Launching Data initiatives is easy.

Bringing value at scale from Data isn't.

10 pitfalls

holding you back bringing value from data at scale

DEVOXX

hymaia

Hi, there !

Yoann Benoit

Head of Data &
Co-Founder @Hymaïa

@YoannBENOIT

yoann-benoit

Questions to reflect on

Do you have all the skills to create end-to-end Data Products ?

Do you see Data as a way to confirm your “business intuitions” ?

Do you often have people asking you “*is there something wrong with the Data*” ?

Trap #1

Narrowing Down Data to DS + DE

Narrowing Down Data to DS + DE

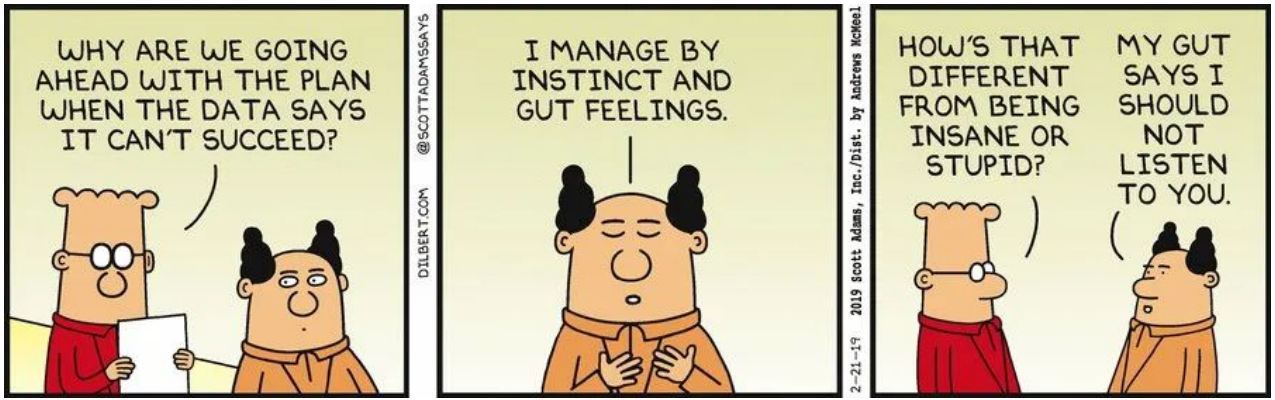
Pitfall #1

Not having a diverse enough Data Team

Narrowing Down Data to DS + DE

Pitfall #2

Being trapped in one's own biases



Narrowing Down Data to DS + DE

Pitfall #1

Not having a diverse enough Data Team

Pitfall #2

Being trapped in one's own biases

Narrowing Down Data to DS + DE

Pitfall #3

Believing data is trustworthy by default

Narrowing Down Data to DS + DE

Pitfall #1

Not having a diverse enough Data Team

Build a diverse Data Team, including business and product skills

Pitfall #2

Being trapped in one's own biases

Don't create a team where everybody comes from the same background

Pitfall #3

Believing data is trustworthy by default

Invest in Data Quality and Governance to generate trust

Takeaway

Data is plural.

Go beyond “profile” and think “skills”.

