Editorial
The First Piece of the Pattern

It is our great honor and privilege to welcome you to the very first issue of *Patterns*, Cell Press's new, open access journal of data science. The past 11 months have been an amazing time of work, talking to people, learning, and making decisions, and it's been thrilling.

At this stage, one could ask: why a new journal of data science? And why have you launched something that's so broad scope and, frankly, ambitious?

We're very glad you asked.

Data science can often feel like a new topic, a new domain springing out of computer science or mathematics, but in reality, it is a fundamental part of what it means to do research. Science has always been about data—after all, how can you test a hypothesis without making measurements, and how can you support a conclusion without showing your results? Only a few short decades ago, it took a great deal of effort to create data. Researchers would send out surveys, carefully design experiments, and spend long hours taking repeated and precious measurements.

How quickly things have changed. Now the average person in the street is generating data, often without being aware of it, simply by walking down the street with a smartphone and their geolocation tagged. For researchers, the ability to collect greater volumes of more detailed data about their subject of interest has come as a great blessing but has also caused problems.

As our ability to collect data grows, the ways we need to deal with it get more complicated, and the tools we need to understand it also become more detailed. If we are to make good decisions about the problems that we are facing as a species, whether that's climate change, pandemics, or even how we decide who gets a job or not, then we need good data, and we need to properly understand those data.

On the other side of all this new detail are the tremendous possibilities we have as a result of this data deluge. We can learn so much more about the structure of the universe, our environment, and also how our own bodies work, because we have access to new, better, and more detailed sources of data.

Patterns is about understanding—not only communicating between authors and readers but also breaking down those disciplinary boundaries to share the data science technologies that can be used to solve the problems that span domains. *Patterns* is about discussion, sharing opinions and ideas, and developing new standards of practice. *Patterns* is about community, bringing together people with similar interests, regardless of what their original domain is.

That is why, in this first issue, we bring together a wide range of authors to discuss an even wider range of topics from data-intensive research, computer science, and data curation. The thing they all have in common is data: how to use them, how to manage them, and how to use new technologies to learn from them.

The Research Articles are the core focus of *Patterns*—they take data science techniques and apply them to real-world problems to discover new knowledge. They span the entirety of the data-intensive research and data science space and cover topics as diverse as sentiment analysis of conservation studies (Van Houtan), intelligent electromagnetic sensing (Li), and how to identify data sharing and reuse (Khan).

Our Perspective pieces are a useful way of explaining, especially in such a fast-growing field as data science, the current situation with a given topic and what the authors think should be done to address the situation. In this issue, we deal with the thorny problem of information entropy (Habermann) and look at the potentials, possibilities, and complications of interdependent networks (Amini).

Because everyone creates and uses data, there are almost as many opinions about data as there are people. Opinion pieces are a way for the community to start discussions, to outline ideas and find collaborators, and to generally share information in a quick and easily understood way. This issue features a wide variety of opinions, ranging from top tips on how to use games to teach important data management concepts (McCutcheon) to how to tell if the next big technology is really suitable for the problem you want to solve (Crowcroft).

With Opinions, we also get into the philosophical aspects of data science, giving an example of how researchers can work across disciplinary boundaries to enrich a valuable (and irreplaceable) climate science database (Slonosky).

Fundamentally, *Patterns* is all about the people as well as the data. Data ethics is a topic we take very seriously, hence the Opinion piece "Who Should Do Data Ethics?" (Wylie). We are firm believers that everyone should do data ethics. Data have the potential to improve life for everyone, when they're used to empower and support communities (Cerit) or to help mitigate and recover from the effects of global pandemics (Perakslis).

Last, but by no means least, this issue looks even closer at the people at the cutting edge of data science, and how they are and were impacted by it. We take both a historical and contemporary perspective, identifying those who have come before us to lay the foundations of the data science field (Inman) and those who are working on the front lines of data science today (Gordon).

There are so many people we need to thank for their efforts and support in launching a new journal that is this broad and ambitious. From our wonderful, supportive colleagues at Cell Press to our amazing advisory board and the reviewers who kindly gave so much of their time and expertise to review for a journal that they didn't know. And of course, to the authors, whose desire to tell their data stories has meant that we have stories to tell and a beautiful array of different articles to publish.

A pattern is something that is repeated. We've laid the first piece with this first issue, and now it's time to see how *Patterns* will grow.

Thank you for sharing these data stories with us.

Sarah Callaghan and Tessa Darbyshire
Patterns Editorial Team
https://doi.org/10.1016/j.patter.2020.100020