Driving organizational change

Choose your own adventure...

Adam Richards, PhD
Head of Enterprise Data Science
Galvanize, Inc
adam.richards@galvanize.com

galvanize

- Computational Biologist
- Industry, Academia, National Labs
- Healthcare, Genomics
- Built AI Platforms
- Educator, Data Scientist

Currently leading a diverse team of data scientists and data engineers. We help build and enable the leading industry data talent.

I consider one of the most important duties of any scientist the teaching of science to students and to the general public. ---Isaac Asimov *The Tragedy of the Moon* (1973), p. 224

Agenda

1 On the importance of Storytelling

One of the most common mistakes in data professionals make

2 On change management and realizing value

There is a close relationship between realizing project value and organizational change

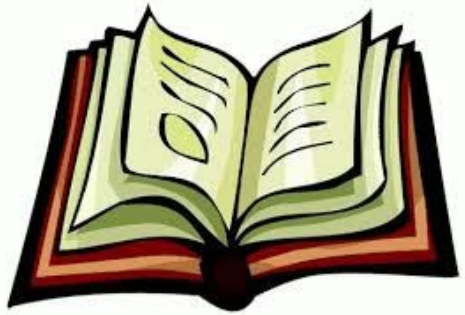
3 Let's play a game

You are part of a s fictional company that produces sells power to the grid using green energy and acts as an on-demand data center

Can you make the change management decisions that are best for you company?

On the importance of Storytelling

Elements of a good story



1. **Characters** → end-users, customers, stakeholders, data professionals
2. **Setting** → backstory, business scenario, data
3. **Plot** → central theme, business opportunity

How can a data product address a given business opportunity?

4. **Conflict** → What are the obstacles?
5. **Resolution** → What is the plan? What benefit or ROI can we expect?

The tools to help data professionals build their story

- **Build vs Buy** --- The best data products will be created by folks with BOTH domain expertise and modern flexible tools.
- **Open-source** --- The reality today is that the best tools are written using open-source technologies
- **Talent goes beyond technical skills**: pair-programming, Agile, microservices architecture, CI/CD, technical communication, team-player.
- **A consistent ecosystem** for data teams has a synergistic effect on ROI when combined with modern software engineering practices like version control and pair-programming


Storytelling

- **Very few people remember statistics**

> According to the McKinsey Global Institute data-driven organizations are 23 times more likely to acquire customers, 6 times as likely to retain customers, and 19 times as likely to be profitable.

- But it is likely that the ability to tell and consume stories has been selected for positively during human evolution.

> Here is a story about how a healthcare start-up got its beginnings…
>
> Three friends decided to go on a climbing trip together in *El Potrero Chico*. One was a data scientist, one a talented mobile developer and another a respected neurologist….

Storytelling and data visualization

- The business opportunity is the and the central hypotheses guide the stories.
- The days of Homer and Brothers Grimm (even slideshows) are long past.
- A good story **interacts with the audience** (interactive code).
- Data visualization is the single most effective prop to help us spin our tales

> The best stories are told by professionals with BOTH domain expertise and modern flexible tools.

Data Science by the numbers

This McKinsey report estimates that AI will have a 1.2-2 Trillion dollar impact on the Supply chain management and manufacturing sector. In order of importance….

1. Predictive Maintenance
2. Yield, Energy, Throughput
3. Procurement, Spend Analytics
4. Inventory and Parts Optimization
5. Logistics Network and Warehouse Optimization
6. Sales and Demand Forecasting

Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don't think AI will transform in the next several years. --Andrew Ng

On change management and realizing value

Measuring value from data projects

- Data Analyst
- Data Translator
- Data Scientist
- Data Engineer
- ML Engineer
- ML Researcher

- Jupyter Notebooks
- Flask/Shiny Apps
- Reports, Dashboards
- Docker
- Packages
- APIs

