

Models in the Wild

David Grant

MIT Philosophy

Our Topic for Today

How data models can **harm** others, and
what **you** can do about it

“We need every practitioner of **data science** to understand the potential impact of their work on society, to take responsibility for it, and to have enough understanding of the critical issues to be able to exercise good judgment.”

- Professor H.V. Jagadish,
University of Michigan

Overview and Lecture Goals

The goal for the day is to help you answer these questions for yourselves:

- **Predictive Modeling Activity:** what features make some colleges better than others?
- **Models in the Wild:** why should data scientists think about ethics as they conduct their research?
- **The Data Science Process:** how does ethics fit in to the data science process?
- **With Great Power...:** how much moral responsibility do data scientists have for what other people do based on their research?

Predictive Modeling Activity

What **features** make some colleges
better than others?

Your client: US News and Reports

US News' description of their college rankings:

“The host of intangibles that makes up the college experience can't be measured by a series of data points. But for families concerned with finding the best academic value for their money, the U.S. News Best Colleges rankings provide an excellent starting point for the search.

They allow you to compare at a glance the relative quality of institutions based on such widely accepted indicators of excellence as first-year student retention and graduation rates and the strength of the faculty. ...

Many factors other than those spotlighted here will figure in your decision, including location and the feel of campus life; the range of academic offerings, activities and sports; and cost and the availability of financial aid. But if you combine the information on usnews.com **http://usnews.com** with college visits, interviews and your own intuition, the U.S. News

rankings can be a powerful tool in your quest for the right college.”

Your project: generate better rankings

Your task is to begin the process of building a predictive model or models to help improve their rankings.

- The research question: “what features of colleges predict better outcomes for students?”
- Your task is come up with a list of candidate model inputs and outputs to guide data collection.

Your task for today

What do you want to predict or estimate? One half of the current the task is to help them decide what kinds of student outcomes to prioritize in the new rankings.

- The relevant outcomes should be both:
 - **Valuable**—the kinds of outcomes we care about when thinking about what college to attend.
 - **Available for analysis**—data on those outcomes could feasibly be collected by the organization.
- The outcomes you choose will be the target variables for subsequent analysis.

Your task for today

Which data are relevant? The other half is to use your current domain knowledge to brainstorm about which data might be relevant to predicting those outcomes.

- What input variables might be both:
 - **Relevant**—to predicting the outcomes you selected in the first phase?
 - **Available for analysis**—as before?

Your task for today

In the next ten minutes, work with your team to come up with a list of:

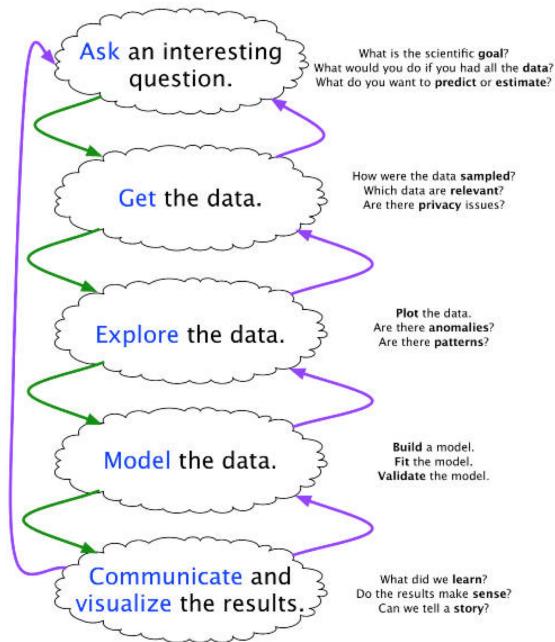
- **Two target variables** that you think are the most important empirical measures of good outcomes for students.
- **Five predictor variables** for those target variables (five total, not ten).

Your client will then—at considerable expense—acquire relevant datasets and provide those to you for analysis.

Models in the Wild

Why should **data scientists** think
about **ethics**?

