

Perception is Key

February, 2013

Mark Madsen

www.ThirdNature.net

@markmadsen



The problem with bad framing



Leads to bad assumptions about use, inappropriate features, poor understanding of substitutability and the impacts it will have.

Bad framing leads to misunderstanding

“Big data is unprecedented.”

- Anyone involved with big data in even the most barely perceptible way



There's a difference
between having no past
and actively rejecting it.

It's nice, but it'll never replace playing outside in the fresh air and getting plenty of exercise.



TANSTAAFL

Technologies are not perfect replacements for one another. Often not better, only different.

When replacing the old with the new (or ignoring the new over the old) you always make tradeoffs, and usually you won't see them for a long time.

A history of databases in No-tation

1970: NoSQL = We have no SQL

1980: NoSQL = Know SQL

2000: NoSQL = No SQL!

2005: NoSQL = Not only SQL

2013: NoSQL = No, SQL!

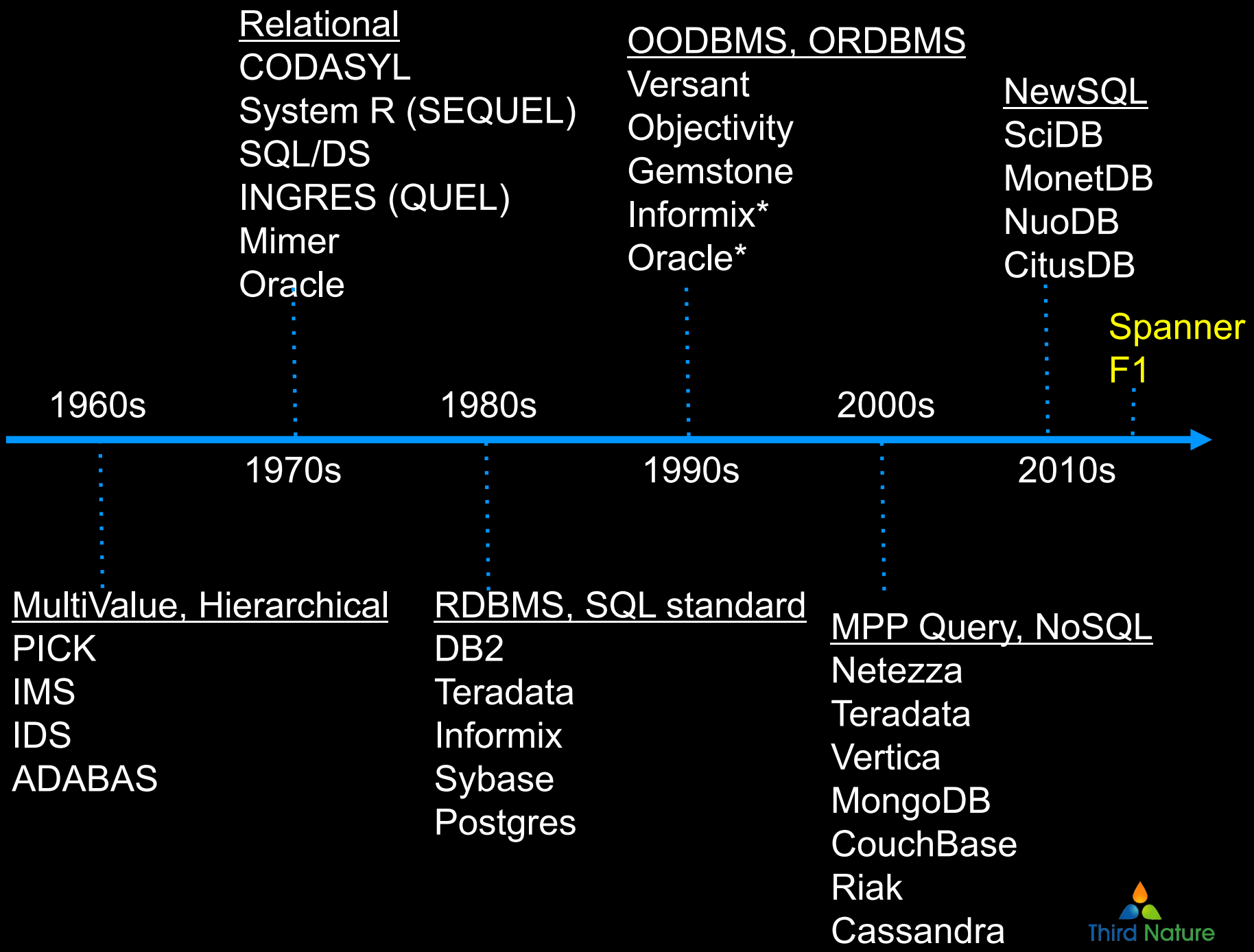
(R)DB(MS)

Maybe...

We need one that speaks pig latin.

These aren't the databases we're looking for.





Disruption vs Destabilization

Disruption is a bad framing for most software.

