
Understanding Human-Data Relationships: Data as Property

Casey Fiesler

Information Science

University of Colorado Boulder

casey.fiesler@colorado.edu

Jed R. Brubaker

Information Science

University of Colorado Boulder

jed.brubaker@colorado.edu

Abstract

Data is structured by governance systems—both technical and legal—that are limited in their ability to capture and codify the surrounding, evolving technology and social ecosystems. As a result, these governance systems often fail to reflect how end users think about online data. In this half of our argument, we focus on one view (particularly dominant in policy) of the relationship between people and their data is that of data as property. But how well does this really capture the way that people think about data? People’s understandings of their relationships to the content they create and their digital traces—including their rights and the rights of others—has important implications for design, policy, and data science research ethics.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s). CSCW '16, February 27 - March 02, 2016, San Francisco, CA, USA.

Author Keywords

Copyright; Data; Human-data relationships; Human-computer interaction; Law; Policy; User-generated content

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

Introduction

Both legal and technical systems have existing infrastructure in place for dealing with digital data. Historically, metaphors have been important in the construction of both, a necessity in the context of rapidly changing technology. Metaphors have always dominated user interfaces—desktops, windows, files, spreadsheets—a way to make abstract computing tasks more familiar [5]. Similarly, metaphor permeates legal discourse, allowing for the construction of legal realities and the application of existing policy to new contexts. Traditionally, property (of different types) has been the dominant metaphor for human-data relationships in these systems.

However, the property metaphors used by those creating and interpreting policy (or even by underlying technology architecture) may not be the same metaphors that *humans* use when thinking about either

the content they purposefully share or the digital traces they leave. For example, what model might someone have for their relationship to a tweet? Is it more like physical property similar to their car, or copyrighted like a song? Is their Twitter account more like real estate that they own or is it an extension of their identity? Knowing how people *actually* model human-data relationships is important for the construction of both technology and policy, but it is also critical to data scientists making decisions about how as researchers they should ethically treat data and content that comes from real people.

Background and Motivation

The law has had a fraught relationship with digital content, in large part due to the new challenges that technology often poses as the law struggles to keep up with these changes. Legal approaches to regulating digital content and the Internet often look a lot like the regulation of physical property. For many years (and still ongoing in some contexts), the law essentially transposed existing property concepts to the Internet [1,6,11,12]. However, these metaphors have been used inconsistently and imprecisely by courts, leading to confusion about relevant legal rights [11,12]. Because of this inconsistency, legal scholars have also argued that it would be beneficial for policies to be chosen based on user experiences of technology rather than metaphors of formal property-based constructs [13].

One concrete example for a legal treatment of data that may conflict with dominant perceptions is that of what happens to one's data after they die—an issue that is not currently well settled under current laws [19,20]. Typically framing the problem in terms of “assets,” the

legal system has treated this situation in a number of (not necessarily mutually exclusive) ways: allowing service providers to govern inheritance in private contracts such as Terms of Service [2], treating certain types of data as original work that is copyrighted and therefore inheritable as intellectual property [20], or allowing fiduciary access to an account (similar to a locksmith opening a door to a deceased's house) [19]. However, considering data as “assets” itself relies on the property metaphor. Brubaker's proposed stewardship model as an alternative to property for the management of post-mortem social media accounts therefore may not track well to existing legal frameworks [4].

Another reason we know that human understandings of their relationships with data do not always track to legal systems is prior work around misunderstandings of the law. Much of Fiesler's research has been around the impact of copyright law in online communities, and in part has involved user conceptions of ownership. In studies of online content creators, she found that understandings of law related to content ownership and re-use do not actually track well to the law, but more to social norms and ethical intuitions [7], and that misunderstandings and conflicts in rules lead to chilling effects or other negative outcomes [8]. An analysis of copyright licenses in Terms of Service for user-generated content and social media sites reveals that people do not necessarily have good models about what rights a website has to their content [9]. In sum, her work has revealed that people's intuitions about how the law handles ownership issues with respect to online content often does not match either the actual law or instantiation in website policies.

Similarly, other work within HCI and CSCW has revealed that the ways people understand the *ownership* of digital things is complex. For example, Odom et al. investigated how people think about their possessions moving from the physical to the digital [17,18]. Their data shows confusion about what people own or don't own, as well as relevant rights such as giving access to other. In the context of social data, Marshall and Shipman have examined how social media users think about the ownership of their content, across different websites; their work shows that social media users do *not* have clear, consistent ideas about their relationships to their data [14–16]. We propose that these complexities are representative of a larger set of divergent understandings of human-data relationships.

Ongoing Work

As follow-up to Fiesler's prior work on understandings of copyright law and Brubaker's prior work on understandings of post-mortem data [4,7,8,9], we are conducting studies examining human-data relationships more generally around social data, including an interview study and a historical examination of policy and technical architectures governing these relationships. Below, we discuss in more detail a survey study intended to illuminate attitudes about rights related to social media content.

