Nichols, T.P. (forthcoming). Data literacies. In M. Winn & L. Winn (Eds)., *Bloomsbury handbook of social justice in education*. Bloomsbury Academic.

Data Literacies

T. Philip Nichols
Baylor University

Data literacies are practices that reflect or cultivate an individual's understanding and control of datafied environments. Though the term is occasionally used to refer to personal competencies for working with data (i.e., analyzing datasets), it is more commonly theorized as a strategic or pedagogical response to increased "datafication" in society – that is, the technology-aided translation of social activities into extractable data that can be used in decision-making and governance.

The spread of datafication into more and more aspects of school, work, home, and civic life, and the uneven social impacts that result, have made "data literacy" an emerging priority among justice-oriented educators. While there are variations in how the concept is mobilized, there is growing consensus that critical perspectives on data, and the conditions of its production and use, have an important place in equitable teaching and learning.

Datafication and Social (In)Justice

The practices associated with data literacies are most easily understood in the context of the larger phenomenon they are meant to address: datafication. Datafication is not new; but rather, it is an outgrowth of a longstanding impulse toward quantifying social existence so as to render it legible for comparison, classification, and control. Philosopher Ian Hacking (1991) argues that the emergence of statistical probability in the seventeenth century gave rise not only to the idea of a mathematical "norm" (and, by extension, the possibility of deviating from it), but also a view of the world as something amenable to calculated prediction. Such innovations have since been applied both to innocent and perverse ends – from weather forecasting to eugenics and capitalist extortion – making 'quantification' less a neutral computational practice than an epistemic shift with significant stakes for social (in)justice.

Datafication is an extension of quantification that uses digital technologies to amplify the scale and speed at which social interactions are coded and processed as "data." The spread of mobile and connective media (e.g., smartphones, laptops, tablets) now makes it possible to quickly accrue massive stores of information about all aspects of life – including many once seen as too complex to quantify. Every click, swipe, keystroke, and login produces data trails that can be accumulated into larger datasets and mined for patterns about users' identities, interests, relationships, anxieties, movements, and consumption habits. Sometimes this information is applied directly back to individual users; for instance, by sending them targeted advertisements

for products they've browsed. However, more often it is pooled together with others', allowing data owners to glean (and sell) population-level insights about the behaviors and preferences of large groups of people – often along lines of race, gender, income, and other formations of difference. Such analyses are tremendously valuable not only to those trying to market products, but also for those looking to win elections or sway public opinion, and to govern or police communities.

As the driving current for producing and parsing these large datasets, datafication has profound implications for civic and planetary wellbeing. Crucially, it is not just the *outcomes* of datafication that are bound up with concerns for social justice, but also the *process* by which digital technologies harvest and mobilize data. Three features of this process are especially salient in "data literacies" education: reduction, abstraction, and individualization (cf. Pangrazio et al., in press).

Reduction involves the condensing of complex phenomena into a form that can be shared, compared, or stored. It is a necessary step in any kind of quantification – one that trims away the excesses of reality to make it comport with available systems of classification. While the term can carry a negative connotation, it is not, by itself, sinister. With no way to grasp the totality of a subject, humans always rely on reduction to produce and order knowledge out of partial information. This becomes problematic, however, when reductions are mistaken for totalities. In education, for instance, when attendance and test performance data are used as proxies for "learning," these reductions can be reified over time, nudging teachers and students toward behaviors that align with one narrow sense of "learning" at the expensive of others. Reductions, then, are political because they always involve judgment: what gets included/excluded; and how will it be simplified, represented, and circulated. In a datafied society, such reductions can have impacts not only on educational equity, but also on who qualifies for, or has access to, insurance, healthcare, social welfare programs, and, in some cases, even fundamental civil rights (cf. Eubanks, 2018).

