

Towards a Framework for Justice-Oriented Data Science Education in K-12 Schools

Authors

Abstract: Data science is increasingly relevant to daily life and has garnered significant attention in education. While data science education has been traditionally focused on technical training, justice considerations are increasingly brought up given growing concerns over fairness and justice in data science. This paper introduces a framework for justice-oriented data science education that comprises five areas grounded in a broad range of literature. To explore and refine the framework in authentic contexts, we applied it to discourse data from one participatory design workshop with teachers. Analysis demonstrated the presence of this framework's areas and their rich connections in teachers' thinking. The framework offers educators a tool to integrate data science, justice issues, and disciplinary content in K-12 classrooms.

Objectives

Data science is an interdisciplinary field that integrates statistics, computing, and disciplinary knowledge to examine phenomena with data (Biehler et al., 2022; Hayashi, 1998). Data science is increasingly relevant to everyday life and has garnered attention in both higher education (Tang & Sae-Lim, 2016) and K-12 (V. Lee & Delaney, 2021). One productive approach to introducing data science in K-12, as shown by prior efforts, is to not treat data science as a replacement or add-on but rather as an opportunity to engage learners in authentic problems, practices, and actions relevant to data investigations (Wilkerson & Polman, 2020). A salient theme under this approach is to mobilize data science methods to create new opportunities for learners to examine justice issues related to society and environment (Linardi et al., 2022). However, more guidance is still needed to support the integration of data science, justice issues, and disciplinary content in K-12 classrooms. This study attempts to bridge this gap by proposing a justice-oriented framework for K-12 data science education.

Perspectives

Data Science and Its Fundamental Connection to Justice

Data science has garnered significant interest across virtually all social sectors creating a high demand for data science expertise in many organizations. Data science education has been traditionally focused on preparing data science professionals. To become a data scientist, one is expected to go through comprehensive training in statistics, computer programming, and related domain areas (NASEM, 2018). Data scientists are involved in data practices—such as collecting, visualizing, analyzing, interpreting, or communicating data—to create powerful tools (e.g., language models, predictive analytics, facial recognition) that are infused in modern societies.

However, this dominant approach to data science education—which is focused on technical training of a homogeneous group of people—problematically presents to students a neutral stance on the power and impact of these tools on society (see Philip & Sengupta’s 2021 critique). First, a growing body of research and public dialogue have questioned implicit biases that are embedded in data science methods and applications that may harm society or certain subpopulations (O’Neil, 2016; Taylor, 2017). Data science must embrace a political orientation and actively incorporate justice considerations instead of treating them as an afterthought (Green, 2021). Second, since almost every member of society is impacted by data science—when placing an online order, doing a web search, or using an automatic soap dispenser—a just, data-infused society depends on data-literate citizens who have data acumen, recognize injustices in algorithmic systems, and take actions to advance justice goals (Biehler et al., 2022). It becomes imperative to create opportunities for everyone to engage in data work as well as imagination of more just futures through working with data.

Data Science Education in K-12

Given these considerations, there are urgent pushes for data science education in K-12 that supports not only career development, but students’ civic engagement supported by data science practices (Irgens et al., 2020; Wilkerson & Polman, 2020). Emerging efforts in this area intend to reach every student; it also goes beyond teaching data science knowledge and skills to help students connect data science with their personal experiences, cultural identities, and sociopolitical positions (V. Lee et al., 2021; Pangrazio & Selwyn, 2021). In line with critical reflection on data science at the societal level, data science in K-12 provides an avenue for students to examine justice issues they care about, leading to potential gains in not only mathematical, statistical, and information literacy skills but also critical skills for agentive, politically minded, and civic participation.

A Framework for K-12 Justice-Oriented Data Science Education

Over the past year, our team has been working on an NSF-funded project to develop a justice-oriented data science program for secondary schools, called DataX. In this project, we use a participatory design approach (DiSalvo et al., 2017) to co-create and iteratively refine DataX with school teachers. Over a series of five half-day design workshops, the teachers and the research team are collaboratively designing the DataX curriculum and refining the DataX online learning platform.

Formative work informing these workshops included a comprehensive review of K-12 data science curricula that were accessible in English. Of the 25 curricula that were reviewed, only nine incorporated justice, and they did so in three distinct manners: (a) asking teachers to generally discuss bias in data collection and analysis; (b) incorporating an independent unit on data ethics; and, more interestingly, (c) designing multiple units to explore justice in contexts such as the criminal justice system. As shown by the review, more guidance is needed to meaningfully integrate justice in K-12 data science education.

