

Knowing All Data Science: a minutes long tour of the limits of the space of data science

Ron Neely

2018.Oct.18

<https://github.com/SpeedCoder5/notes>

Is it possible to know all of Data Science?

In minutes?

- What?
- How?
- Why?

Wide – not deep.

For whom?

- New to Data Science
- Business Folks - \$
- Seasoned Data Scientists

Agenda: a perspective – open to suggestions

Covered

What is data science?

- Learn: A method to learn fast and recall more
- The space of sci-kit learn
- A capability maturity model for data science
- Deliverables
- Limits
- Arguably: How NN, AI and ML fit it

Not Covered

Details of

- Neural Networks (NN)
- Artificial Intelligence (AI)
- Machine Learning (ML)

What is data science? Is there anything new? Has it all been done before?

[Wikipedia](#): **Data science** is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from data in various forms, both structured and unstructured.

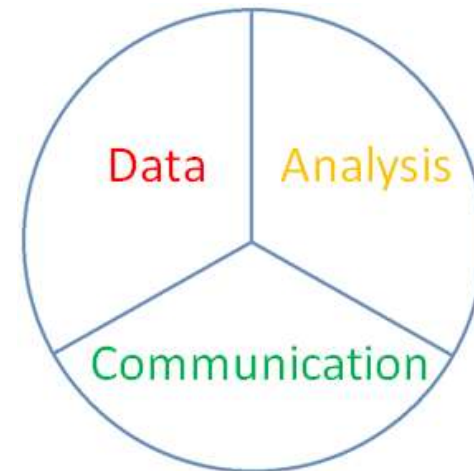
Anything *missing* from above definition?

What is data science? Is there anything new? Has it all been done before?

[Wikipedia](#): **Data science** is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from data in various forms, both structured and unstructured.

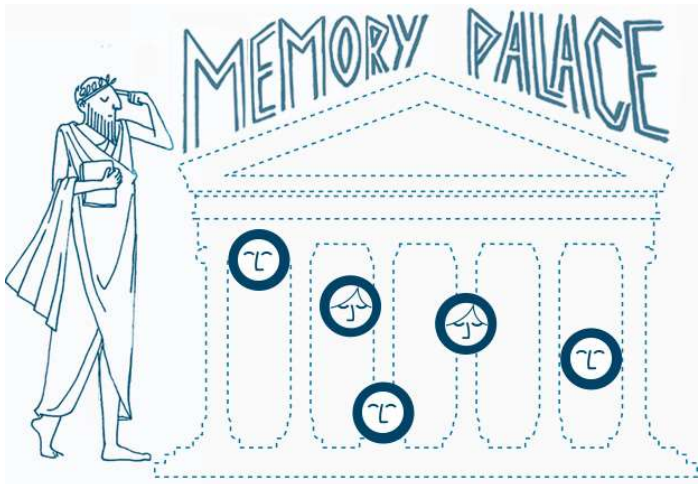
- **Data** => Information => Knowledge
- **Analysis**
 - Visualization
 - Predictive Model
- **Communication**
 - Advice
- **Repeat**