Questions to reflect on

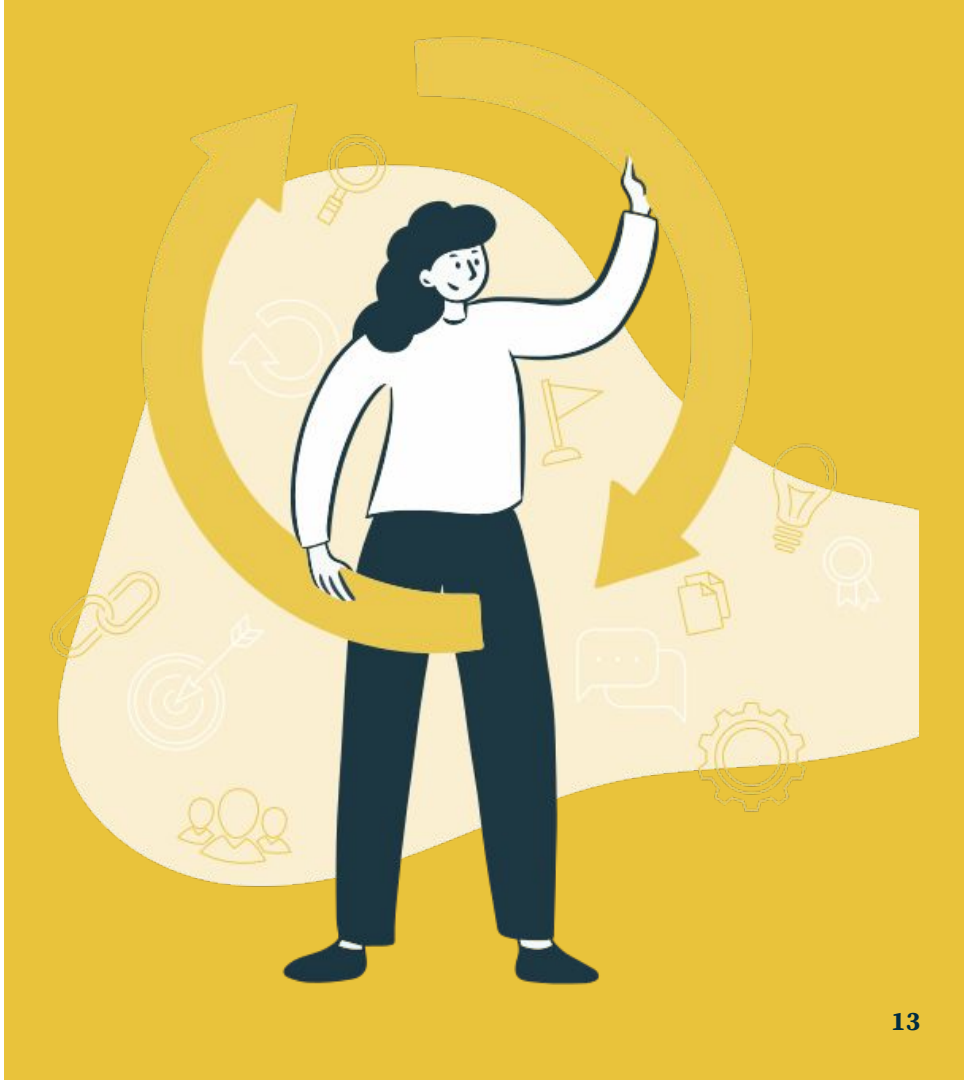
Do you manage Data Products like you manage Software ?

Is your centralized Data Team becoming a bottleneck ?

How painful is the deployment and management of your ML models ?