The Data Science Process



Ethics and the data science process

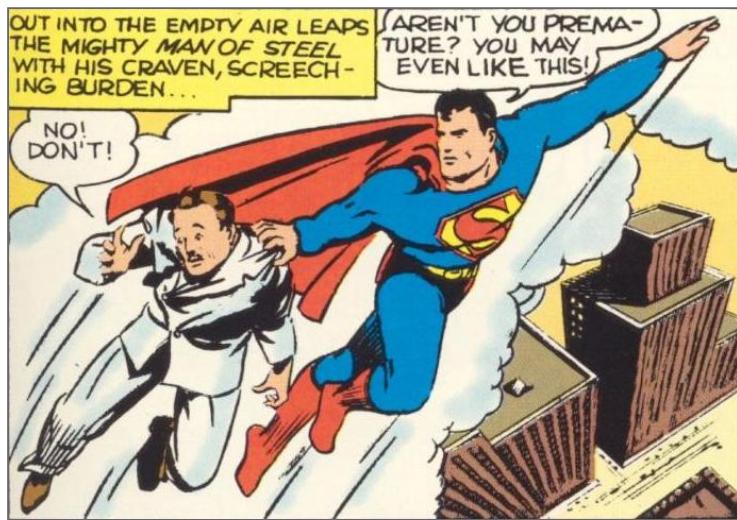
Claim: data scientists need to consider the ethical implications of their research at every stage of the data science process.

- Is this right?
- Should your team have talked more about ethical considerations during the team activity? If so, how might ethical considerations have been relevant to the choices you made?

Why integrate ethics into the data science process?

Three reasons to integrate ethics into the data science process:

- Data models are **powerful**.
- That power can be **destructive**.
- That destruction can be **invisible**.



Data Models are Powerful

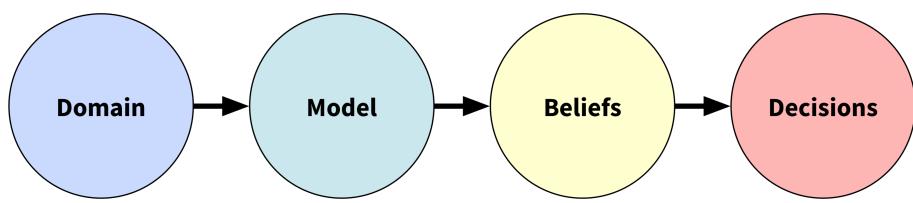
The social role of data scientists

What do data scientist do?

- Collect, manage, and analyze data.
- Develop models of empirical phenomena based on those data.
- Generate summaries of their findings.

For what purpose?

- To guide decision-making by others.



“The evidence is clear: Data-driven decisions tend to be better decisions. Leaders will either embrace this fact or be **replaced by others who do.**”

- Harvard Business Review

College rankings

The USN&R rankings have been extremely influential. They are used by:

- Students and families to make decisions about where to go to college.
- Colleges to decide how to allocate their resources.
- Alumni to make decisions about how much to donate.

Recidivism prediction models

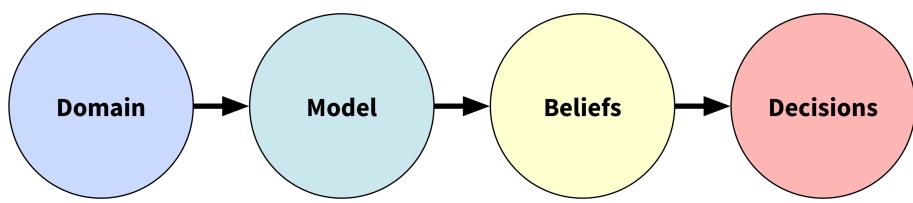
The LSI-R (Level of Service Inventory-Revised) is a survey given to convicted criminals across the US.

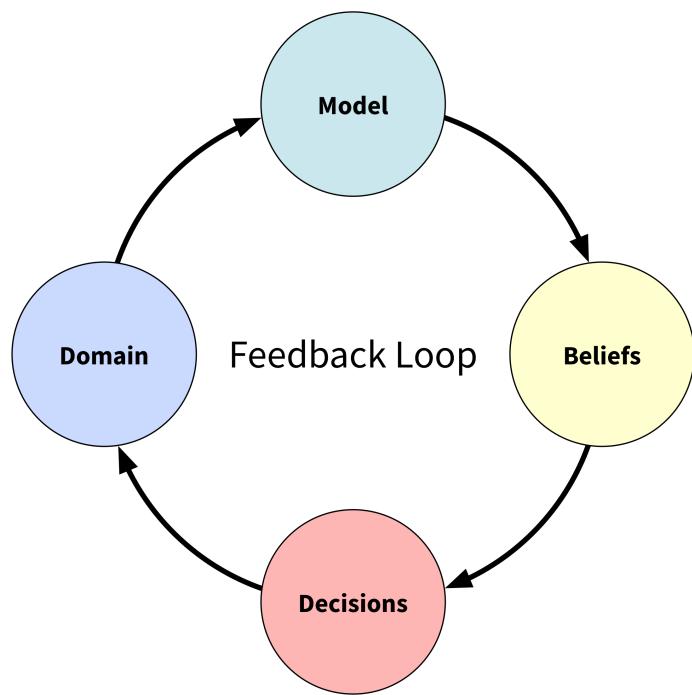
- Sample questions:
 - "How many prior convictions have you had?"
 - "When was the first time you were ever involved with the police?"
 - "What part did others play in the offense?"
 - "What part did drugs and alcohol play?"