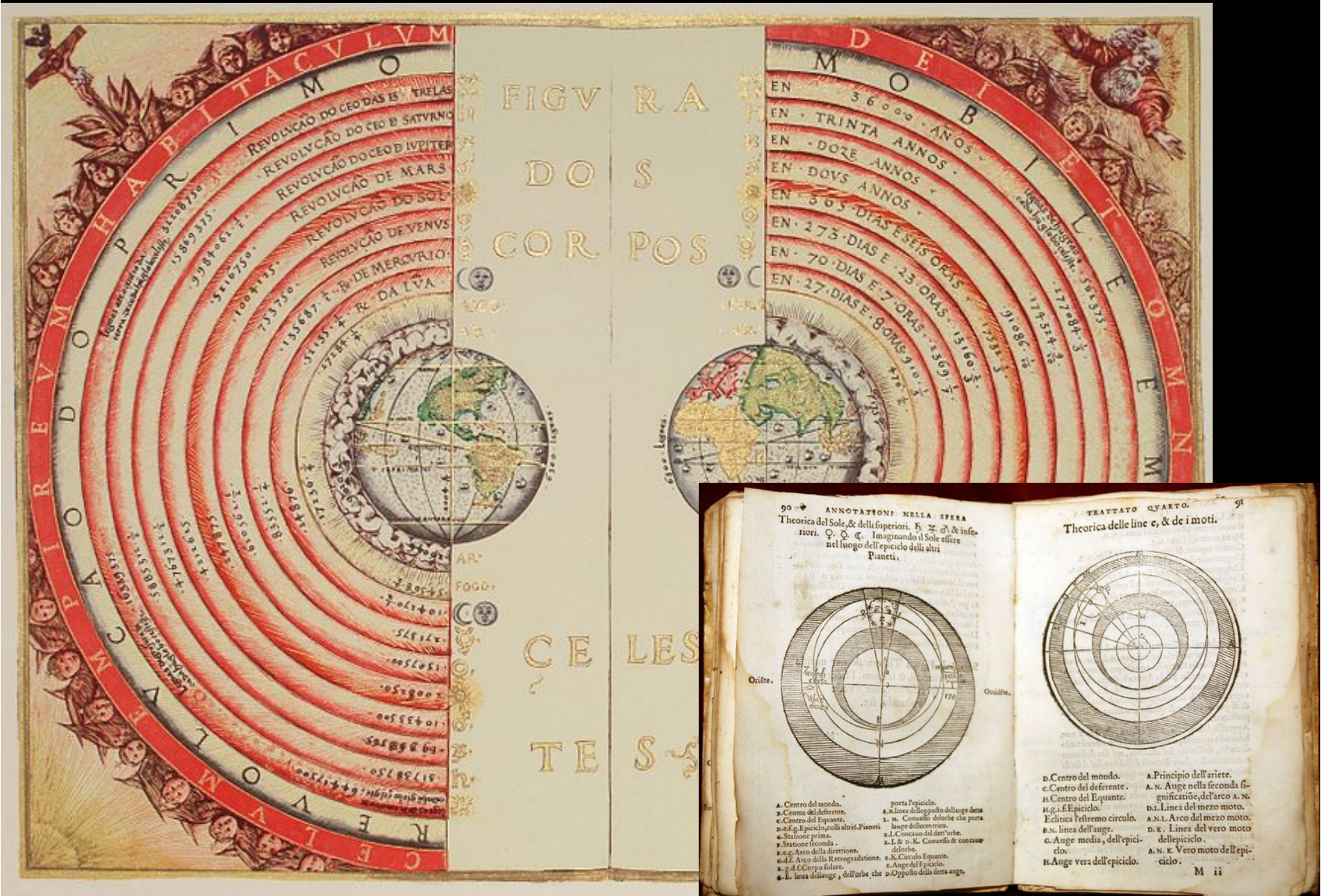
It exists in a software ecosystem with many dependencies.

Better to think of new technology as **destabilizing**.
This does not imply direct replacement.

Technology is not additive, it is transformative



The scene in the 1500s



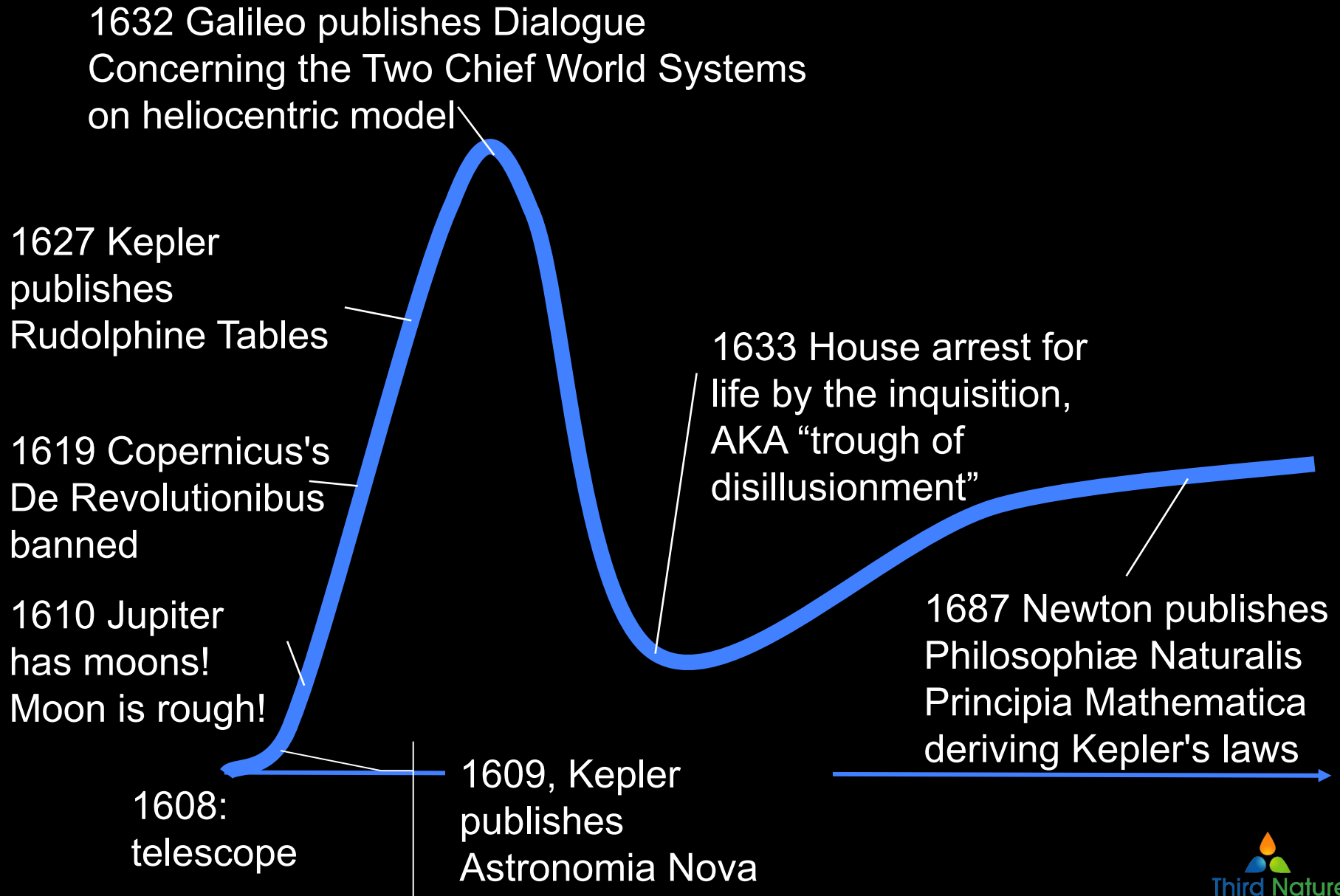
Pre-telescope, Copernicus works on discrepancies



