

# 5 Antipatterns in Scaling Enterprise AI

Sarah King

Director of Product, Molecula



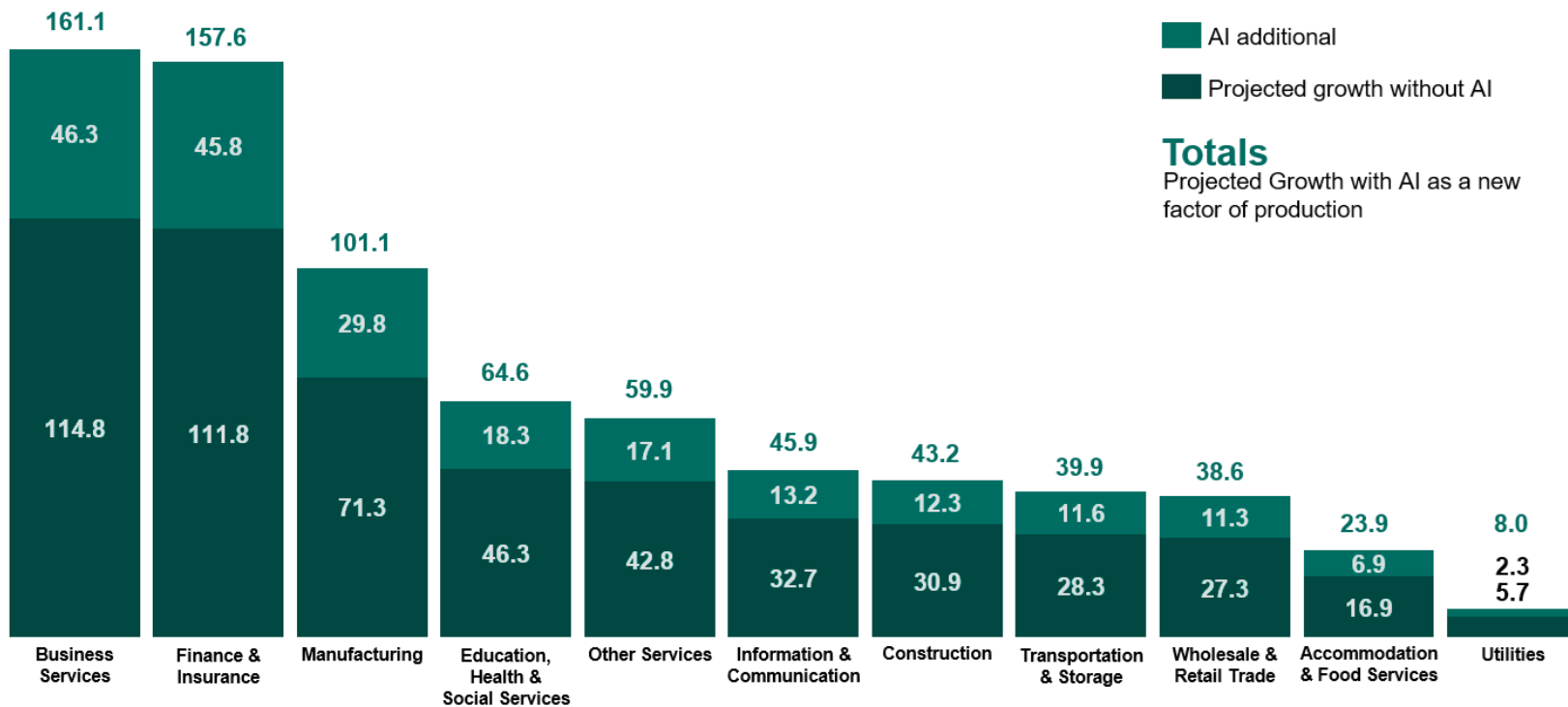
@ Molecula.

Background in Sales and Community Management.