Trap #2

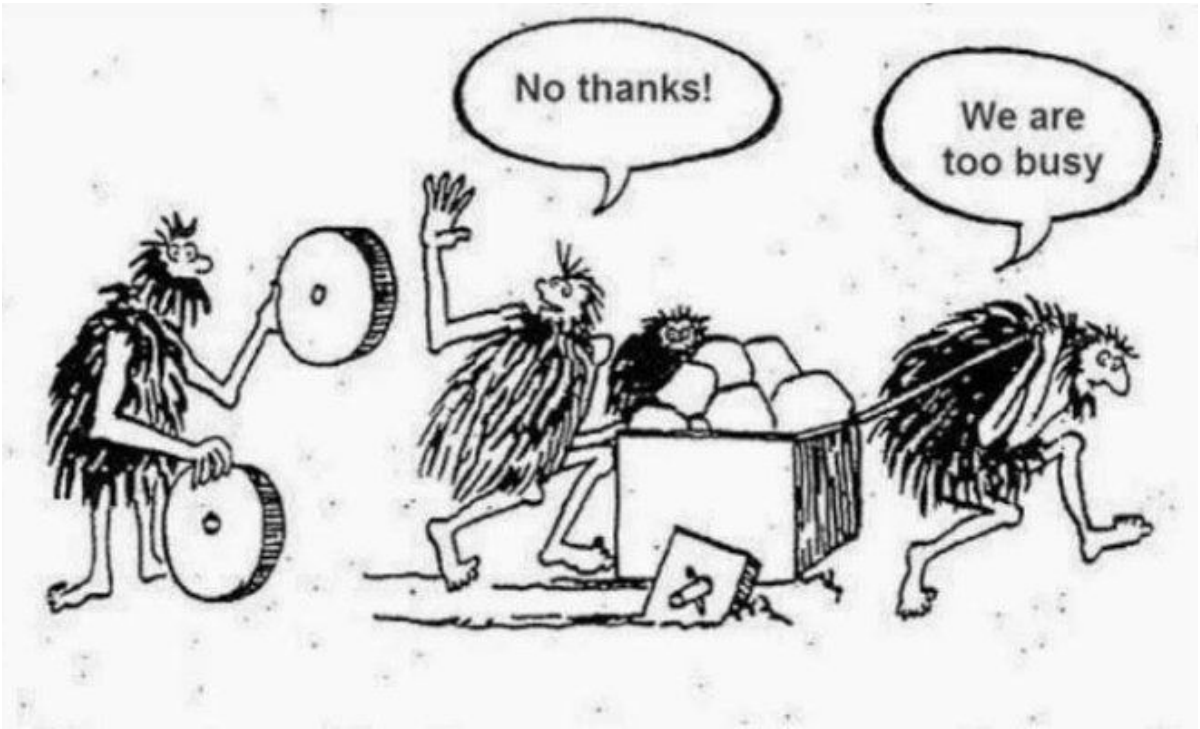
Reinventing the wheel



Reinventing the wheel

Pitfall #4

Believing Data is exempt from Software Engineering best practices



Reinventing the wheel

Pitfall #5

Letting Data Teams become a bottleneck

Reinventing the wheel

Pitfall #4

Believing Data is exempt from Software Engineering best practices

Pitfall #5

Letting Data Teams become a bottleneck

Reinventing the wheel

Pitfall #6

Thinking "one-shot" when moving Machine Learning to production

Reinventing the wheel

Pitfall #4

Believing Data is exempt from Software Engineering best practices

Data & ML Engineering : Software Engineering applied to Data

Pitfall #5

Letting Data Teams become a bottleneck

Data Mesh : Domain Driven Design applied to Data

Pitfall #6

Thinking "one-shot" when moving Machine Learning to production

MLOps : DevOps applied to Data

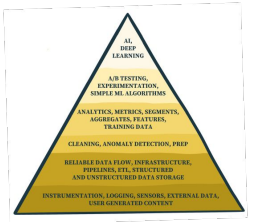
Takeaway

Data is Software.

Steal from software engineering and DevOps design patterns & best practices.

Questions to reflect on

Do you have a hard time knowing if you must invest more time in a PoC of if you should stop it ?



Did you generate trust on simple needs before building more complex Use Cases ?