Galileo sees some things in 1609-10



Discoveries follow a familiar curve



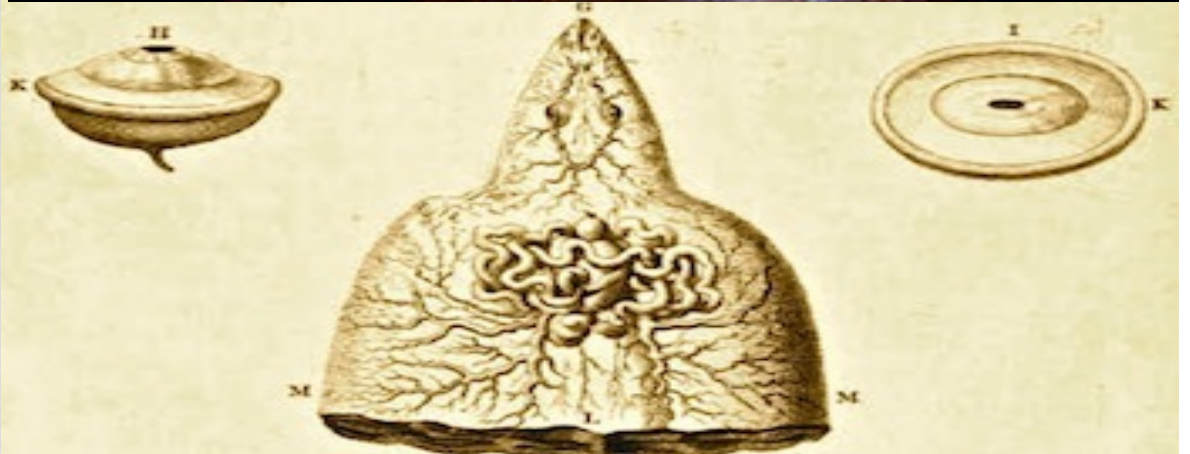
The scene by the late 1600s

- A. We are not at the center of the universe
- B. We're much smaller than we thought
- C. Universal law: fixed code, variable data, i.e. determinism, if you know the initial conditions you can predict anything

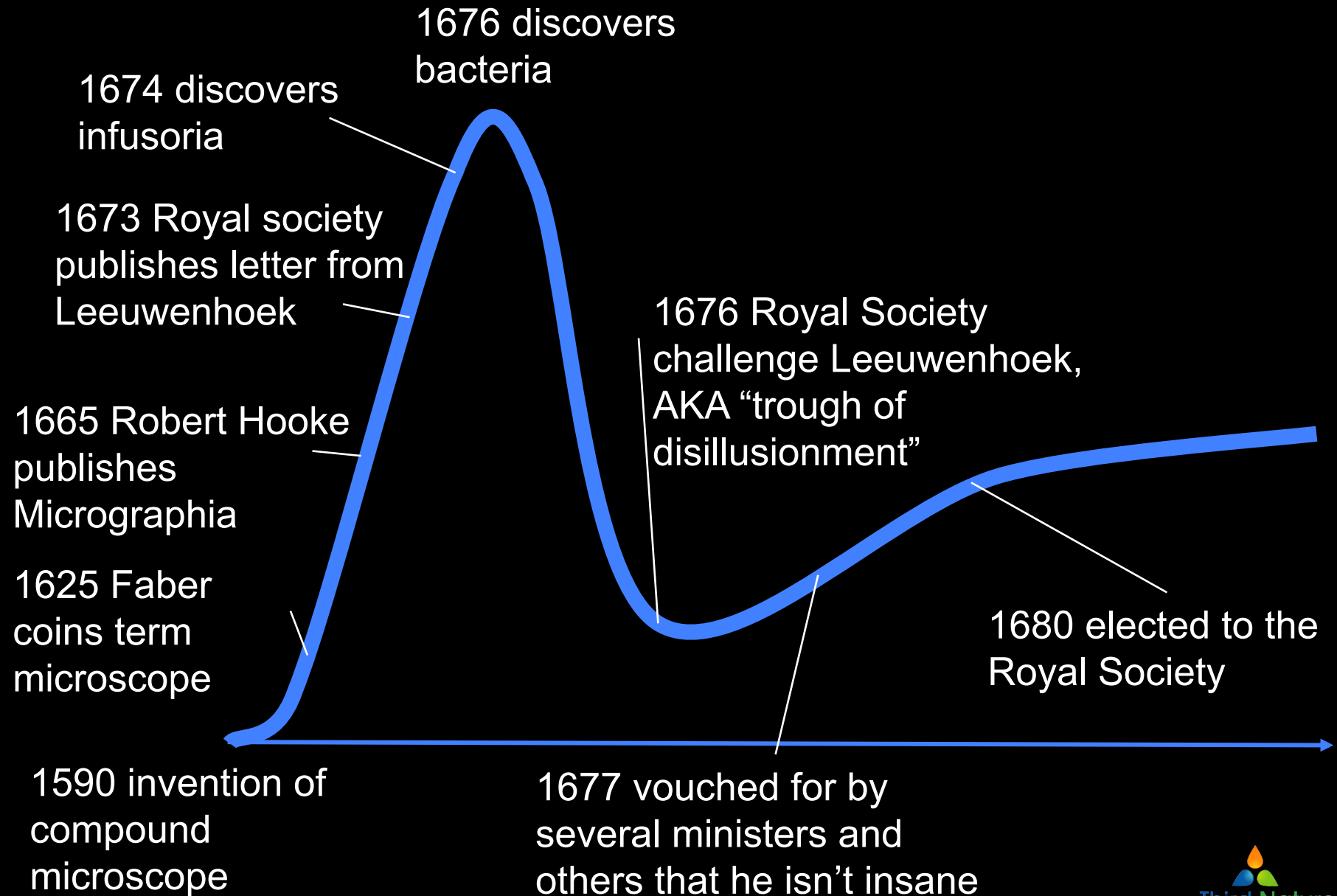
Leeuwenhoek builds a microscope



He discovers a lot of things nobody suspected



Discoveries follow a familiar curve



Two scopes that changed perspective

- A. Telescopes expanding our observable world, changed our concepts of scale, our place
- B. Microscopes expanded the observable world, changed our concepts of scale, our relationship to our surroundings
- C. These tools allow us to look deeper, challenge assumptions, generate information

Like the telescope and microscope, computing technology has become a tool of observation