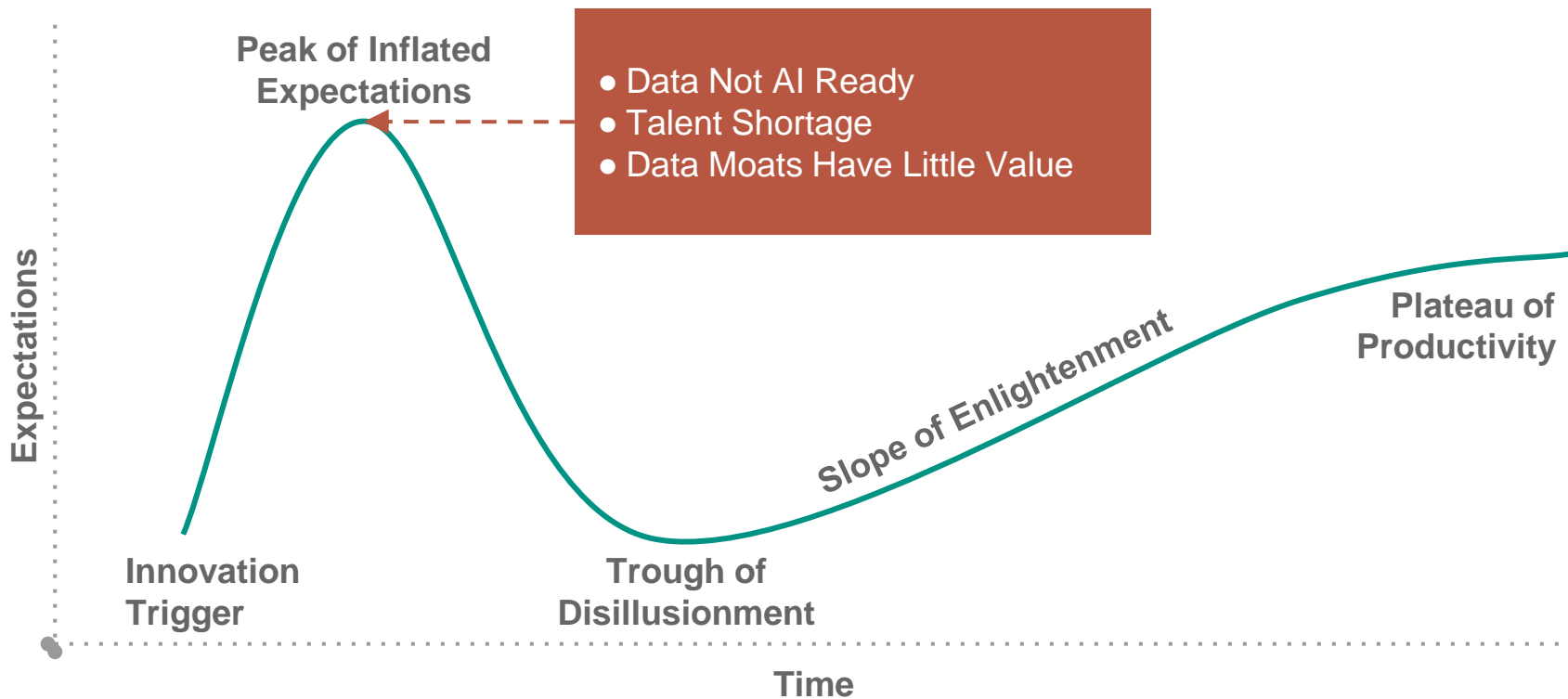
Cat Lady.

Krav Maga.

# At The Cusp Of The 5th Intelligence Era



# But, “AI” Is Falling Off The Hype Cliff



# What is an antipattern?

There must be at least two key elements present to formally distinguish an actual anti-pattern from a simple bad habit, bad practice, or bad idea:

1. A commonly used process, structure, or pattern of action that despite initially appearing to be an appropriate and effective response to a problem, has more bad consequences than good ones.
2. Another solution exists that is documented, repeatable, and proven to be effective.

*Coined by Andrew Koenig and inspired by [Design Patterns](#)*

# Many antipatterns in scaling Enterprise AI

1. Machine Learning projects happen in a vacuum.
2. “Laptop Data Science”
3. All Data Scientists are seen as the same.
4. Putting Expensive Resources on Model Management.
5. Not monitoring or communicating outcomes.

Other Areas for Discussion: Bias, Data Privacy, Feature Extraction and Reusability...

# 1. Machine Learning projects happening in a vacuum.



# 1. ML projects happening in a vacuum.

- People are protective of data in their business units
- Silos are very difficult to navigate
- No collective interest or buy-in



# 43%

Of Enterprises lack a clear strategy for AI.

*\*McKinsey, 'AI adoption advances, but foundational barriers remain,' Nov '18*

# 1. Collaborate cross-functionally and communicate the initiative company-wide.



CRO



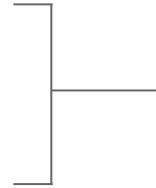
CTO



COO



CDO



Prioritize Business Objectives and empower a cross functional team to execute together.



Data Engineering



Data Science



Marketing



Sales



Tactical team collaborates on identifying Business Problems aligned with executive priority that can be solved by ML.

## 2. “Laptop Data Science”



## 2. “Laptop Data Science”

- Highly contested - big data sucks to work with.
- Many times, working off of small data is not representative of your business' production data.
- COPIES COPIES COPIES

# 85%

Of data in 2023 will be copies according to IDC.