Trap #3

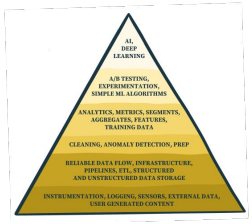
Rushing Headlong : Fear of Missing Out

Rushing Headlong - Fear of Missing Out

Pitfall #7

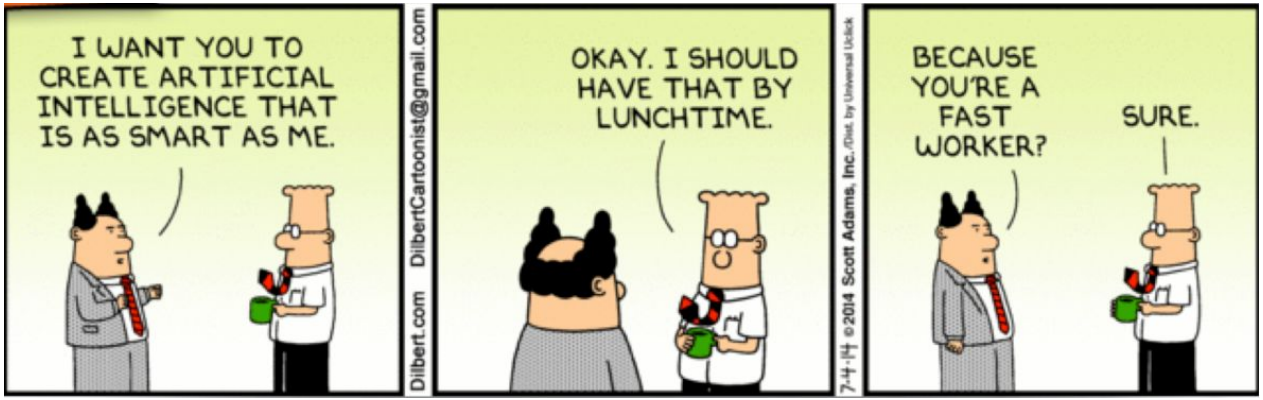
Falling in the never-ending-PoC trap

Rushing Headlong - Fear of Missing Out



Pitfall #8

Jumping too quickly into complex Use Cases

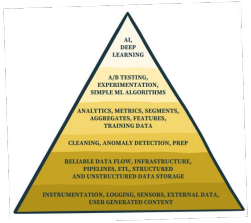


Rushing Headlong - Fear of Missing Out

Pitfall #7

Falling in the never-ending-PoC trap

A Data Product is only valuable if it brings business value



Pitfall #8

Jumping too quickly into complex Use Cases

Start with simple Use Cases, but end-to-end

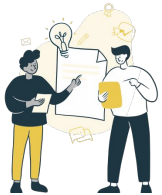
Takeaway

Data is Product.

Steal from Product Thinking and Product Management best practices.

Questions to reflect on

Do you all have the same definitions of roles, responsibilities and concepts concerning Data ?



Is there a “voice of data” in strategic instances and decisions ?