Majority use of computing over time

1930s-1950s: Calculate

1960s-1980s: Automate

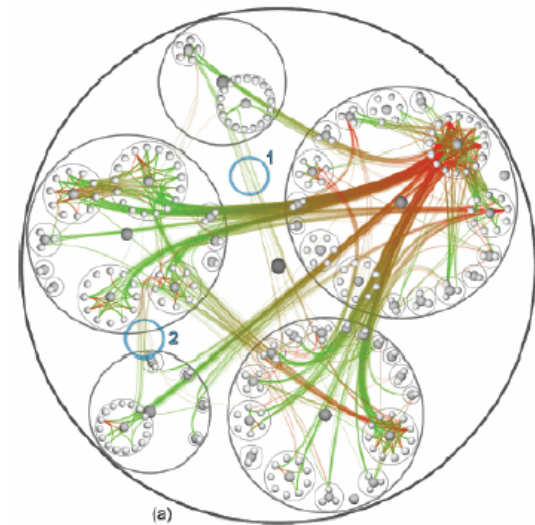
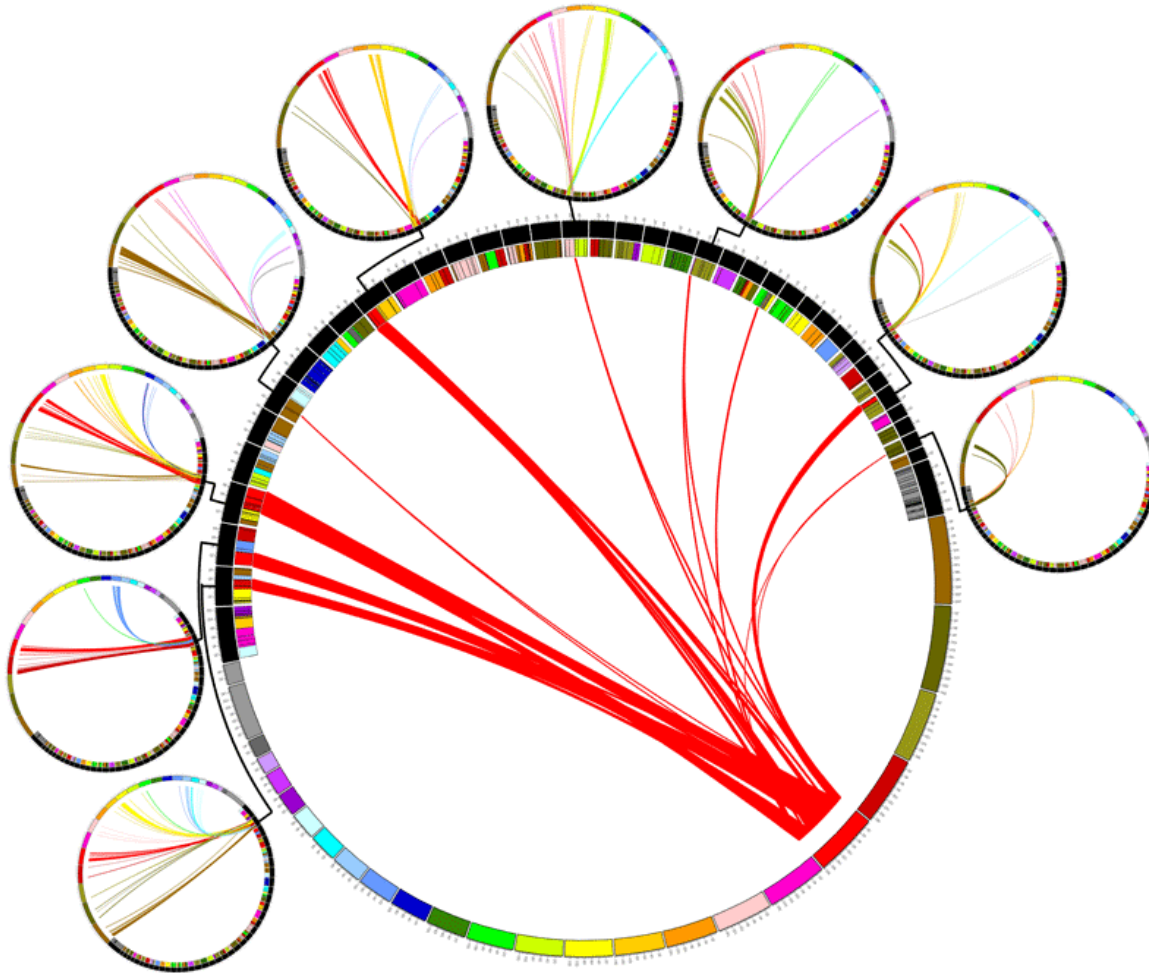
1990s-2010s: Informate