Departing from this curriculum review, the research team developed a framework for advancing justice-oriented data science education in an organic, parallel process before and alongside our workshops with teachers. First, before the workshops with teachers began, we analyzed (and continue to iteratively review) the relevant literature and research about data science education. Next, the research team members collected our own individual reflections about equity-centered data science education and discussed our ideas with each other. These steps informed the creation of the framework itself.

To represent the rich intersection of justice issues, data science (DS), and interdisciplinary inquiry, this framework calls out five key areas that collectively form a dynamic space for exploring justice-oriented data science (JODS). These areas in this framework are briefly described below, with pertinent literature listed in Table 1:

1. *DS practices*: Learners work with data in authentic ways, including wrangling data, making data moves, generating data representations, and interpreting findings.
2. *Disciplinary inquiry with DS*: Learners engage in meaningful disciplinary or interdisciplinary inquiries in which they pose questions and tackle them by analyzing data, while interacting and communicating with others.
3. *Examining justice through DS*: Learners develop their understanding of a range of justice issues (e.g., racial, climate) and their intersections through data investigations; learners mobilize DS to develop tools to tackle justice issues.
4. *Critical reflection on DS*: Learners consider the nature of DS as a field of research and practice, the ways in which fairness and biases are reflected in DS, and connections between DS and societal discourse.
5. *Identity and cultural practices*: Learners see themselves as people who can use data for purposes that interest them, recognize connections between DS and themselves and their communities, and identify ways to engage in DS in culturally congruent manners.

As illustrated in Figure 1, these areas intersect each other and do not have clear boundaries, in alignment with visions of justice and equity as multiperspective and heterogeneous (Bang & Vossoughi, 2016; Gutiérrez & Jurow, 2016). By referring to them as *areas* and illuminating their overlaps, we highlight multiple entry points to JODS. We imagine a JODS learning experience to be dynamic and emergent. A learner could enter the JODS space from one particular area (such as climate change) before connecting to overlapping areas (e.g., climate justice) and then expanding again into another area (e.g., data investigations of heatwaves in neighborhoods with different socioeconomic status).

This framework is still in its infancy and needs to be examined and refined in authentic empirical contexts. We begin this process in this study by reporting findings from the second participatory design workshop of the project. Before this workshop, one workshop had already taken place, with a primary focus on discussing what data science is and sharing its connections to equity and the teachers' current classroom practices. It was during the second workshop that we shared the JODS framework with the teachers and engaged in an emergent conversation about

it. By examining the second workshop, we asked: *In which ways were the framework areas reflected in the participatory design workshop's conversation?*

Methods

Participants included three secondary school teachers, one teaching science and two teaching social studies, and four university researchers who actively participated in the workshops. The teachers work in the same school district and are all keen on integrating data in curricula, had prior experience in doing so, and expressed varied comfort levels with data science. The data for this study is the transcript from the workshop, conducted via Zoom for 180 minutes. The workshop covered the goals of the DataX project, definitions of data science, reflections on the JOTS framework, discussions of several chapters of Hammond's (2014) book on CRT, individual work time envisioning justice-oriented data science learning activities, and pilot testing of the DataX software. The workshop was recorded, auto-transcribed by Zoom, and reviewed and revised to clean up transcription errors.

We conducted a content analysis of the transcript by using the individual framework areas and this project's design components (curriculum, pedagogy, and platform) as deductive codes (Miles et al., 2018). Using these established codes, two coders independently analyzed one excerpt from the transcript (appx. 10 minutes), debriefed with the research team, and clarified code interpretations. Then they independently coded the remainder of the dataset and cross-checked each other for consistency. Discourse content could be coded as multiple areas. For example, one teacher participant's talk about investigating a dataset from the U.S. colonial times of lists of passengers traveling to New England was coded as both *Disciplinary inquiry with DS* and *Justice inquiry with DS*; this segment was also coded as *Curriculum - Social Studies* in terms of design component. To further examine the context in which different areas were discussed during the workshop, we constructed a network of codes based on their proximity in the coded transcripts. Two codes were connected if their distance in the transcript was within 300 characters (equivalent to one discourse unit in the dataset) which would include the cases of partial and perfect overlaps.

Findings

First, the content analysis found all five areas of the framework represented in the transcript of the workshop conversation. Some areas were more dominant than the others. For instance, *Identity and cultural practices* was the most discussed area, appearing 27 times and accounting for 13.63% of the transcript. *Disciplinary inquiry with DS* (12.13%) and *Justice inquiry through DS* (11.55%) followed closely. The other two areas, *Critical reflection on DS* (3.01%) and *Data practices* (1.52%) were the least represented in the transcript.