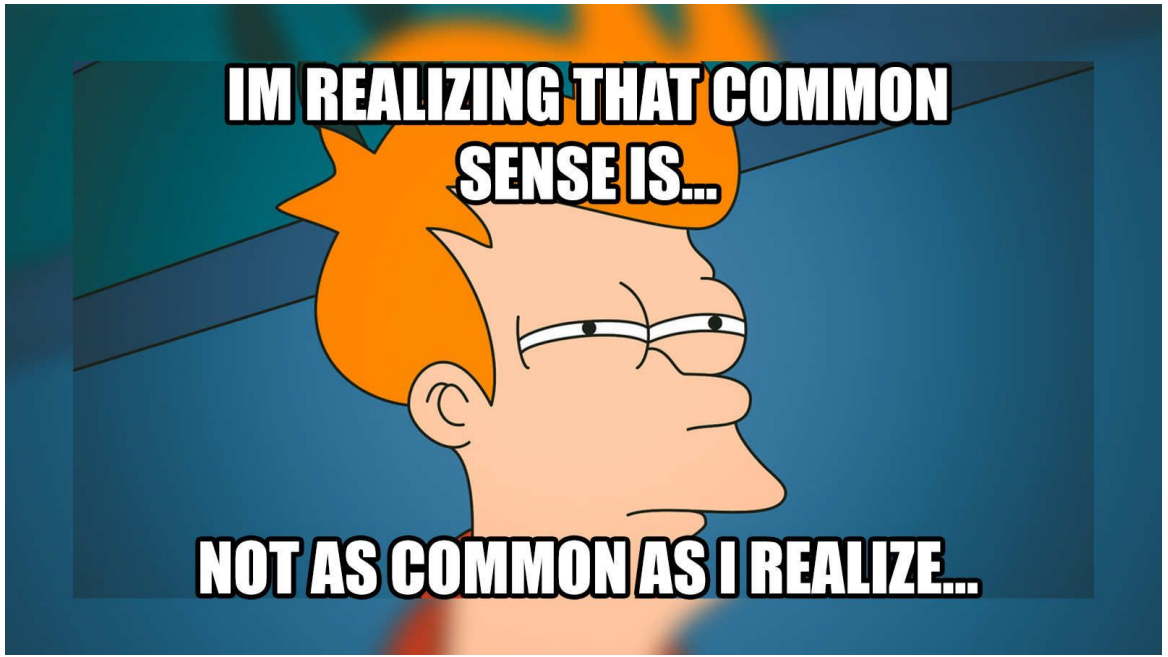
Trap #4

Feeding the Data Mystification

Feeding the Data Mystification

Pitfall #9

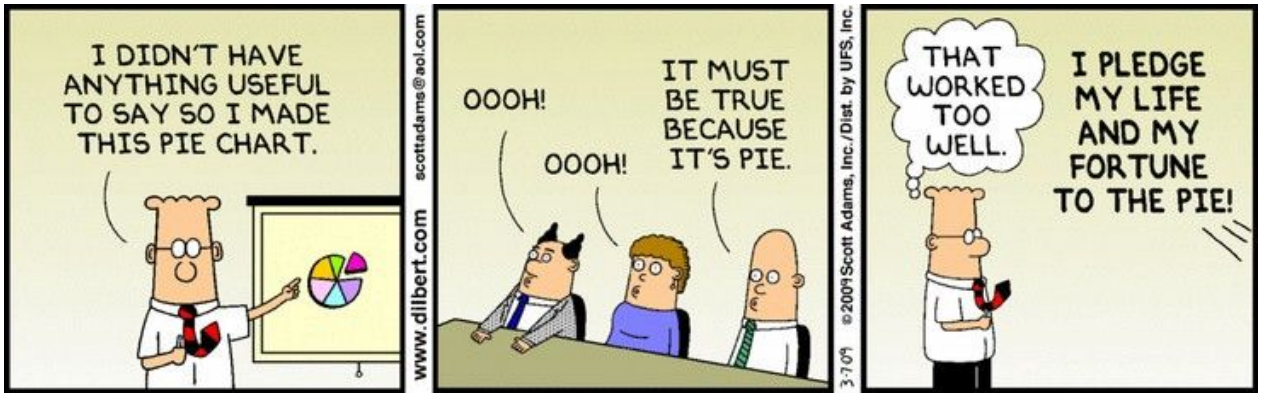
Believing Data Culture concerns only the Data Team



Feeding the Data Mystification

Pitfall #10

Not including Data in strategic decisions



Feeding the Data Mystification

Pitfall #9

Believing Data Culture concerns only the Data Team

Build a multi-level data acculturation program

Pitfall #10

Not including Data in strategic decisions

Give purpose to the Data Team's work

Takeaway

Data is Culture.

“The next era for Data & AI is to democratize access.”

Andrew Ng

Data Traps

Narrowing down Data to Data Science & Data Engineering

Data is PLURAL

Reinventing the wheel

Data is SOFTWARE

Rushing Headlong because of the fear of missing out

Data is PRODUCT

Feeding the Data Mystification

Data is CULTURE

10 pitfalls holding you back bringing value from data at scale

Not having a diverse enough Data Team

Build a diverse Data Team, including business and product skills

Believing data is trustworthy by default

Invest in Data Quality and Governance to generate trust

Letting Data Teams become a bottleneck

Data Mesh : Domain Driven Design applied to Data

Falling in the never-ending-PoC trap

A Data Product is only valuable if it brings business value

Believing Data Culture concerns only the Data Team

Build a multi-level data acculturation program

Being trapped in one's own biases

Don't create a team where everybody comes from the same background

Believing Data is exempt from Software Engineering best practices

Data & ML Engineering : Software Engineering applied to Data

Thinking "one-shot" when moving Machine Learning to production

MLOps : DevOps applied to Data

Jumping too quickly into complex Use Cases

Start with simple Use Cases, but end-to-end

Not including Data in strategic decisions

Give purpose to the Data Team's work

Conclusion

Data is not a goal in itself.

It's a means to move towards your purpose as a company.

Thank You !

E-Book (in french) from the talk