2010s+: Actuate

Rising organizational complexity

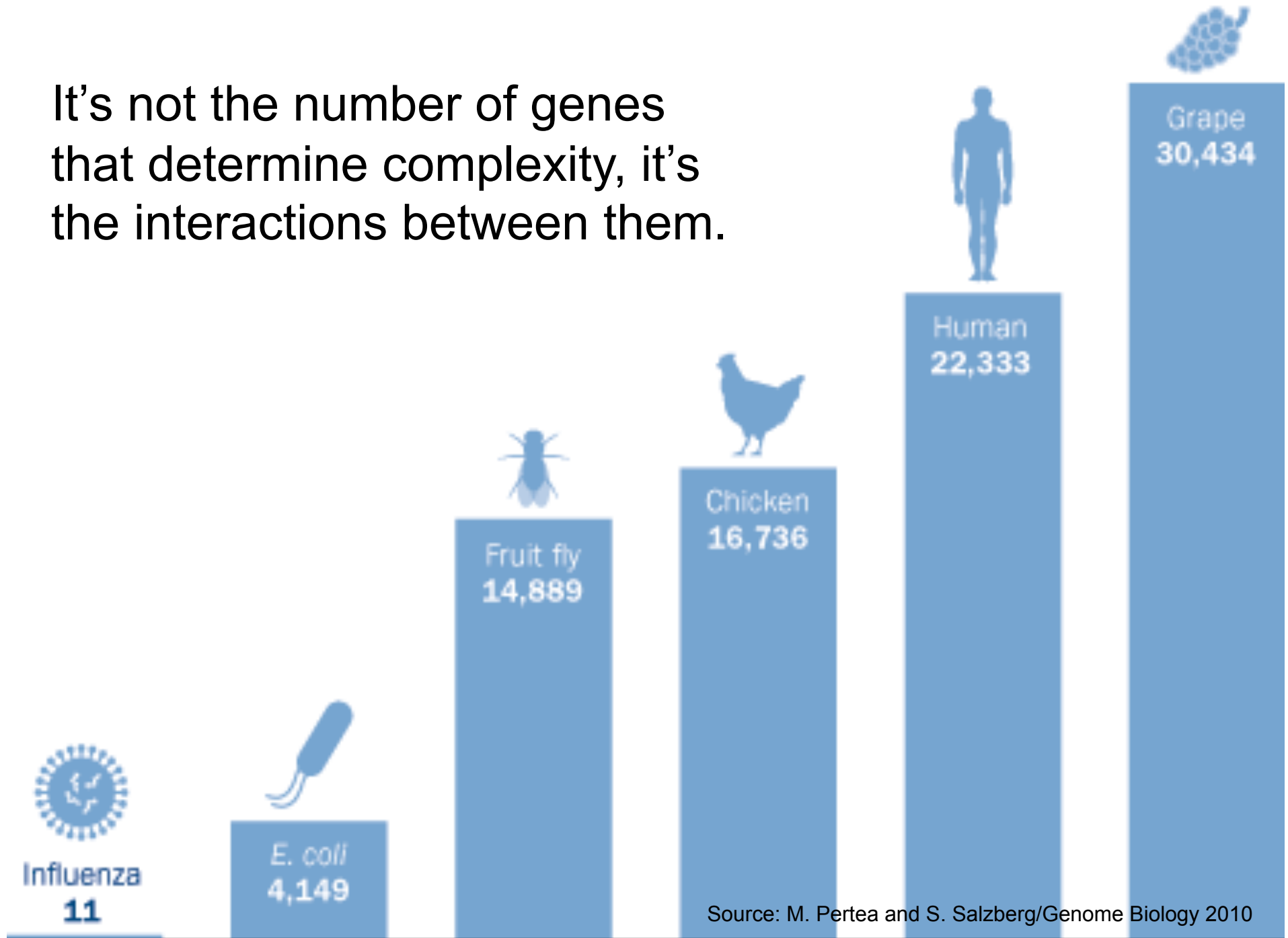
Interconnection and Dependency

Dynamic models are missing from most data systems today. These drive new workloads, generate different data, need new techniques.



Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data, Danny Holten

It's not the number of genes that determine complexity, it's the interactions between them.



Source: M. Pertea and S. Salzberg/Genome Biology 2010

**When you first give people access to information
that was unavailable...**



**OH GOD
I can see into forever**

After a while it becomes the new normal



As practices evolve based on new capabilities...



OMG WTF

A new level of complexity develops over top of the older, now better understood processes, leading to new data and analysis needs.

How do you manage the business in today's environment?

Our simplistic notions of decisions with stable models, ordered data and predictability are being replaced by concepts from decision support and complex adaptive systems (CAS).



Simple

Assumption: Order

Cause and effect is repeatable
& predictable

Known

Standard processes, clear
metrics, best practice

Sense, categorize, respond

Like a bike

How do you manage the business in today's environment?

Our simplistic notions of decisions with stable models, ordered data and predictability are being replaced by concepts from decision support and complex adaptive systems (CAS).



Simple

Complicated

Assumption: Order	Assumption: Unorder
Cause and effect is repeatable & predictable	Cause and effect is separated in time & space, repeatable, learnable
Known	Knowable
Standard processes, clear metrics, best practice	Analytical techniques to determine options, effects
Sense, categorize, respond	Sense, analyze, respond
Like a bike	Like a car

How do you manage the business in today's environment?

Our simplistic notions of decisions with stable models, ordered data and predictability are being replaced by concepts from decision support and complex adaptive systems (CAS).



Simple

Complicated

Complex

Assumption: Order	Assumption: Unorder	Assumption: Disorder
Cause and effect is repeatable & predictable	Cause and effect is separated in time & space, repeatable, learnable	Cause and effect is coherent in retrospect only, modelable but changing
Known	Knowable	Unpredictable
Standard processes, clear metrics, best practice	Analytical techniques to determine options, effects	Experiment to create possible options
Sense, categorize, respond	Sense, analyze, respond	Test, sense, respond
Like a bike	Like a car	Like economic systems

Situational context governs data use

Data technology support varies for these contexts

Batch to online
operational reporting

Data warehousing
and BI

Big data, analytics

Assumption: Order

Cause and effect is repeatable
& predictable

Known

Standard processes, clear
metrics, best practice

Sense, categorize, respond

Reporting, dashboards

We do this well

Assumption: Unorder

Cause and effect is separated
in time & space, repeatable,
learnable

Knowable

Analytical techniques to
determine options, effects

Sense, analyze, respond

Ad-hoc, OLAP, data discovery

We do this ok

Assumption: Disorder

Cause and effect is coherent
in retrospect only, modelable
but changing

Unpredictable

Experiment to create possible
options, test hypotheses

Test, sense, respond

Casual analysis, simulation

We are still learning

Over time, processes are (usually) better understood, so there is a movement of decisions from right to left, where support is more automated.

Simple

Complicated

Complex

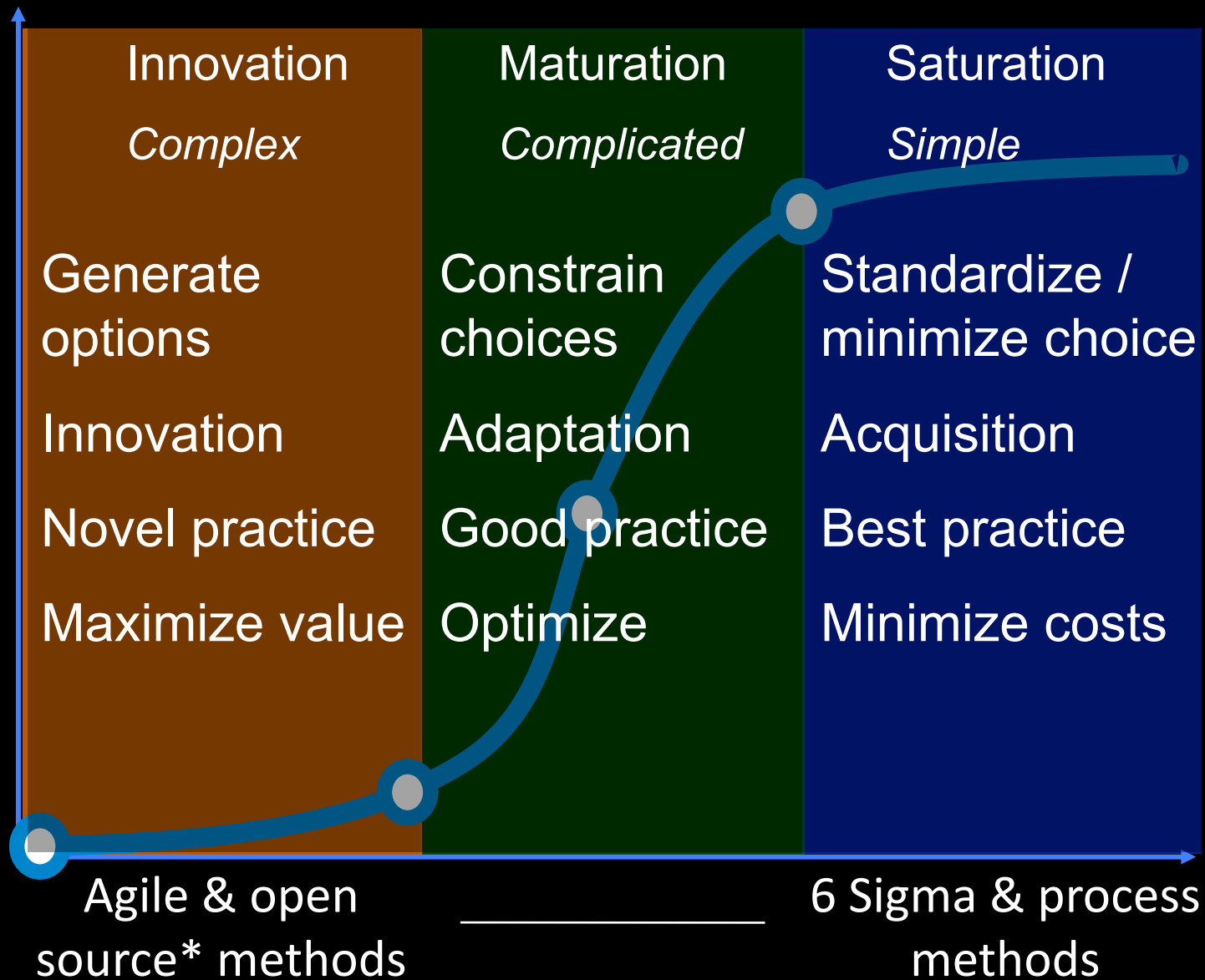
Assumption: Order	Assumption: Unorder	Assumption: Disorder
Cause and effect is repeatable & predictable	Cause and effect is separated in time & space, repeatable, learnable	Cause and effect is coherent in retrospect only, modelable but changing
Known	Knowable	Unpredictable
Standard processes, clear metrics, best practice	Analytical techniques to determine options, effects	Experiment to create possible options
Sense, categorize, respond	Sense, analyze, respond	Test, sense, respond
Like a bike	Like a car	Like economic systems