Recidivism prediction models

How the LSI-R is used:

- The LSI-R is used in many states to feed a predictive model that estimates the likelihood the person would reoffend if released.
- In some states it is used to better target anti-recidivism programs.
- In some states (incl. Idaho and Colorado) the results of the model are used to determine the severity of the person's criminal sentence.







That power can be destructive

The four horsemen of big data

Four concerns that come up over and over again:

- Privacy.
- Negative externalities.
- Discrimination.
- Feedback loops.

Privacy

What can you tell about someone from what the "like" on Facebook?

Negative externalities

US News & Reports rankings →

- led to unforeseen uses in decision-making
 - led to dramatic increases in college costs
 - led to defunding of poorly ranked but nonetheless high-quality colleges

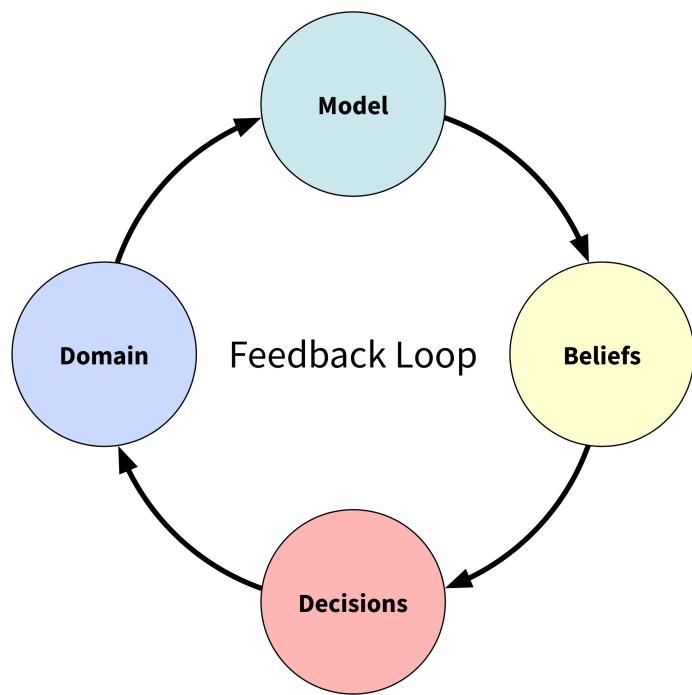
Discrimination

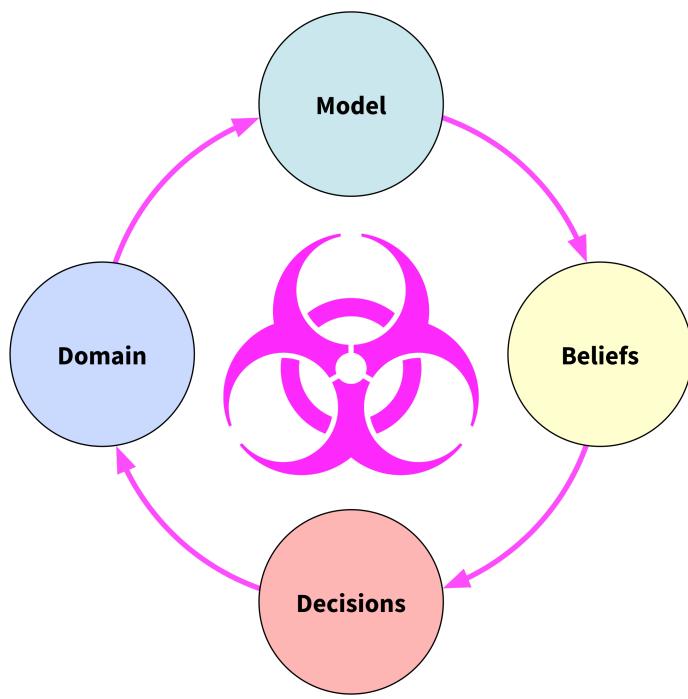
Due to sampling bias caused by social bias.