Abstraction involves identifying and interpreting patterns within and across the reductions in a dataset. Because of the scale and speed at which datafication occurs, there is virtually no limit to the kinds of abstractions that can be culled from digitally mediated activity. Between the data that users voluntarily share, the metadata about their usage, and the analytic data that can be generated by feeding this information through probability-based algorithms, data-owners are able to detect all kinds of correlations which might offer insights into users' identities and behaviors. However, because these inferences are based on reductions, they are not always accurate. Abstractions formed from weak associations, for instance, can lead to users being served advertisements for products that are comically ill-suited to their needs, or that they have already purchased. Sometimes, though, the results are more troubling. Scholars have demonstrated that abstractions from datasets regularly reproduce the dominant norms of white, capitalist, heteropatriarchy – for instance, by delivering search results that reinforce racialized stereotypes or that nudge users toward reactionary content (Noble, 2018). Even though datafication is

premised on its potential for algorithmic objectivity, in other words, its methods for abstracting insights from data are outgrowths of human decision-making and, as such, reflect the biases, ideologies, and power asymmetries of society writ large.

The third feature of datafication, individualization, refers to the cycle by which patterns derived from reduced and abstracted data are then filtered back onto individuals – coaxing or conditioning their attention, attitudes, or activities in particular ways. For data owners, the value of datafication is not just the population-level insights that can be inferred from large data-streams, but the capacity of these insights to be mobilized toward different ends. Targeted advertisements, for instance, can be used to boost product sales; and personalized newsfeeds can be used to reinforce or elide particular political narratives. But individualization is also key to the spread of data systems for more dramatic interventions in public and private life. The rise of "smart cities" and "smart homes," for instance, promise to act on individualized data in ways that conserve energy, reduce inefficiencies, and protect property. This drive is also behind the use of "predictive analytics" in public institutions. In law enforcement, such technologies use abstracted insights from demographic and behavioral data to surveil and police communities that have been algorithmically classified as "at risk;" and in schools, similar methods are now used to track student progress and anticipate disciplinary concerns. Scholars have documented the uneven impacts of such systems and, particularly, the deleterious consequences for communities of color – where individualization has led to heightened surveillance, policing, and discipline (Benjamin, 2019). This is further evidence that the implications of datafication for social (in)justice inhere not just in its outcomes, but also in the guiding logics – reduction, abstraction, and individualization – that underwrite the translation of social activity into actionable data.

Data Literacies

"Data literacies" is an umbrella term for the skills and practices that people use to understand, analyze, control, or assert agency in systems and processes of datafication. As with parallel concepts like "digital literacies" and "media literacies," there is no uniform sense of what data literacies are, much less how they should be cultivated or taught; however, approaches to data literacies in education can be broadly seen as falling into three interrelated categories: reflexive, critical, and interventionist (Pangrazio & Sefton-Green, 2020).

Reflexive approaches to data literacies tend to focus on developing learners' awareness of datafied systems so they can make more informed, intentional, and responsible choices related to privacy and information-seeking. They tend to emphasize the "individualized" features of datafication – that is, the forms of personal data that systems collect and the personalized ways such systems apply data. This is the orientation most common in school curricula related to data and "digital citizenship." In such programs, students are instructed in how to properly manage private data (e.g., selecting appropriate passwords, being judicious in what information they share, managing their "digital footprint" online) and how to navigate data-driven information

environments (e.g., recognizing targeted advertisements, mitigating against "filter bubbles" in personal newsfeeds). The most sophisticated versions also cultivate more explicit literacy practices, like "reading" the data-streams that online activities produce or altering privacy settings to avoid "writing" data one might wish to keep private. Pedagogically, reflexive orientations to data literacies have found the most traction in formal school settings, in part, because the focus on personal responsibility lends itself to structured and sequenced lessons that can be taught, practiced, and assessed. However, such approaches are not always able to engage the deeper societal impacts of datafication that extend beyond individuals alone.