Data Science Wheel



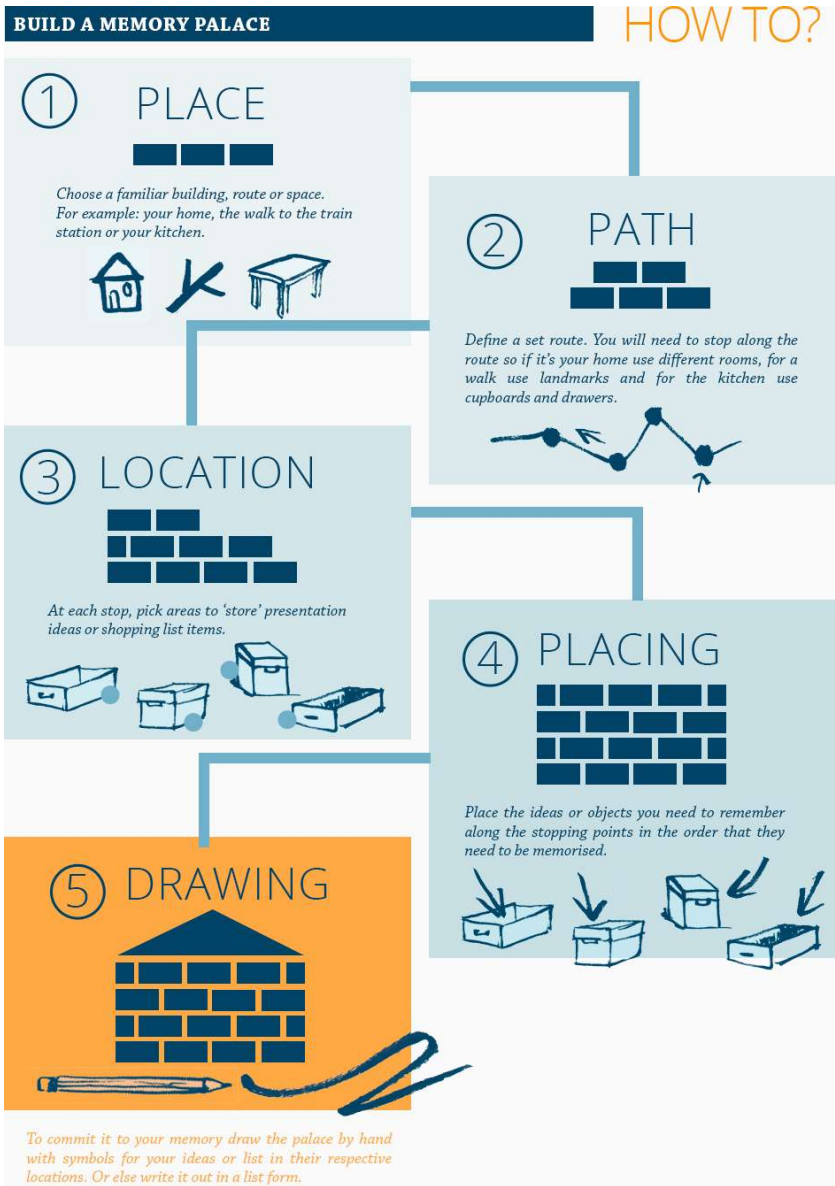
How do you learn quickly? How do you remember what you learned?