Second, these framework areas were brought up in connection with each other and with the project's design components. Figure 2 shows the network of codes constructed based on proximity. The width of an edge indicates the number of adjacent appearances. As shown in Figure 2, *Disciplinary inquiry with DS* and *Justice inquiry through DS* frequently appeared near

each other, so did *Critical reflection on DS* and *Identity and cultural practices*. These framework areas were also discussed in connection with this project's design components, including the curriculum, pedagogy, and technology. For instance, *Identity and cultural practices* appeared close to *DataX pedagogy*; all five framework areas were connected with discussions of the justice aspect of our curriculum.

To illustrate the ways in which the framework areas were discussed, Table 2 shows an excerpt from the transcripts that was condensed for clarity. In a six-minute segment actively participated by two teachers and three researchers, all of the framework areas were touched. More importantly, they naturally transitioned to each other as teachers operated in a “design mode,” interpreting these areas and making rich connections with their own practice to advance DataX's design components.

Significance

Data science education is emerging in K-12 education and is informed by scholarship in statistics education, STEM education, learning sciences, literacies, and so forth. While it is widely accepted that data science provides a rich opportunity for authentic engagement with real-world problems, K-12 educators need concrete guidance to support their creative exploration of the nexus of data science, disciplinary learning, and justice. By introducing the JODS framework, we highlight five key areas and their rich overlaps. Analysis of a design workshop with high school teachers demonstrated the presence of these areas and their rich connections in teachers' thinking. The framework offers educators, designers, and researchers a means to systematically integrate disciplinary learning, data science practices, and justice issues when designing data science learning opportunities. Future work will seek to refine the proposed framework and enact it in curriculum designs.

Word count: 2000

References

- Atenas, J., Havemann, L., & Timmermann, C. (2020). Critical literacies for a datafied society: Academic development and curriculum design in higher education. *Research in Learning Technology*, 28. <https://doi.org/10.25304/rlt.v28.2468>
- Bang, M., & Vossoughi, S. (2016). Participatory Design Research and Educational Justice: Studying Learning and Relations Within Social Change Making. *Cognition and Instruction*, 34(3), 173–193. <https://doi.org/10.1080/07370008.2016.1181879>
- Bhargava, R., Kadouaki, R., Bhargava, E., Castro, G., & D'Ignazio, C. (2016). Data Murals: Using the Arts to Build Data Literacy. *The Journal of Community Informatics*, 12(3).
- Biehler, R., De Veaux, R., Engel, J., Kazak, S., & Frischemeier, D. (2022). Editorial: Research on data science education. *Statistics Education Research Journal*, 21(2), 1. <https://doi.org/10.52041/serj.v21i2.606>
- Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209–227. <https://doi.org/10.1007/s10676-013-9321-6>
- DiSalvo, B., Yip, J., Bonsignore, E., & DiSalvo, C. (Eds.). (2017). *Participatory Design for*

Learning. Routledge.