Fiesler's prior work shows that people may not have a good idea about the rights that social media sites (e.g., Twitter) have in their content, and Marshall & Shipman have shown similar confusions around what re-use rights other people might have in that content [9,15,16]. We also know from prior work that people tend to have poor mental models of their social media audiences [3]. In order to better understand the

nuances of how people feel about certain uses—by websites, by other users, by journalists, and by *researchers*—we are conducting a survey targeting specific Twitter content.

Participants are recruited from Twitter and asked questions about a *specific* piece of content. For one particular tweet, we ask how they feel about certain uses of that tweet. What if another user tweeted the text without attributing it to you? What if it appeared in an image meme? What if it appeared in a BuzzFeed article? What if a researcher scraped this tweet and used it in aggregate data analysis? What if the tweet itself appeared in a research paper? Would you feel differently about any of these if you were asked permission? Do you think of this content as your *property*? Do your feelings about these uses differ than what you think your *legal rights* in preventing such uses would be?

One goal of this research is to better situate the content that data scientists study in context. Public discourse around the Facebook emotional contagion study [10] along with others suggests that we need a more nuanced understanding of how people perceive their own relationships with data and how they feel about that data being studied. More nuanced understandings of human-data relationships (using metaphors of property or otherwise) are essential to the context in which data scientists work.

References

1. Shyamkrishna Balganesh. 2006. Common Law Property Metaphors on the Internet: The Real Problem with the Doctrine of Cybertrespass. *Michigan Telecommunications and Technology Law Review* 12, 265–333.

2. Natalie M. Banta. 2014. Inherit the Cloud: The Role of Private Contracts in Distributing or Deleting Digital Assets at Death. *Fordham Law Review* 83, 799.
3. Michael S Bernstein, Eytan Bakshy, Moira Burke, Brian Karrer, and Menlo Park. 2013. Quantifying the Invisible Audience in Social Networks. 21–30.
4. Jed R. Brubaker, Lynn S. Dombrowski, Anita M. Gilbert, Nafiri Kusumakaulika, and Gillian R. Hayes. 2014. Stewarding a legacy: responsibilities and relationships in the management of post-mortem data. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*.
5. Wendy Hui Kyong Chun. 2011. *Programmed Visions: Software and Memory*. MIT Press, Cambridge, MA.
6. Julie E. Cohen. 2007. Cyberspace As/And Space. *Columbia Law Review* 107, 210–256.
7. Casey Fiesler and Amy S. Bruckman. 2014. Remixers' understandings of fair use online. *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*.
8. Casey Fiesler, Jessica Feuston, and Amy S. Bruckman. 2015. Understanding Copyright Law in Online Creative Communities. *Proceedings of the 2015 ACM Conference on Computer-Supported Cooperative Work and Social Computing - CSCW '15*.
9. Casey Fiesler, Cliff Lampe, and Amy S. Bruckman. 2016. Reality and Perception of Copyright Terms of Service for Online Content Creation. *CSCW '16: Proceedings of the 2016 ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*.
10. A.D.I. Kramer, J.E. Guillory, and Jeff T. Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24, 8788–8790.
11. Jacqueline D. Lipton. 2003. Mixed Metaphors in Cyberspace: Property in Information and Information Systems. *Loyola University Chicago Law Journal* 35, 235–274.
12. Jessica Litman. 2006. *Digital Copyright*. Prometheus Books.
13. Michael J Madison. 2002. Rights of Access and the Shape of the Internet. *Boston College Law Review* 44, 2, 433–507.
14. Catherine C Marshall and Frank M Shipman. 2011. The Ownership and Reuse of Visual Media. *Proceedings of the 11th ACM/IEEE-CS Joint Conference on Digital Libraries*, 157–166.
15. Catherine C Marshall and Frank M Shipman. 2011. Social Media Ownership: Using Twitter as a Window onto Current Attitudes and Beliefs. *Proceedings of the ACM conference on Human factors in computing systems - CHI '11*.
16. Catherine C Marshall and Frank M Shipman. 2015. Exploring the Ownership and Persistent Value of Facebook Content. *Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing - CSCW '15*.
17. William Odom, Abi Sellen, Richard Harper, and Eno Thereska. 2012. Lost in translation: understanding the possession of digital things in the cloud. *Proceedings of the ACM conference on Human factors in computing systems - CHI '12*.
18. William Odom and John Zimmerman. 2013. Fragmentation and transition: understanding perceptions of virtual possessions among young adults in Spain, South Korea and the United States. *Proceedings of the ACM conference on Human factors in computing systems - CHI '13*.
19. Suzanne B. Walsh. 2014. Coming Soon to a Legislature New You: Comprehensive State Law Governing Fiduciary Access to Digital Assets. *Charleston Law Review* 8, 429–449.
20. Ashley F. Watkins. 2014. Digital Properties and

Death: What Will Your Heirs Have Access to After
You Die? *Buffalo Law Review* 62, 194–235.