Critical approaches to data literacies expand the term to include not just awareness of data systems and the skills needed for responsibly navigating them but also critical analysis of their underlying power relations. Drawing on traditions of "critical pedagogy" and especially "critical literacies" – the reading and re-writing of the word and world in ways that confront, resist, or upend power hierarchies – critical data literacies interrogate how the processes of reducing, abstracting, and individualizing data might be used to perpetuate (or address) ongoing social injustices (Nichols et al., 2021). This includes, for instance, nurturing meta-awareness of the ways that data technologies encode logics of capitalist extraction or racialization into ostensibly neutral educational "tools," like search engines, classroom apps, or learning analytics platforms. Scholars suggest that such practices can help educators become more "racially literate" about data, and "data literate" about race (Philip et al, 2016) – an orientation with implications not only for classroom-level teaching and learning, but also for how school-, district-, and state-level administrators can identify the ideologies that underpin growing imperatives for data-driven decision-making in public education. Critical approaches to data literacies also emphasize not just critical *readings* of datafication, but critical *writings* as well. Researchers have demonstrated, for example, that attunement to the workings of existing data systems can inspire young people to imagine and create alternate data processes that center the experiences and interests of nondominant communities (Stornaiuolo, 2020). While critical pedagogies like these are less common in schools than their reflexive counterparts, they have gained increased attention among researchers and practitioners. This suggests that, as challenging as critical data literacies are to integrate in the curriculum, their stakes for social justice are becoming too great to ignore, or to relegate merely as matters of personal responsibility.

Finally, a third category of data literacies focuses on what can be called interventionist tactics – that is, practices that go beyond analyzing or critiquing the processes and impacts of datafication to actually intervene or assert agency within large data systems. Such tactics often extend out of critical approaches to data literacy. For instance, attending to the ways raced, classed, and gendered identities are formulated in processes of data abstraction can be an important precursor to envisioning possible strategies for countering their latent ideologies. Interventionist data literacies along these lines might include identifying policy levers that could be used to regulate how certain kinds of data classifications are made and used, or devising workarounds or "hacks" that obstruct or redirect the design of such processes. Tactics like these

can be effectively mobilized at different scales of educational systems. Students, for example, can be taught to look for vulnerabilities where data systems that work against their interests or values might be subverted. Likewise, schools and districts can model collective responses that organize opposition to exploitative commercial data practices: for instance, by refusing to contract with data analytics providers whose standards for data collection and privacy are not transparent (Garcia & Nichols, 2021). Pedagogically, interventionist tactics are arguably the least defined of the different approaches to data literacies. This is, in part, because we have only recently begun to grapple with the scale and impacts of datafication, and as such, our methods for remaking these relations are relatively nascent.

Importantly, there are interdependencies across all three of these approaches to data literacies. While formal instructional programs may favor one, or a combination, they can also be seen as working in tandem. Just as interventionist tactics are often prefigured by critical analysis of datafication, so too can awareness and reflexivity anticipate, and extend from, critique and intervention. In this way, data literacies are better understood as a repertoire of practices for confronting datafication rather than a singular set of skills.

Challenges for Data Literacy Education

Despite the growing resonance of data literacies in justice-oriented education, several enduring challenges can complicate efforts to center such practices in teaching and learning. The most immediate of these difficulties is the mismatch in the scales at which data literacies education and datafication operate. As noted, pedagogies associated with datafication commonly prioritize individual dispositions and actions: nurturing responsibility in personal data use, for instance, in the reflexive tradition; or sharpening individuals' abilities to identify and demystify ideologies, in the critical approach. Powerful as these practices can be, the emphasis on individuals can direct attention away from those aspects of datafication that are not as easily grasped at the scale of singular persons or isolated actions. Datafication, after all, is interested in "individualization," but a form that is refracted through reduction and abstraction – decidedly collective processes. This leads to situations where individuals' data literacy practices can have simultaneously just and unjust implications in large data systems. We have seen, for instance, how the same data technologies that activists use to organize protests and document police brutality also generate and share geolocation and facial-recognition data that law enforcement agencies use to predict, surveil, and disrupt future protests (Brayne, 2020). Such entanglements between individual actions and collective consequences can be difficult to comprehend, let alone teach; as such, they remain a persistent challenge for data literacies pedagogies to confront.