- Erickson, T., Wilkerson, M., Finzer, W., & Reichsman, F. (2019). Data Moves. *Technology Innovations in Statistics Education*, 12(1). <https://escholarship.org/uc/item/0mg8m7g6>
- Finzer, W. (2013). The Data Science Education Dilemma. *Technology Innovations in Statistics Education*, 7(2).
- Green, B. (2021). Data Science as Political Action: Grounding Data Science in a Politics of Justice. *Journal of Social Computing*, 2(3), 249–265. <https://doi.org/10.23919/JSC.2021.0029>
- Gutiérrez, K. D., Baquedano-López, P., & Tejada, C. (1999). Rethinking diversity: Hybridity and hybrid language practices in the third space. *Mind, Culture, and Activity*, 6(4), 286–303. <https://doi.org/10.1080/10749039909524733>
- Gutiérrez, K. D., & Jurow, A. S. (2016). Social Design Experiments: Toward Equity by Design. *Journal of the Learning Sciences*, 25(4), 565–598. <https://doi.org/10.1080/10508406.2016.1204548>
- Hammond, Z. L. (2014). *Culturally Responsive Teaching and The Brain: Promoting Authentic Engagement and Rigor Among Culturally and Linguistically Diverse Students* (1st edition). Corwin.
- Hayashi, C. (1998). What is Data Science? Fundamental Concepts and a Heuristic Example. In C. Hayashi, K. Yajima, H.-H. Bock, N. Ohsumi, Y. Tanaka, & Y. Baba (Eds.), *Data Science, Classification, and Related Methods* (pp. 40–51). Springer Japan. https://doi.org/10.1007/978-4-431-65950-1_3
- Irgens, G. A., Knight, S., Wise, A. F., Philip, T. M., Olivares, M. C., Vakil, S., Marshall, J., Parikh, T., Lopez, M. L., Wilkerson, M. H., Gutiérrez, K., Jiang, S., & Kahn, J. B. (2020). *Data Literacies and Social Justice: Exploring Critical Data Literacies through Sociocultural Perspectives*.
- Jiang, S., & Kahn, J. (2020). Data wrangling practices and collaborative interactions with aggregated data. *International Journal of Computer-Supported Collaborative Learning*, 15(3), 257–281. <https://doi.org/10.1007/s11412-020-09327-1>
- Kahn, J. (2020). Learning at the Intersection of Self and Society: The Family Geobiography as a Context for Data Science Education. *Journal of the Learning Sciences*, 29(1), 57–80. <https://doi.org/10.1080/10508406.2019.1693377>
- Ladson-Billings, G. (2021). Three Decades of Culturally Relevant, Responsive, & Sustaining Pedagogy: What Lies Ahead? *The Educational Forum*, 85(4), 351–354. <https://doi.org/10.1080/00131725.2021.1957632>
- Lee, H., Mojica, G., Thrasher, E., & Baumgartner, P. (2022). Investigating data like a data scientist: Key practices and processes. *Statistics Education Research Journal*, 21(2), 3. <https://doi.org/10.52041/serj.v21i2.41>
- Lee, V. R., & Delaney, V. (2021). Identifying the Content, Lesson Structure, and Data Use Within Pre-collegiate Data Science Curricula. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-021-09932-1>
- Lee, V. R., Pimentel, D. R., Bhargava, R., & D'Ignazio, C. (2022). Taking data feminism to school: A synthesis and review of pre-collegiate data science education projects. *British Journal of Educational Technology*, bjet.13251. <https://doi.org/10.1111/bjet.13251>
- Lee, V. R., Wilkerson, M. H., & Lanouette, K. (2021). A Call for a Humanistic Stance Toward K–

- 12 Data Science Education. *Educational Researcher*, 0013189X211048810. <https://doi.org/10.3102/0013189X211048810>
- Linardi, S., Gradeck, R., Idoko, R., & Say, B. (2022, April 8). *Developing a Data Science for Social Justice Curriculum*. <http://d-scholarship.pitt.edu/43195/>
- National Academies of Sciences, Engineering, and Medicine. (2018). *Data Science for Undergraduates: Opportunities and Options*. The National Academies Press. <https://doi.org/10.17226/25104>
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- Pangrazio, L., & Selwyn, N. (2021). Towards a school-based 'critical data education.' *Pedagogy, Culture & Society*, 29(3), 431–448. <https://doi.org/10.1080/14681366.2020.1747527>
- Philip, T. M., Schuler-Brown, S., & Way, W. (2013). A Framework for Learning About Big Data with Mobile Technologies for Democratic Participation: Possibilities, Limitations, and Unanticipated Obstacles. *Technology, Knowledge and Learning*, 18(3), 103–120. <https://doi.org/10.1007/s10758-013-9202-4>
- Philip, T. M., & Sengupta, P. (2021). Theories of learning as theories of society: A contrapuntal approach to expanding disciplinary authenticity in computing. *Journal of the Learning Sciences*, 30(2), 330–349. <https://doi.org/10.1080/10508406.2020.1828089>
- Rubin, A., & Mokros, J. (2018). Data clubs for middle school youth: Engaging young people in data science. In *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan*. (pp. 1–5). International Statistical Institute.
- Sander, I. (2020). Critical big data literacy tools—Engaging citizens and promoting empowered internet usage. *Data & Policy*, 2. <https://doi.org/10.1017/dap.2020.5>
- Shen, H., DeVos, A., Eslami, M., & Holstein, K. (2021). Everyday Algorithm Auditing: Understanding the Power of Everyday Users in Surfacing Harmful Algorithmic Behaviors. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 433:1-433:29. <https://doi.org/10.1145/3479577>
- Shreiner, T. L. (2019). Students' use of data visualizations in historical reasoning: A think-aloud investigation with elementary, middle, and high school students. *The Journal of Social Studies Research*, 43(4), 389–404. <https://doi.org/10.1016/j.jssr.2018.11.001>
- Skovsmose, O. (2012). *An Invitation to Critical Mathematics Education*. Springer Science & Business Media.
- Tang, R., & Sae-Lim, W. (2016). Data science programs in U.S. higher education: An exploratory content analysis of program description, curriculum structure, and course focus. *Education for Information*, 32(3), 269–290. <https://doi.org/10.3233/EFI-160977>
- Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 4(2), 2053951717736335. <https://doi.org/10.1177/2053951717736335>
- Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology*, 25(1), 127–147. <https://doi.org/10.1007/s10956-015-9581-5>
- Wilkerson, M. H., & Polman, J. L. (2020). Situating Data Science: Exploring How Relationships