Method of Loci - Learn faster - Recall more

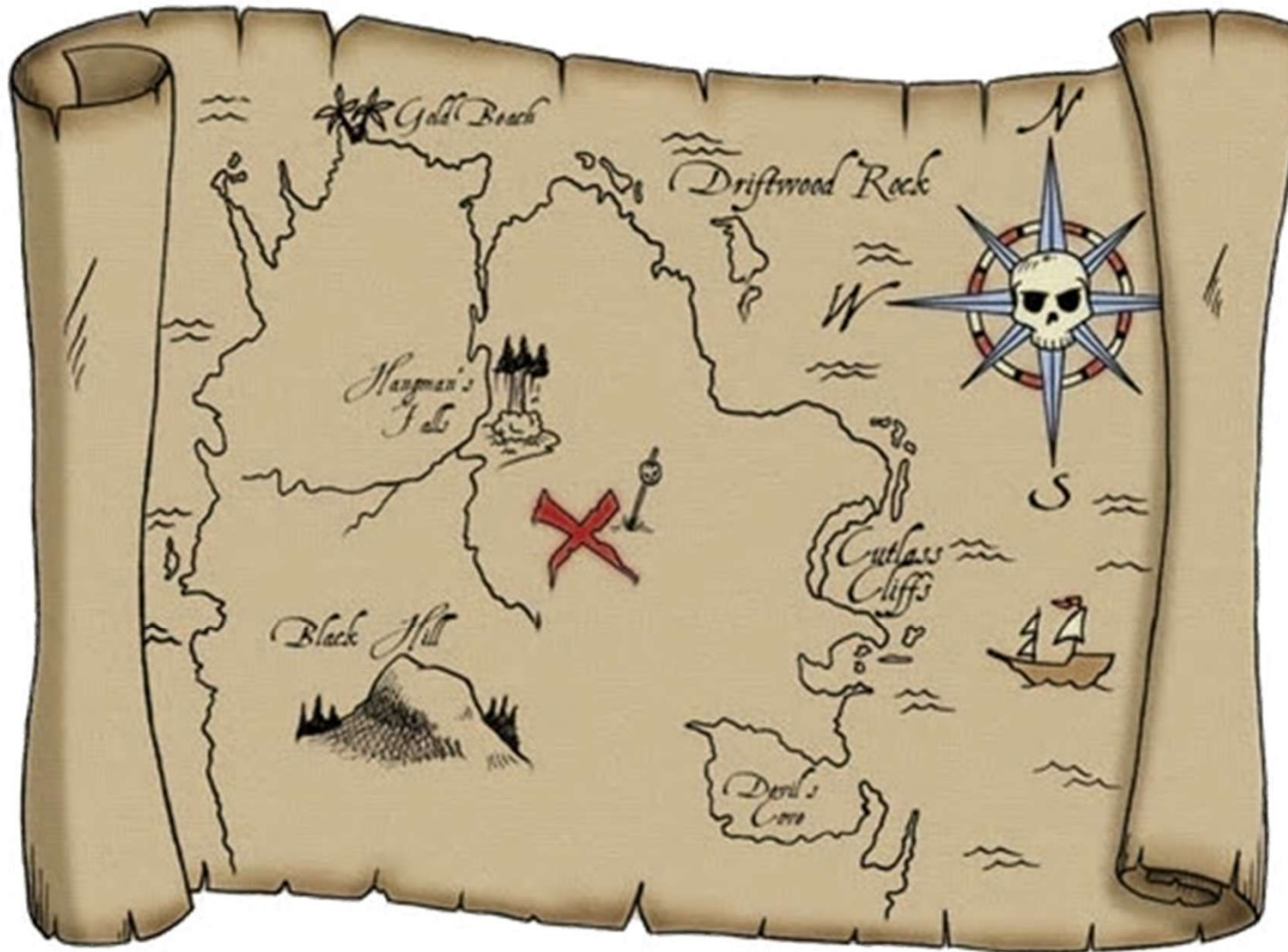


- Who? Ancient Greek poet: Simonides of Creos
- Are we losing geo-spatial capability?
- Use mental “maps” to get to data:
Getting to a desired destination requires both:
 - start point
 - AND basis with origin

