Foreword

Why Datafy Social Media?

With data about each of us having value to so many others (mostly to better sell us things—or refuse to sell us things if we look too “risky”), it seems inevitable that as social media matured, the various players already interested in our personal data would want to mine this type of information to add to the data they have already gathered. Noncommercial research into our connections, opinions, and personal situations as embedded in social media has only recently begun, however, as data mining and other analytical techniques weren’t originally designed for the messiness of social media and the malleable, episodic nature of much of its content. Commercial entities investigating social media in an attempt to “datafy” it usually keep their results proprietary, which doesn’t advance wider research and prevents outside oversight of data ownership and privacy issues. Not helpful.

Why try to abstract data and content from the World Wide Web, multiple social media platforms, microblogging sites, e-mail platforms, geolocation devices, and text and multimedia collections? Why analyze it, visualize it, and apply it to other situations? What do the researchers hope to learn?

The possibilities range widely, with almost as many options as motives for digging into the data. Commercial motives for evaluating data we already know fairly well:

- Selling more effectively by identifying the specific audience for a product
- Modifying a product to better fit consumer needs and wants
- Reducing risk by not selling (insurance, for example) or lending (mortgages) to people who fit an industry’s risk profile
- Developing long-term loyalty relationships with customers
- Fine-tuning tools to search and evaluate data in real time

What about noncommercial motives for this type of research?

- Tracking health emergencies and outbreaks
- Making sense of citizen science for migrations, population counts, temperature variations by time and location, etc.
- Identifying and correcting bottlenecks and flaws in networks or other technology services
- Evaluating the completeness of content collections
- Spotting cultural, political, and other trends as they develop
- Learning breaking news from the people observing it
Foreword

• Sharing key concepts and lessons teased out through analysis and presented using advanced visualization techniques
• Developing tools to search all types of data and metadata in real time and quickly analyze the likely meaning as it becomes visible

Both types of research have many more drivers—these short lists barely touch on the reasons researchers and their backers want to collect, analyze, and visualize data from social media networks and platforms. This book takes a closer look at the authors’ specific interests, providing their insights into a still new type of research that has many benefits—and risks—for all of us. Whether you agree with their conclusions, the datafication of social media continues. It behooves the worldwide community to keep up with what’s happening behind the scenes and speak up with ethical concerns, ideas for further research, and proposals for openness and privacy rules to get a handle on this exploding area of research that peers so closely into our personal and professional lives. Our social media interconnections mean that what researchers, governments, and industries do with our data has vital importance now and in the future.

Nancy Hays
EDUCAUSE, USA