to Data Shape Learning. *Journal of the Learning Sciences*, 29(1), 1–10.
<https://doi.org/10.1080/10508406.2019.1705664>

Wright, P. (2016). Social justice in the mathematics classroom. *London Review of Education*.
<https://doi.org/10.18546/LRE.14.2.07>

Zieffler, A., Garfield, J., & Fry, E. (2018). What Is Statistics Education? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education* (pp. 37–70). Springer International Publishing. https://doi.org/10.1007/978-3-319-66195-7_2

Table 1. Description of areas of the framework

Areas	Description	Related literature
Data science practices	Learners work with data in authentic ways, including wrangling data, making data moves, generating data representations, interpreting findings, and so on.	<ul style="list-style-type: none"> • Data cycle (Finzer, 2013) • Data moves (Erickson et al., 2019) • Data wrangling practices (Jiang & Kahn, 2020) • Statistical thinking (Rubin & Mokros, 2018; Zieffler et al., 2018) • Data practices and processes (H. Lee et al., 2022)
Inquiry with data science	Learners engage in meaningful disciplinary or interdisciplinary inquiries in which they pose their own questions and answer these questions by analyzing data, while interacting and communicating with others.	<ul style="list-style-type: none"> • Mathematics and science (Skovsmose, 2012; Weintrop et al., 2016) • Historical reasoning (Shreiner, 2019)
Examining justice through data science	Learners develop their understanding of a range of justice issues (e.g., racial, climate) and their intersections through data investigations; learners mobilize data science to develop tools to tackle justice issues.	<ul style="list-style-type: none"> • Critical mathematics (Skovsmose, 2012) • Social justice in math (Wright, 2016) • Data feminism (V. Lee et al., 2022) • Critical data education (Pangrazio & Selwyn, 2021)
Critical reflection on data science	Learners consider the nature of data science as a field of research and practice, the ways in which fairness and biases are reflected in data science, and connections between data science and societal discourse.	<ul style="list-style-type: none"> • A contrapuntal approach to learning (Philip & Sengupta, 2021) • Critical “big data” literacy (Atenas et al., 2020; Sander, 2020) • Algorithmic bias, auditing, accountability (Bozdag, 2013; Shen et al., 2021)
Identity and cultural practices	Learners see themselves as people who can use data for purposes that interest them, recognize connections between data science and themselves and their communities, and identify ways to engage in data science in culturally congruent manners.	<ul style="list-style-type: none"> • Culturally responsive teaching (Hammond, 2014; Ladson-Billings, 2021) • Democratic participation, self identity, family history (Kahn, 2020; Philip et al., 2013, p. 2013; Wilkerson & Polman, 2020) • Hybrid language practices (Gutiérrez et al., 1999) • Arts and data (Bhargava et al., 2016)

Table 2. A condensed excerpt from the transcript.

Line	Time	Transcript	Code
464-466	1:14:12-1:14:36	R1: The idea of data science might be intimidating to students. How can we build up their confidence?	Critical Reflection on DS Identity and Cultural Practices
467-470	1:14:46-1:15:30	R2: How do we build self-efficacy and overcome reifying beliefs like “I’m not good at math”, or “I don’t understand graphs” so that students believe they can be successful?	DataX Pedagogy Identity and Cultural Practices
471-474	1:15:41-1:16:24	R1: And how do we include the idea that oral stories can be more valued in different cultures? Maybe it pushes us to reconsider what counts as data?	Identity and Cultural Practices Critical Reflection of DS
475-481	1:16:27-1:17:16	T1: I have an activity where students create data about their favorite superheroes and try to prove which one is best as a low-barrier entry.	DataX Pedagogy DataX Curriculum Identity and Cultural Practices
487-488	1:18:15-1:19:12	T2: I have used maps with data as a different access point.	DataX Curriculum Disciplinary inquiry with DS
489-490	1:19:12-1:19:43	T1: I wonder if it would be possible to have students find bias in their own work or others and see how it could be improved.	Justice Inquiry through DS Critical Reflection on DS
491-492	1:19:43-1:20:31	R3: I want to call out these great connections between critical reflection and data science; what is collected, what is data, who gets to make those decisions?	Justice Inquiry through DS Critical Reflection on DS
493-499	1:20:31-1:21:30	R3: When I think of the anecdotes that [T1] and [T2] have shared and the discussions they are having with students, it seems like you could take a question that is already sparking some kind of argument in the classroom and work at it in some way that really feels like it's covering a lot of the things that we care about.	Disciplinary Inquiry with DS