Business Opportunities

1. **New Opportunities**
2. **Improved or Optimized**
3. **Automated**

Measuring value from data projects

Data Maturity

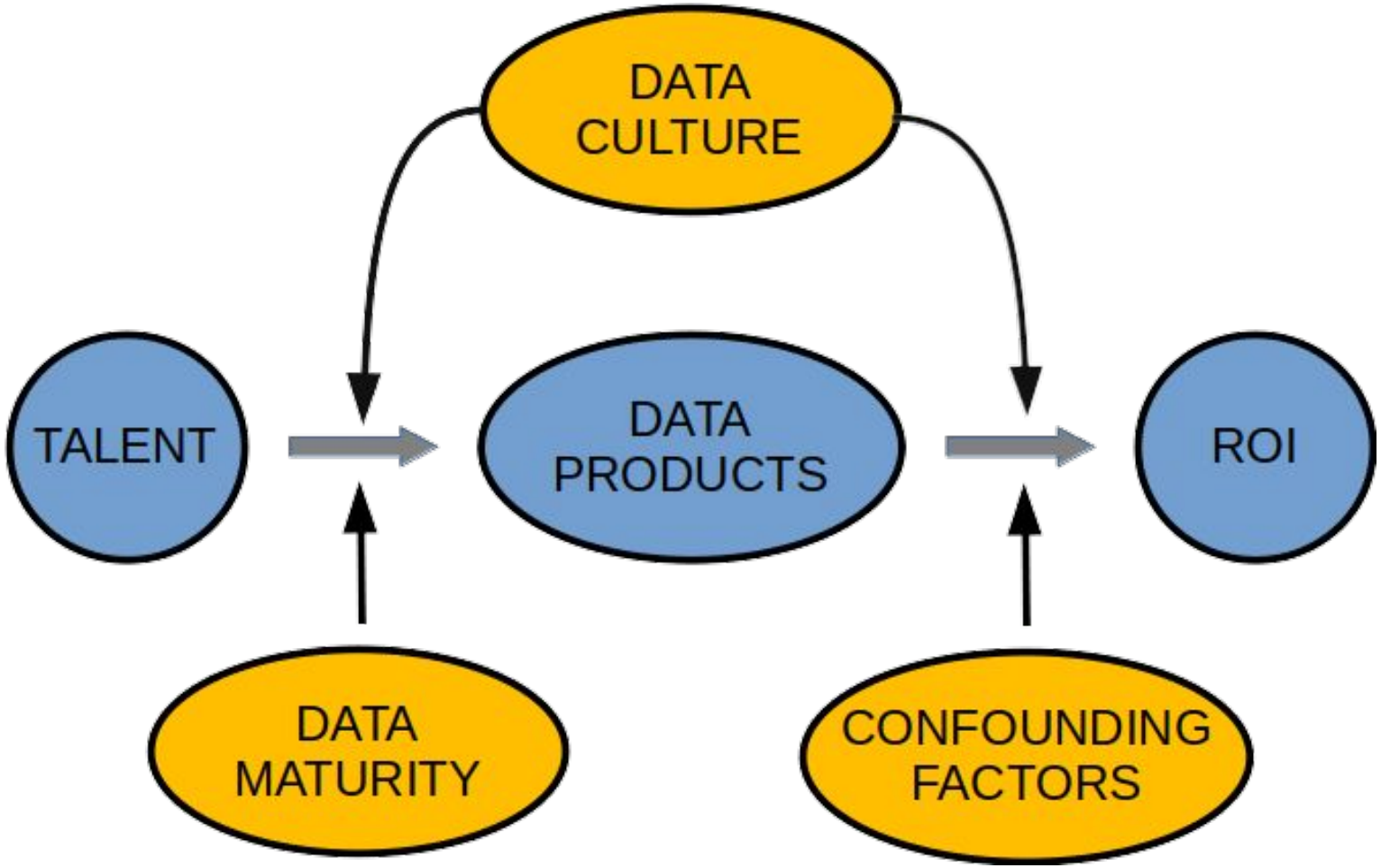
- Centralized Data Lake
- API’s, Data Streams

Data Culture

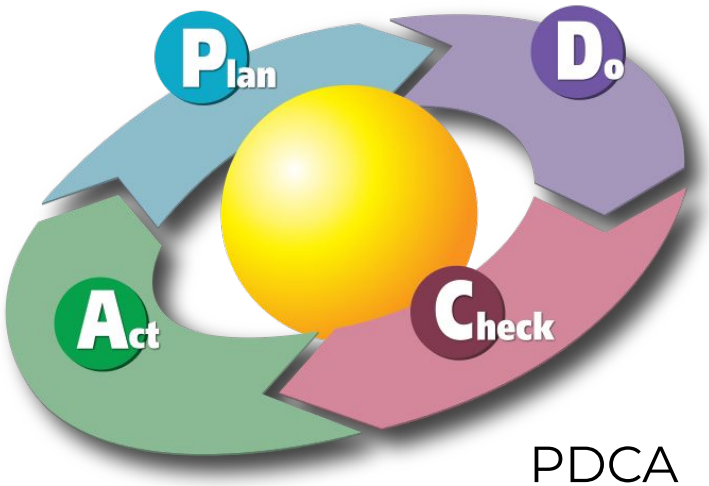
- Infrastructure, Team Org.
- Business Responsiveness

Confounding Factors

- Software releases, bugs
- Marketing, Seasonal trends



The Road to Data-Driven Will Require Change



PDCA

Think of Data as Capital

In economics, capital consists of assets that can enhance one's power to perform economically useful work.