<http://advertisementfeature.cnn.com/think-brilliant/memory-palace.html>

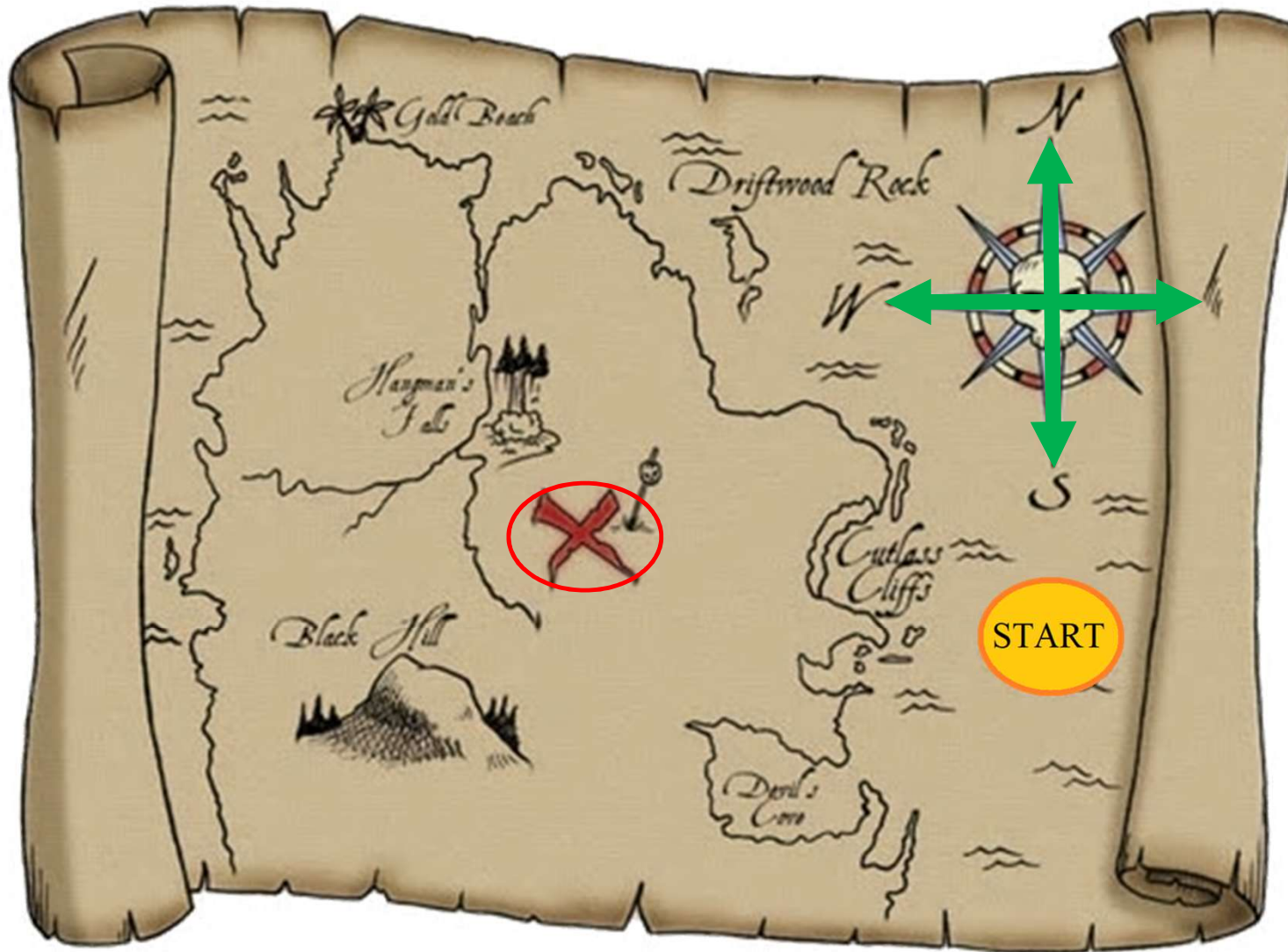