- Early contact with police as a proxy variable in recidivism prediction.

Due to the effect of social bias and social structure on the phenomenon itself.

- SAT scores.





Other examples?



That destruction can be invisible

College rankings, again

What kinds of things would you have had to know in order to anticipate the negative effects of US News' ranking model?

- What questions would you need to have asked?
- How could you have figured out to ask those questions?

Model opacity

Introduced by others' limited understanding of or access to the model:

- The people using a model to make decisions sometimes lack even a basic understanding of how the models work.
- The people whose lives are affected by those decisions frequently are completely unaware that the relevant models are even being used.
- Even if they are aware, it is often difficult or impossible for them to get answers about how the models work.



Death by a thousand cuts

Seemingly insignificant decisions can snowball into large-scale effects.

The Data Science Process

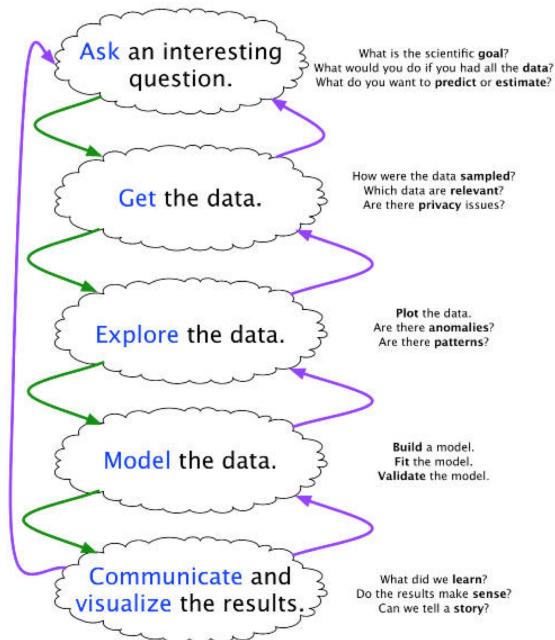
How does **ethics** fit in to the **data science process?**

Ethics and the data science process

How ethics fits into the data science process:

- Data scientists make countless **choices** while conducting their research.
- Many of these are **only partly constrained**, if at all, by scientific considerations.
- **Asking the right questions** at each stage and spending time trying to figure out the answers are essential to avoiding bad outcomes.

The Data Science Process



Ask an interesting question

What is the scientific goal? ➔

- Should that be the scientific goal?
- Should we be conducting this research in the first place?
- What are the motivations of my client in funding the research and how might they use our results in ways that affect others?

Get the data

How were the data sampled ➔

- Are there sampling biases in the data that could lead to discrimination?
 - Boston's pothole detection app.

Model the data

How might researchers' modeling choices interfere with others' ability to understand it?

- Including too many input or output variables can undermine interpretability.
- With some modeling techniques, the relationship between the input variables and the output variables is extremely difficult to determine with precision.
- When proxies are used, it can be difficult to tell whether they are proxies for variables you would never knowingly include as inputs.

Communicate and visualize the data

Which stories can you tell? ➔

- Which of those stories should you tell?
- How should you tell those stories?

Ariely video

With Great Power...

How much **moral responsibility** do
data scientists have?

How much moral responsibility
do data scientists have for what
other people do based on their
research?

How much responsibility do data scientists have?

A short argument that they can have quite a lot:

- Data scientists typically function as a kind of **adviser** to others.
- As advisers, they are **partly responsible** for the harmful effects that advice has, insofar as they could have foreseen those effects had they taken reasonable precautions.

Take-home points

How can I avoid **morally bad outcomes** in my research?

Avoiding bad outcomes requires:

Knowledge

- Of what moral issues might be relevant to research in the target domain.
- Of the broader social context surrounding the target domain and the research project itself.
- Of plenty of case studies, good and bad.

Practice

- Using this knowledge to ask the right questions while conducting research, and figure out how to answer them.

Lots of help

- From people with a wide range of relevant knowledge and experience.