- Executives and leadership play a critical role in advocating and orchestrating change
- Management plays a key role in facilitating and driving change
- Data must be thought of as a capitol

We do not advocate for a specific approach to change management, but we do suggest that the approach be structured. Dr. John P. Kotter of Harvard has proposed an 8-step process and the statistician W. Edwards Deming proposed a iterative 4-step process called Plan-Do-Check-Act Cycle (PDCA).

Change Management: a collective term for all approaches to prepare, support, and help individuals, teams, and organizations in making organizational change (wiki)

Change management and storytelling

A key component to all forms of change management is the ability to win the buy-in of you organization's employees on the change.

- Do not focus on statistics: performance metrics, financials , efficiency numbers
- Use the power of storytelling to help convey these messages
- Use data, models and visualizations to help convey your story

Leadership and management can promote buy-in using stories In the same way data scientists do this to ‘sell’ their solutions to business problems.

Why is change management so important?

- A fortune 20 company hires another company to upskill and modernize its data professionals. More than a year after the training the same company is hired to investigate the impact of the training.
  - With conservative estimates about productionized capstone project either through savings or new products revenue was estimated in the 10s of millions in the first couple of years.
  - **9/10 projects were abandoned**
- A fortune 500 bank tasks a data team with optimizing call center staffing, procedure and infrastructure.
  - Through simulation and modeling a solid solution was developed
  - The project lead presented to leadership and got buy in
  - Savings estimated to be ~10 million in the first year
  - **Project was never implemented**

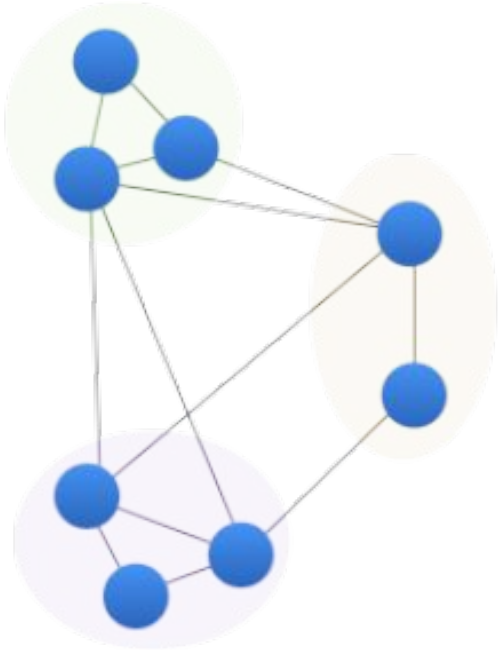
What happened?

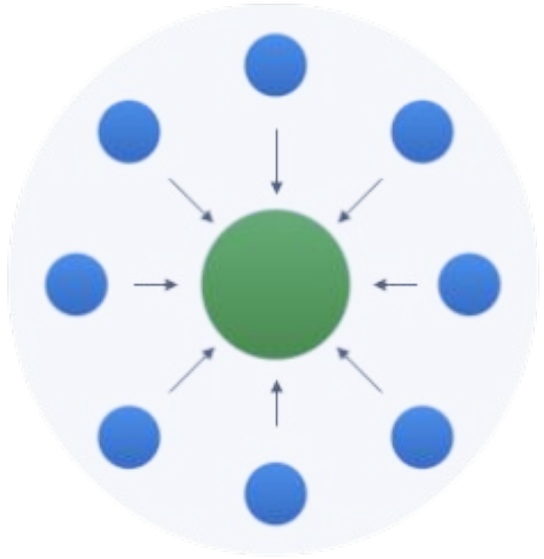
POLL 1

All of the following are reasons data science projects can fail. Which of the following was the most influential in the previous two examples?

(1) Failure to move the data product from prototype to production
(2) Poor communication with business stakeholders during development
(3) Lack of available infrastructure (cloud computing, storage etc.)
(4) Poor within team communication (shared tools, languages etc.)
(5) Limited to no ability to operationalize projects that span business units.