Method of Loci Map



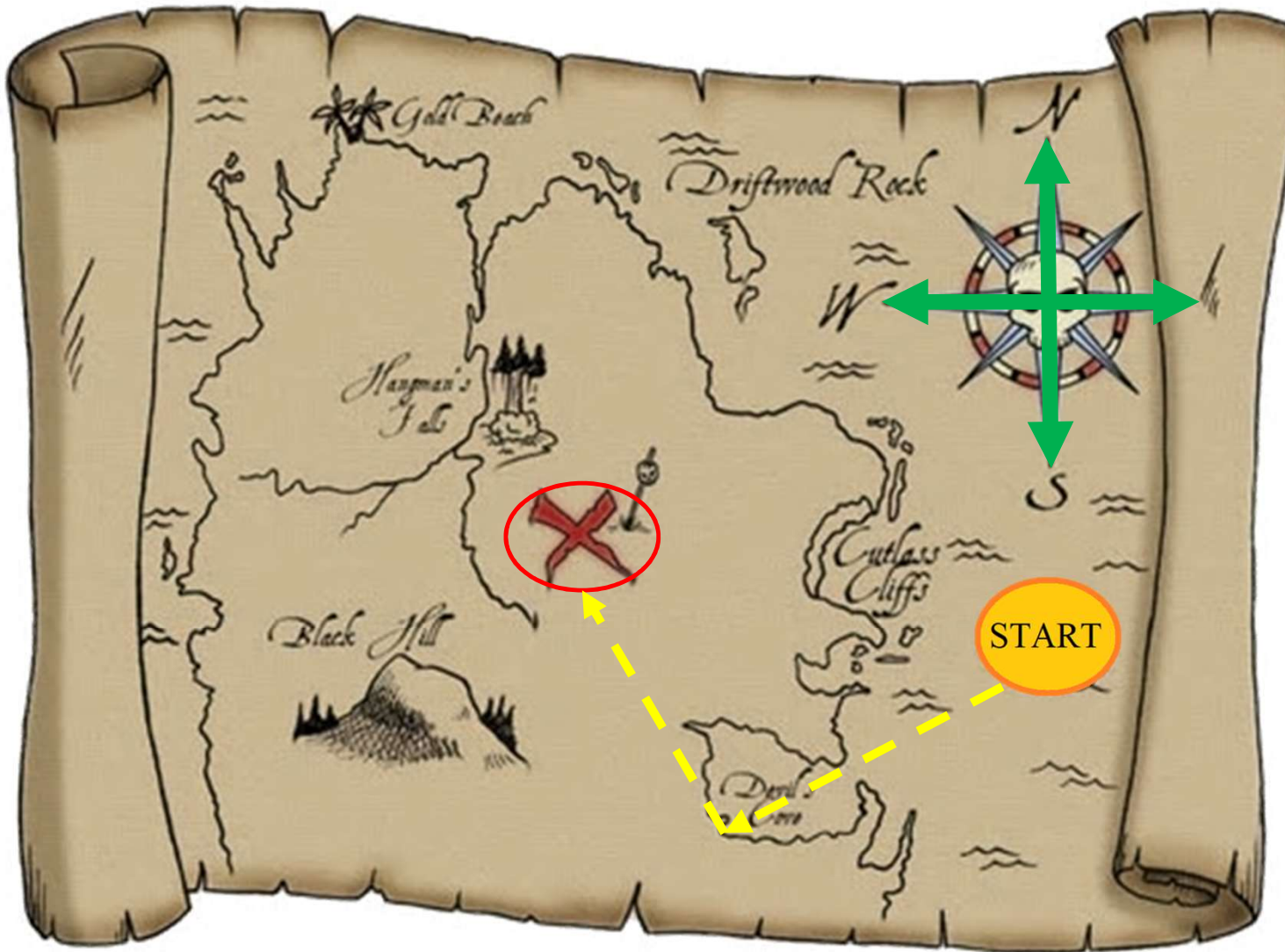
✓ Drawing

Method of Loci Map



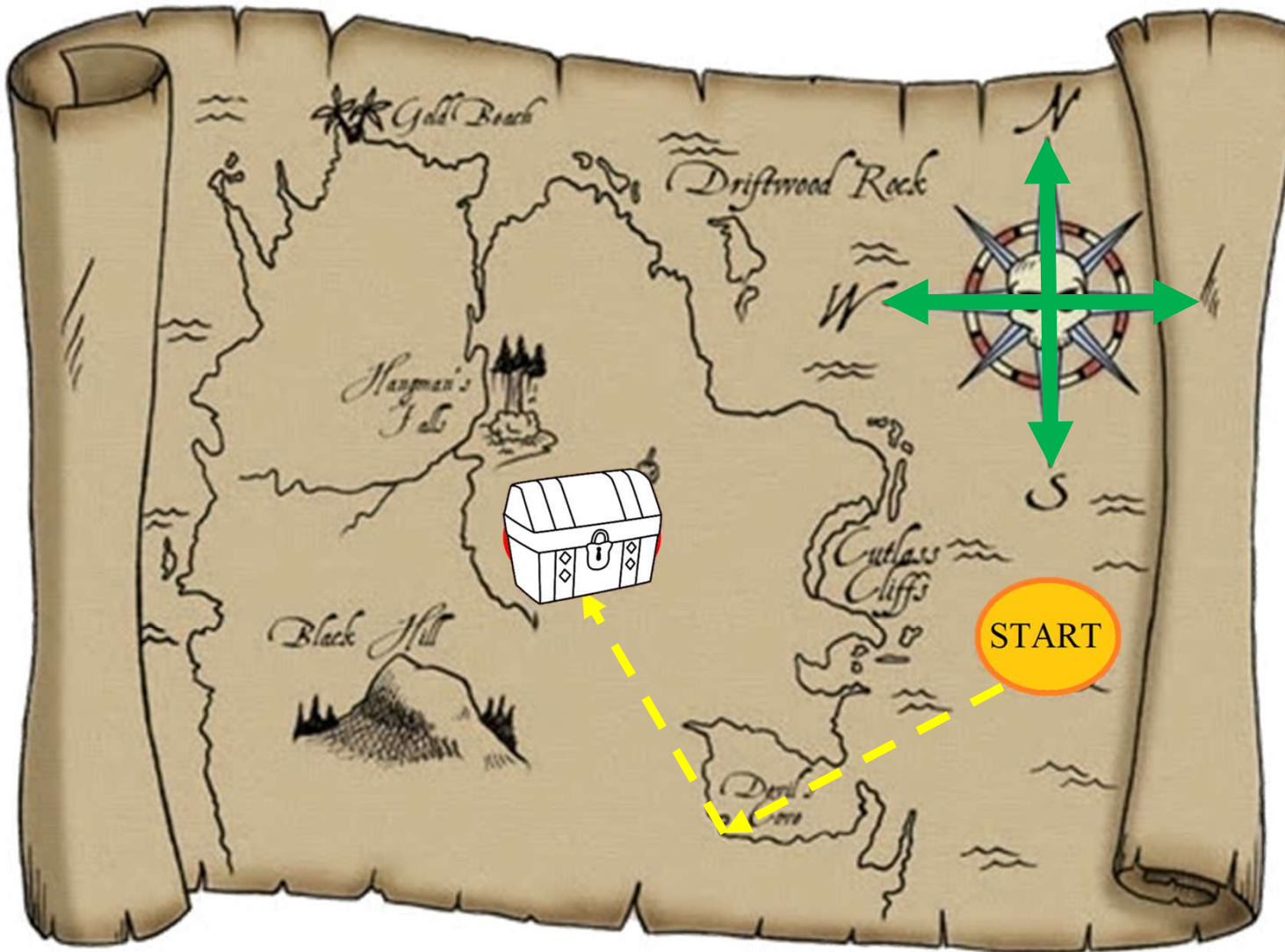