Complexity will continue to increase



Technology captures observations. These change our understanding. New understanding changes practices. Practices drive changes to technology, needing more data

As a technology moves from emerging to commodity the nature of acquiring, using and managing it changes



How we develop best practices: survival bias

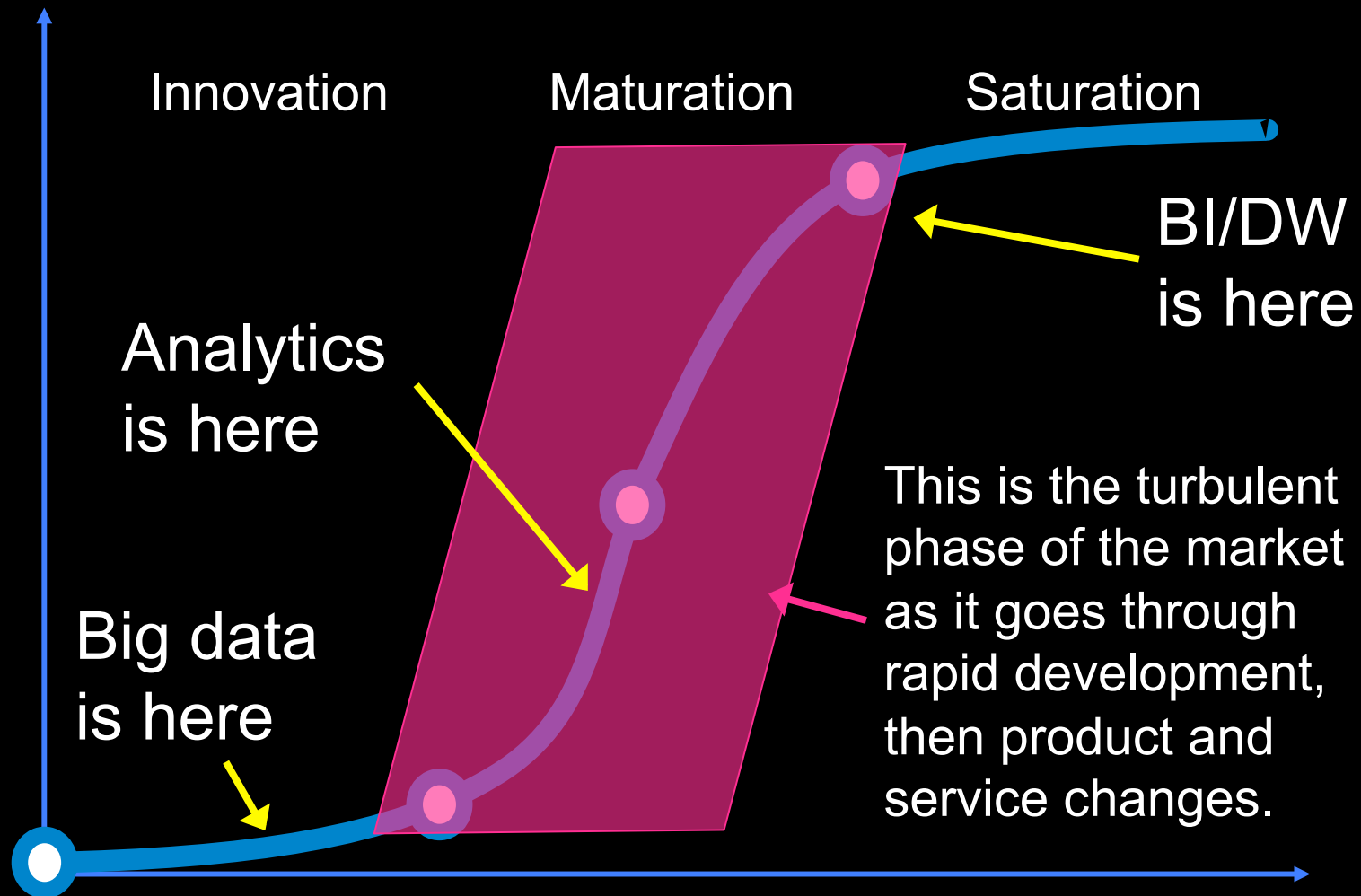




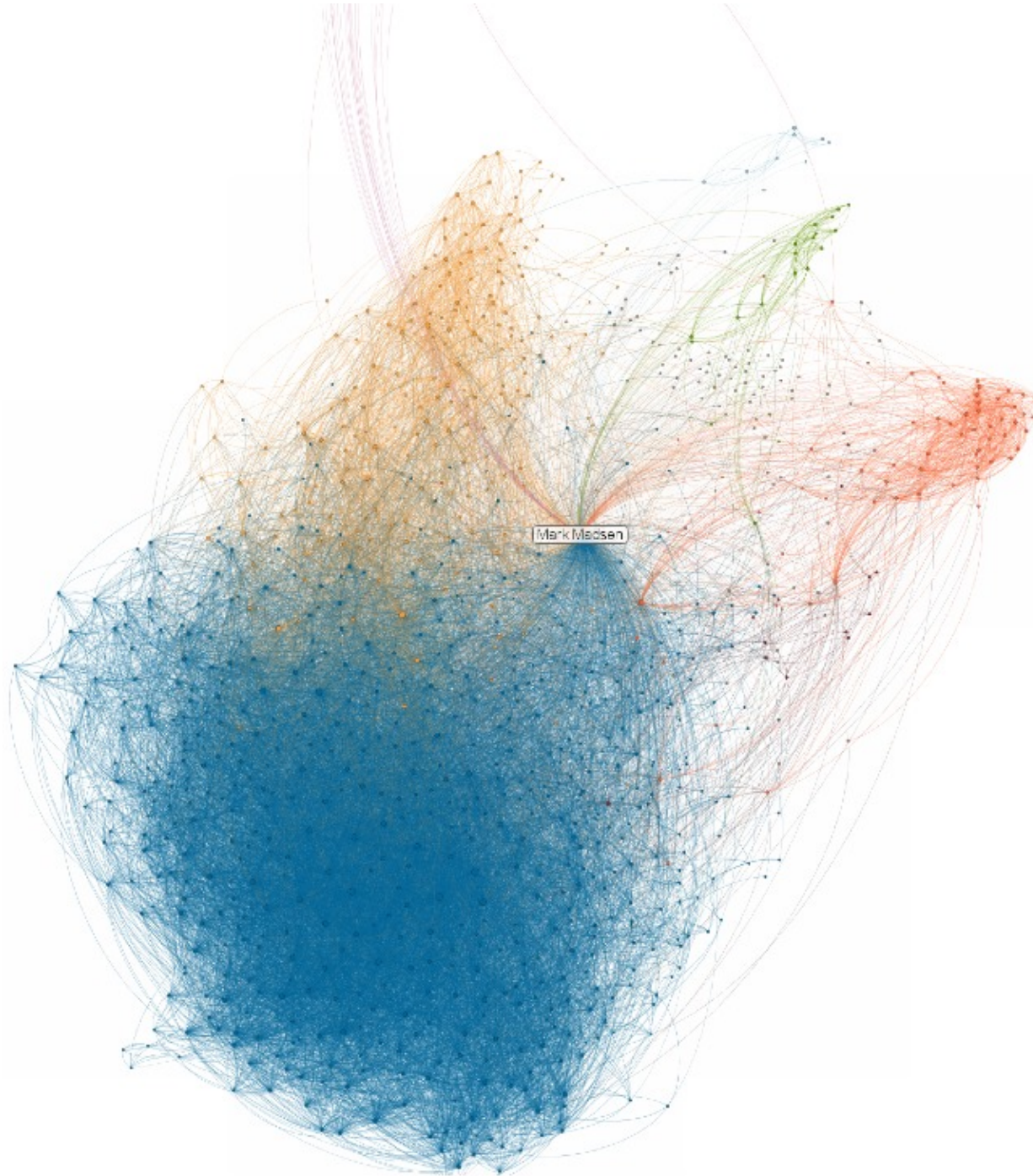
Beware best practices