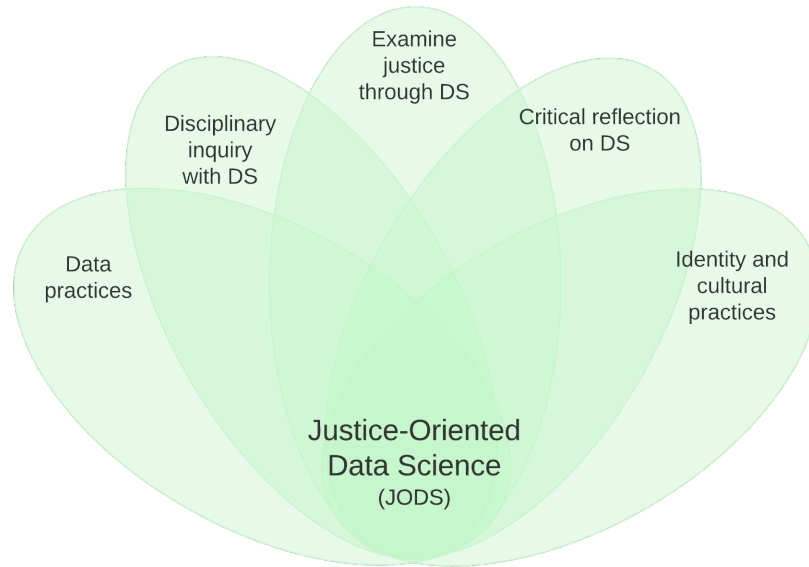


Figure 1. Illustration of the framework for justice-oriented data science education (JODS).

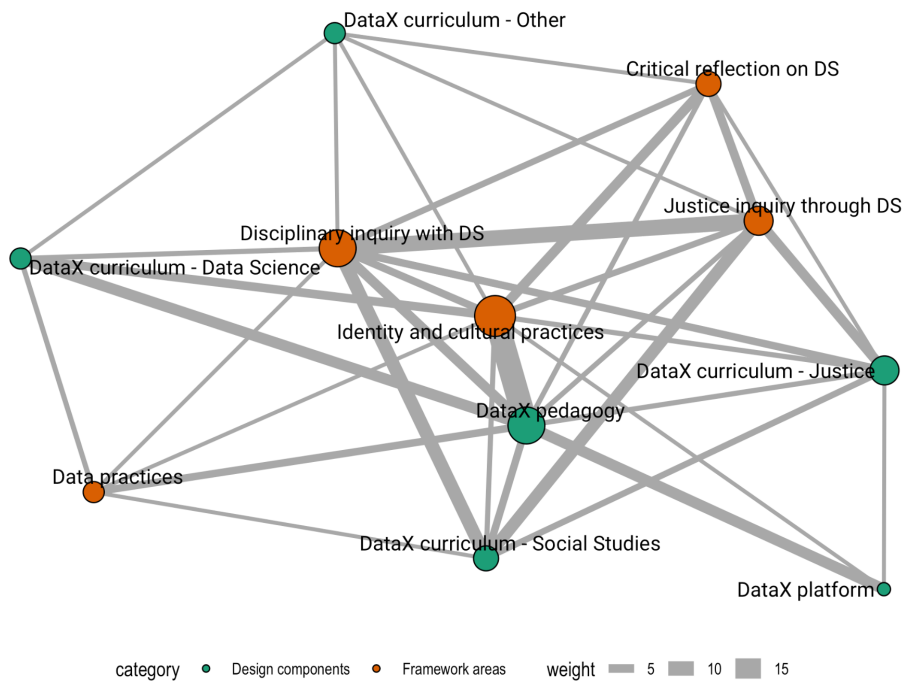


Figure 2. Network of the codes based on their proximity in the transcripts.