- ✓ Drawing
- ✓ Place

Method of Loci Map



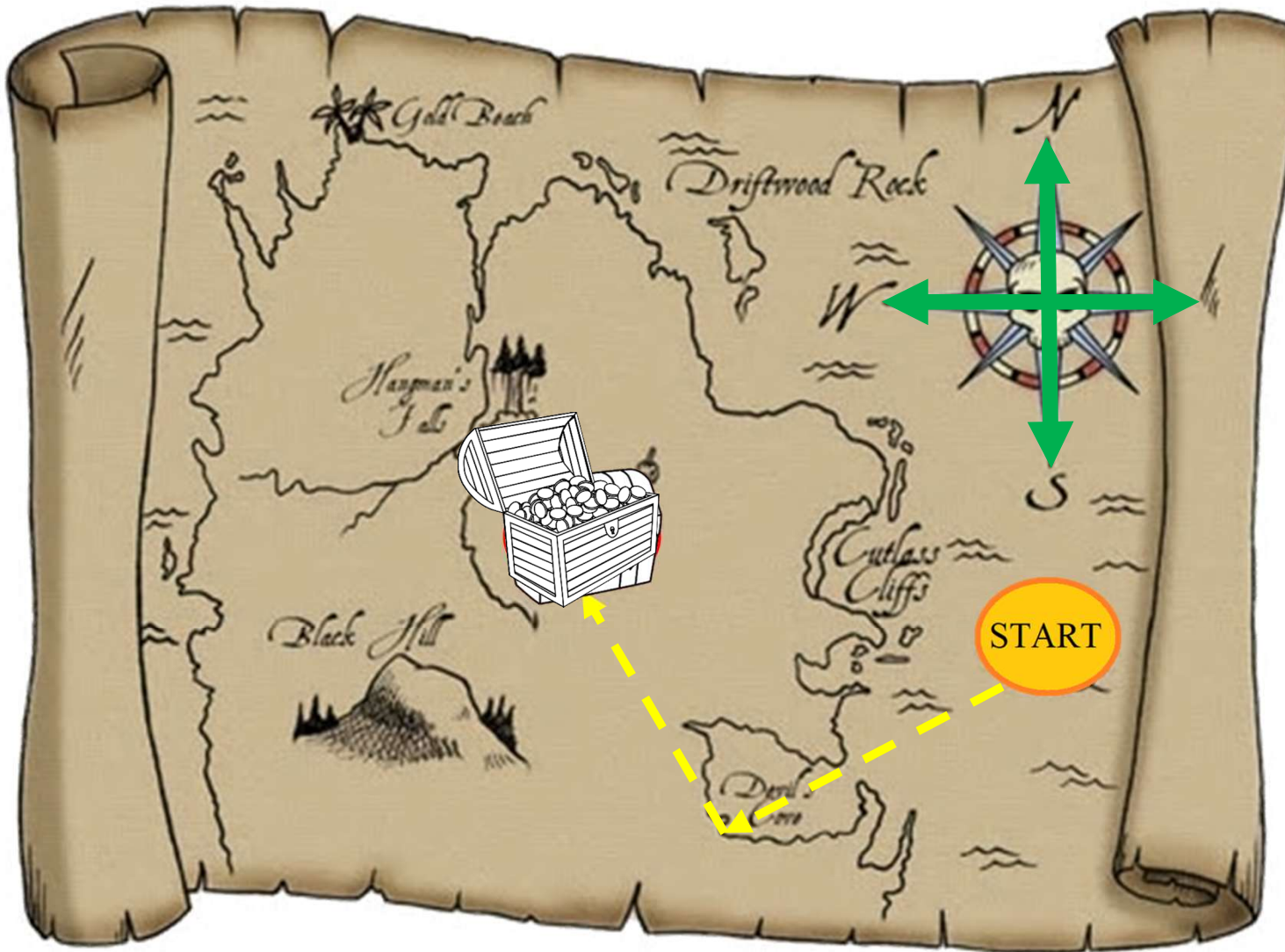
- ✓ Drawing
- ✓ Place
- ✓ Path