The paths of trailblazers are often luck. Emulating what they do can be the equivalent of superstition.

Today: repeating the experience of the prior eras



The Internet combined with commodity computing is forcing a new business and IT structural evolution, already underway.



This shows the slow change of culture and practice.

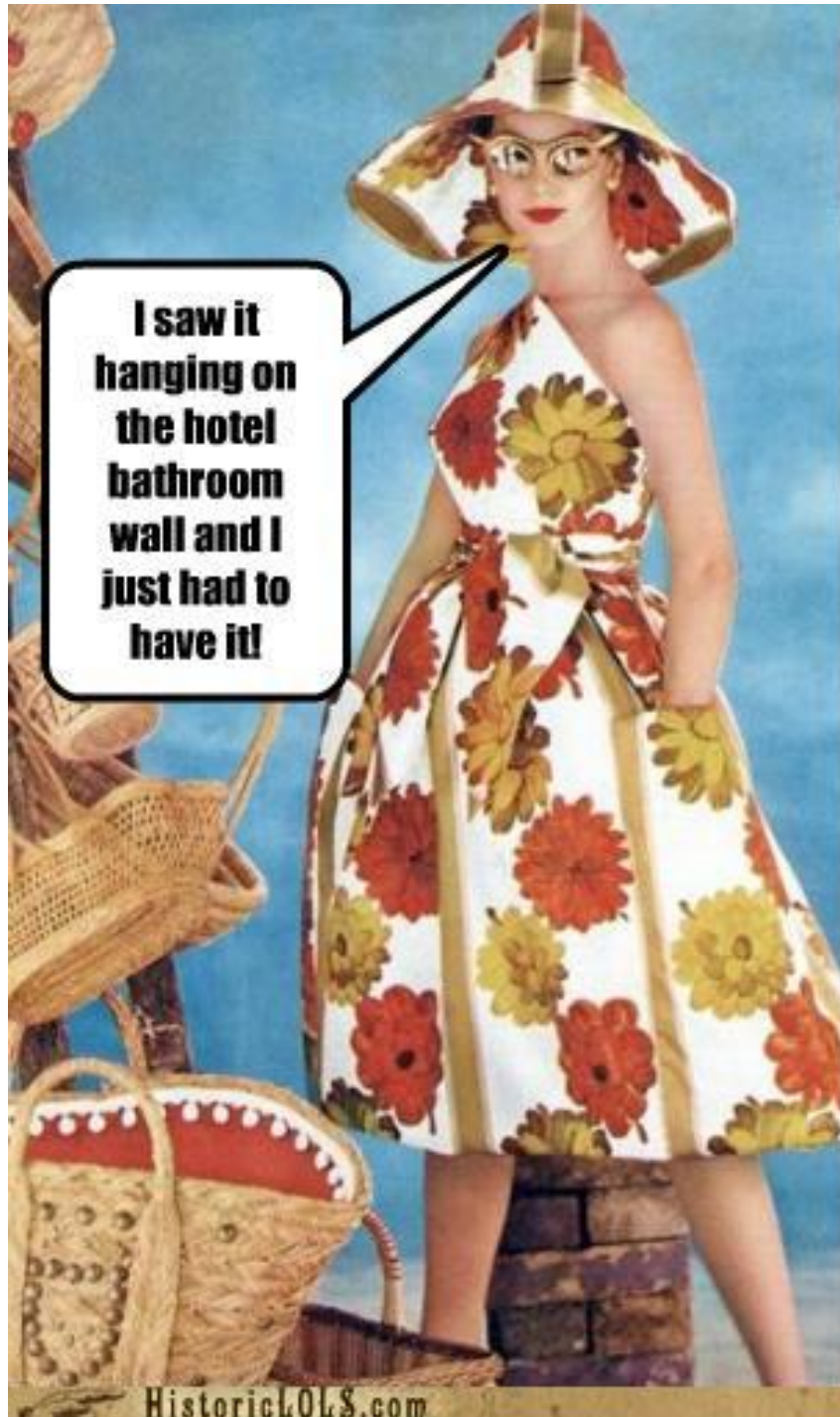
There are two opposing conclusions one could draw after analyzing this, depending entirely on perspective.