A second, related challenge for data literacies pertains to the appropriateness of "literacies" as a framework for addressing datafication. Traditionally, the concept of "literacy" has been closely associated with the production and interpretation of texts – most commonly alphabetic print, though its meaning has also been extended to include other media forms like films and

video games. However, in adopting this terminology for addressing datafication, it is not immediately apparent what the "text" is that students are being taught to create, interpret, or remix (cf. Pangrazio & Sefton-Green, 2020). Is data itself the text? And if so, is it individual datapoints, or collective categorizations that are reduced or extracted from them? Questions like these are further complicated by the fact that data systems aren't fixed or stable. As we have seen, they recursively collect, analyze, produce, and apply data in ways that don't hold still long enough for conventional "reading" and "writing" practices to be neatly applied to them. This suggests that while the idiom of "literacy" may be the most convenient term currently available to us for teaching and learning about datafication, there may be need over time to develop more precise language and repertoires of practice that better address the relations among data systems, individuals, and society.

Conclusion

Even with its uncertainties related to terminology and pedagogy, "data literacies" has proven to be a generative concept for engaging with datafication and its evolving impacts on human flourishing. If, in time, the term is replaced by something less closely tied to "texts," its driving interest in reflecting on, critiquing, and intervening in data systems should remain a pressing concern for educators and students oriented toward social justice. As long as datafication continues its creep into public institutions and private life, it merits significant pedagogical attention. Likewise, while current approaches to data literacies education may be too narrowly focused on individual responsibility and action, the potential for building more collective responses to datafication may depend on the foundations that our present counter-strategies – reflexivity, critique, and intervention – have laid. What we now call "data literacies," in other words, can serve as a powerful precursor to other forms of pedagogy that not only confront the challenges of datafiction but engage us, collectively, in the work of imagining new data systems – ones rooted in justice and liberation rather than extraction and profit.

Key References

Benjamin, Ruha, *Captivating Technology: Race, Carceral Technoscience, and the Liberatory Imagination of Everyday Life* (Durham: Duke University Press, 2019).

Brayne, Sarah, *Predict and Surveil: Data, Discretion, and the Future of Policing* (New York: Oxford University Press, 2020).

Eubanks, Virginia, *Automating Inequality: How High-Tech Tools Profile, Policy, and Punish the Poor* (New York: St. Martin's Press, 2018).

Nichols, T. Philip et al., "Critical Literacy, Digital Platforms, and Datafication," in *The Handbook of Critical Literacies*, ed. Jessica Zacher Pandya et al. (New York: Routledge), 345-353.

Noble, Safiya, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018).

Pangrazio, Luci et al., "Datafication Meets Platformization: Materializing Data Processes in Teaching and Learning," *Harvard Educational Review*, (forthcoming).

Pangrazio, Luci and Julian Sefton-Green, "The Social Utility of Data Literacies," *Learning, Media, and Technology 45*, no. 2 (2020): 208-220.

Philip, Thomas, Maria C. Olivares-Pasillas, and Janet Rocha, "Becoming Racially Literate about Data and Data-Literate about Race: Data Visualizations in the Classroom as a Site of Racial-Ideological Micro-Contestations," *Cognition and Instruction 34*, no. 4 (2016): 361-388.

Stornaiuolo, Amy, "Authoring Data Stories in a Media Makerspace: Adolescents Developing Critical Data Literacies," *Journal of the Learning Sciences 29*, no. 1 (2020): 81-103.

Author

T. Philip Nichols, Department of Curriculum and Instruction, Baylor University, USA