## 2. Big data is easy to work with.

- Access to 100% of data.
- Accessing data CAN be close to real-time.
- Eliminate the back and forth that can add days and weeks to your project.
- Implement policies to protect against aimless copying.

# Many solutions for working with big data

Molecula™

 pilosa

 ARROW

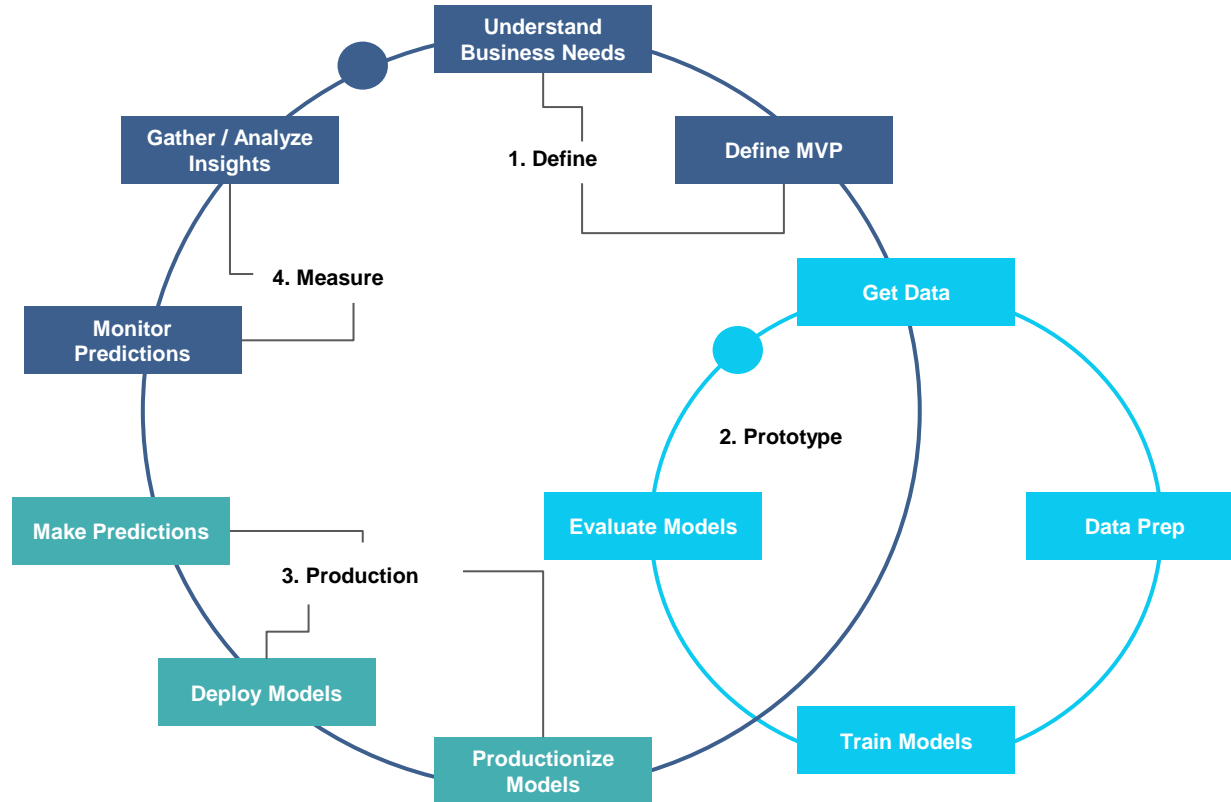
  
Parquet

3. All Data Scientists are seen as the same.





# 3. All Data Scientists are seen as the same.



3. Take an inventory of skill sets on your team and hire/structure accordingly.

- Data Products need maintenance that requires a broad range of skill sets.
- Writing Algorithms
- Data Engineering
- Machine Learning Engineers/Ops
- DevOps
- Data Governance

# 4. Putting the wrong people on production model management.



## 4. Make models manage models.

- Automate this process
- Data Skewness Criterion
- Model Skewness Criterion
- Less expensive Dev resources can then maintain

# 5. Not monitoring or communicating outcomes.



## 5. Not monitoring or communicating outcomes.

It's nice to look at shiny things!



# 5. Monitor WITH and WITHOUT.

*“Nearly every client looks at the deliverable vs. the outcome. It’s really easy to forget how critical it is to monitor a process with machine learning vs. without it.”*

8am outside my classroom waiting to meet my new students. I'm so excited. 3pm I met them!



# Thank you!

Sarah King  
Director of Product, Molecula  
sarah@molecula.com  
@sarahking\_atx  
@molecula