Reaction #1: Stick to what you know



Reaction #2: Use the new technology for everything





HistoricLOLS.com

Technology Adoption

Some people can't resist getting the next new thing because it's new.

Many IT organizations are like this, promoting a solution and hunting for the problem that matches it.

Better to ask "What is the problem for which this technology is the answer?"

The three way workload break

1. **Operational**: OLTP systems
2. **Analytic**: OLAP systems
3. **Scientific**: Computational systems

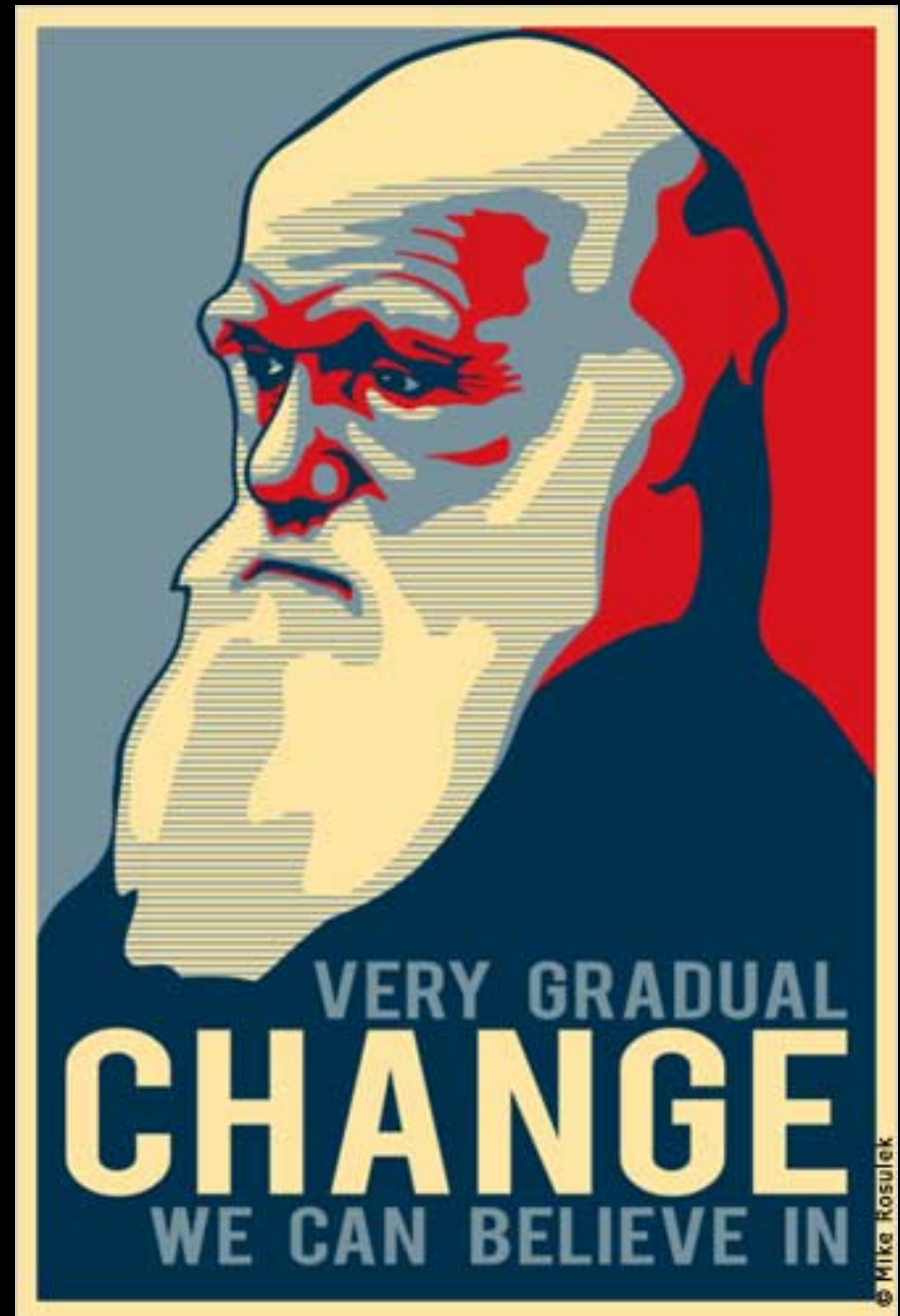
Unit of focus:

1. Transaction
2. Query
3. Computation

Different problems require different platforms

Change happens slower than you think. Data processing is ~70 years old and still evolving.

Technology adoption is gated by complexity: what came before, skills, dependencies, practices.



The big data revolution, more of an evolution



About the Presenter

Mark Madsen is president of Third Nature, a research and advisory firm focused on analytics, information strategy and data management. Mark is an award-winning author, architect and CTO whose work has been featured in numerous industry publications. Over the past ten years Mark received awards for his work from the American Productivity & Quality Center, TDWI, and the Smithsonian Institute. He is an international speaker, a contributor at Forbes Online and Information Management. For more information or to contact Mark, follow @markmadsen on Twitter or visit <http://ThirdNature.net>



About Third Nature



Third Nature is a research and consulting firm focused on new and emerging technology and practices in analytics, business intelligence, information strategy and data management. If your question is related to data, analytics, information strategy and technology infrastructure then you're at the right place.

Our goal is to help organizations solve problems using data. We offer education, consulting and research services to support business and IT organizations as well as technology vendors.

We fill the gap between what the industry analyst firms cover and what IT needs. We specialize in product and technology analysis, so we look at emerging technologies and markets, evaluating technology and how it is applied rather than vendor market positions.





SOME RIGHTS RESERVED

CC Image Attributions

Thanks to the people who supplied the creative commons licensed images used in this presentation:

prospector-1954.jpg - <http://www.flickr.com/photos/jthorstad/2203256491/>

Circos, Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data, Danny Holten