Organizing data teams

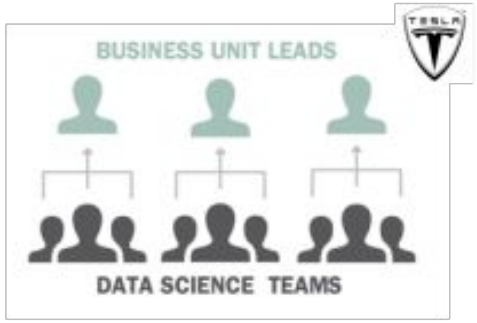




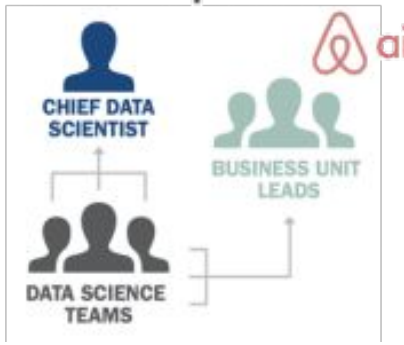
Centralized

at&t

Embedded



Hub & Spoke



airbnb

Let’s play a game

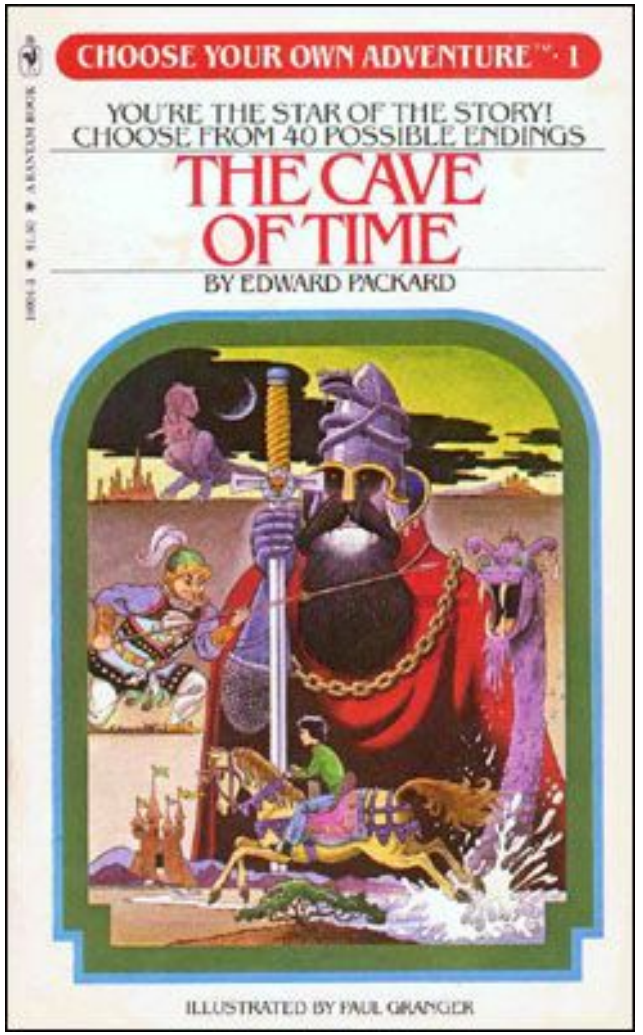
Setting the stage for our game

Choose your own adventure (*Elige tu propia aventura*) books are more generally known as gamebooks (Librojuegos).

Setting

- We are a small company that uses solar farms and other sources of renewable energy to sell back to the grid
- We also provide data center services on-demand

- We use AI to predict demand (**energy and service**)
- We leverage **supply chain optimization** techniques to make smart decisions about how to maximize revenue



credit: wiki

POLL 2

We have to hire a team of data professionals which arrangement makes the most sense for our company? We have a limited budget so here are the options.

