLEDGEMENTS

We want to acknowledge and thank our scientific and community team members, also listed at the above URL. For questions or comments please reach out to Lisa.schilling@ucdenver.edu.

6. REFERENCES

- [1] Valle, N., Brishke, J., Ritzhaupt, A. D., Evans, C., Nelson, D. R., & Shenkman, E. (2018). Using Instructional Design to Support Community Engagement in Clinical and Translational Research: a Design and Development Case. *Journal of Formative Design in Learning*, 2(1), 20-35.
- [2] Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *Bioscience*, 59(11), 977-984
- [3] Johnstone, S. M. (2005). Open educational resources serve the world. *Educause Quarterly*, 28(3), 15.
- [4] Scaife, M., Rogers, Y., Aldrich, F., & Davies, M. (1997, March). Designing for or designing with? Informant design for interactive learning environments. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (pp. 343-350). ACM
- [5] Van der Velden, M., & Mörtberg, C. (2015). Participatory design and design for values. *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, 41-66.
- [6] Fleurence, R. L., Curtis, L. H., Califf, R. M., Platt, R., Selby, J. V., & Brown, J. S. (2014). Launching PCORnet, a national

patient-centered clinical research network. *Journal of the American Medical Informatics Association*, 21(4), 578-582.

- [7] Concannon, T. W., Meissner, P., Grunbaum, J. A., McElwee, N., Guise, J. M., Santa, J., ... & Leslie, L. K. (2012). A new taxonomy for stakeholder engagement in patient-centered outcomes research. *Journal of general internal medicine*, 27(8), 985-991.
- [8] Fleurence, R., Selby, J. V., Odom-Walker, K., Hunt, G., Meltzer, D., Slutsky, J. R., & Yancy, C. (2013). How the Patient-Centered Outcomes Research Institute is engaging patients and others in shaping its research agenda. *Health Affairs*, 32(2), 393-400
- [9] Atkins, D. E., Brown, J. S., & Hammond, A. L. (2007). A review of the open educational resources (OER) movement: Achievements, challenges, and new opportunities (pp. 1-84). Mountain View: Creative common.
- [10] Caswell, T., Henson, S., Jensen, M., & Wiley, D. (2008). Open content and open educational resources: Enabling universal education. *The International Review of Research in Open and Distributed Learning*, 9(1).
- [11] Mayer, R., & Mayer, R. E. (Eds.). (2005). *The Cambridge handbook of multimedia learning*. Cambridge university press.
- [12] Clark, R. C., & Mayer, R. E. (2016). *E-learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning*. Hoboken: John Wiley & Sons.
- [13] Morrison, G. R., Ross, S. M., Kemp, J. E., & Kalman, H. (2010). *Designing effective instruction*. Hoboken: John Wiley & Sons.
- [14] Kolmar C. These Are The 10 Most Diverse Cities In Colorado For 2019. HomeSnacks [Internet]. 2019 2/28/19. Available from: <https://www.homesnacks.net/most-diverse-cities-in-colorado-128907/>.
- [15] Silver N. The Most Diverse Cities Are Often The Most Segregated. FiveThirtyEight [Internet]. 2015. Available from: <https://fivethirtyeight.com/features/the-most-diverse-cities-are-often-the-most-segregated/>.

About the authors:

Lisa Schilling is the Principal Investigator of this project. She is board-certified in Internal Medicine and Clinical Informatics. She is a practicing general internist with the University of Colorado. Dr. Schilling joined the University of Colorado's Department of Medicine in 1997 where she is currently a Professor in the Division of General Internal Medicine. She is the Co-Director of D2V and the Medical Director of Office of Value Based Performance. She is a co-PI of pSCANNER, a distributed data research network. Her research focuses on issues of data quality, data modeling, record linkage, and building infrastructures to support the secondary use of clinical data to generate meaningful knowledge that impacts health.

Janet Corral is an Associate Professor in Educational Informatics in the CU Department of Medicine. Dr. Corral serves as the Director of Teaching Scholars' Program for the Academy of Medical Educators and the Director of Digital Education for CU's Anschutz Medical Campus' Office of Digital Education.

Bethany M. Kwan is currently an assistant professor in CU's Department of Family Medicine. Dr. Kwan is a health psychologist and health services researcher and serves as the Director, CCTSI Quality and Process Improvement Program and the Investigator and Education Program Lead for the Adult & Child Consortium for Health Outcomes Research & Delivery Science program.

Griselda Pena-Jackson is the Executive Director of 2040. Ms. Pena-Jackson has 12 years specializing in non-profit organizations focused on community engagement and education.

Julie Ressalam served as the project manager of this project. She is the Senior Research Coordinator for the CU Center for Bioethics and Humanities and a project manager for the D2V program.

Seth Russell is a Research Analytics Instructor for the D2V program. He is a well-experienced informatician and software engineer.