(1) 2 data engineers, 1 data analyst, 1 data scientist
(2) 4 data scientists
(3) 1 data engineer, 1 data analyst, 2 data scientists
(4) 1 data engineer, 3 data analyst, 1 data scientist
(5) 4 data engineers

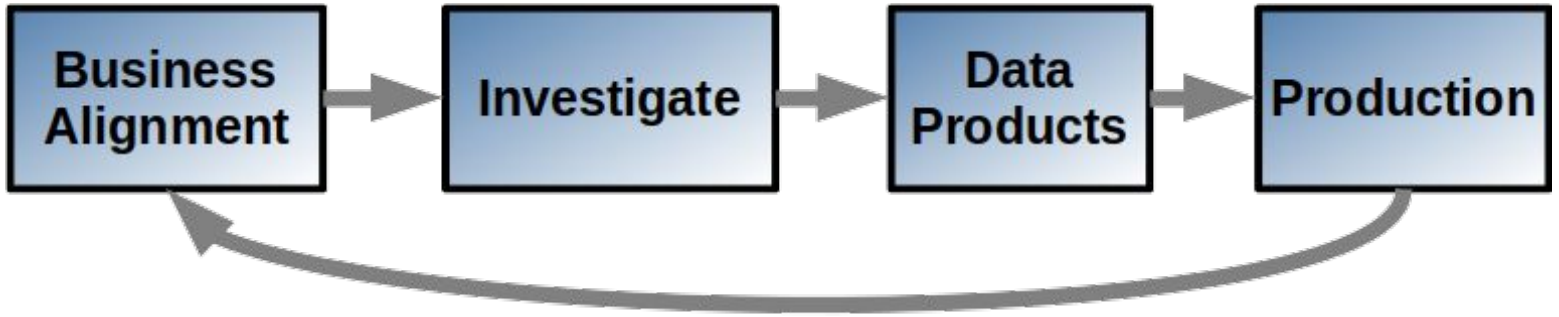
If you chose (2) or (5) then your start-up adventure ends early.

You want to have a diverse team with the expertise of individuals spanning the entire data science project lifecycle. A rationale for any of the other choices can be constructed depending on the priorities of the business.

For example, if there is a long term vision and scale is the priority then you may want to invest up front in data engineers. On the other hand if there is a need to prove to investors that the business is moving in the right direction then you will want more data scientists and data analysts.

Data-Driven Organizations - Process

They use a **formalized process** and have a strong data culture



Data Translator

Data Scientist

Data Engineer

Process Models:

- CRISP-DM
- OSEMN
- Design Thinking

One of the important subplots

The plot of our story here is to run a successful business, but there are many subplots to consider.

A subplot
• Part of our business is running a data center • Every month you see the P&L (profit and loss statements) and it is clear that the data centers incurs significant costs when it comes to the infrastructure related to cooling. • It is time to leverage the data team to help reduce cost

POLL 3

Cooling fans and other equipment fail. Contractors are used to perform maintenance and staff operators are onsite. Given what you know about the cooling issue select the data product that you think will have the largest ROI.

(1) Contractor staffing optimization model based on past failure data
(2) Predictive maintenance model for relevant equipment
(3) Use past failure data to better estimate cost that is then used to smartly allocate electricity between data center and the grid
(4) Incorporate past failure data into a larger AI system that makes smart recommendations about what operators should do and when

All of the answers are viable projects. The best answers are (1) and (2).

Predictive maintenance and staffing optimization are the best answers in this situation because they are the easiest data products to build and deploy of the 4 possibilities.

Of course we did not mention whether or not the right data were available for these projects, which is an often overlooked critical factor.

Google uses AI in their data centers to help operators make decisions, but this is an example of a more involved project.

Predictive Maintenance Demo

POLL 4

Now you have a working predictive maintenance model for your cooling fans. Which of the following will be the least helpful for the change management associated with this project?

(1) Design a plan to track the performance of the model and investigate the effect it has on the business
(2) Convene the data team to re-prioritize everyone's efforts on the next project that will have the best ROI
(3) Continue to iterate on the model including stakeholders from all business units
(4) Automate the performance monitoring, data ingestion, model training and quality assurance components of the project

All of the answers except (2) are reasonable. Including stakeholders at this point in the process here is the best answer.

Continuing to iterate on the project implies that you will be (4) Automating the performance monitoring, data ingestion, model training and quality assurance components of the project and (1) Designing a plan to track the performance of the model while investigating the effect it has on the business.

Recap

1 On the importance of Storytelling

One of the most common mistakes in data professionals make

2 On change management and realizing value

There is a close relationship between realizing project value and organizational change

3 Let’s play a game

You are part of a s fictional company that produces sells power to the grid using green energy and acts as an on-demand data center

Can you make the change management decisions that are best for you company?

Questions?

Appendix

The essential terminology (in a picture)

- Deep-learning is mostly the use of neural networks
- Machine learning is mostly comprised of supervised-learning and unsupervised-learning
- AI is used to describe any of these systems