Method of Loci Map



- ✓ Drawing
- ✓ Place
- ✓ Path
- ✓ Location

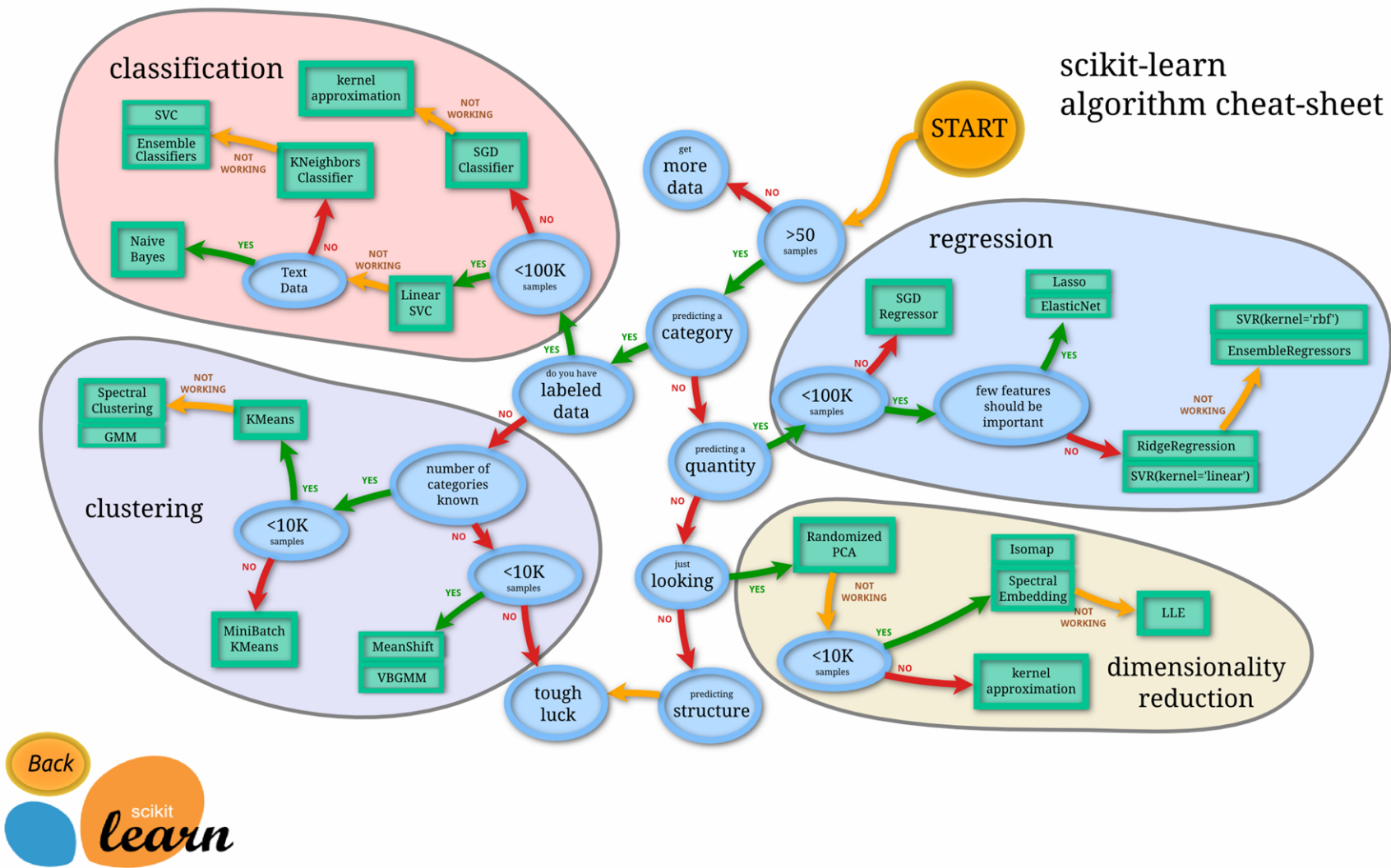
Method of Loci Map



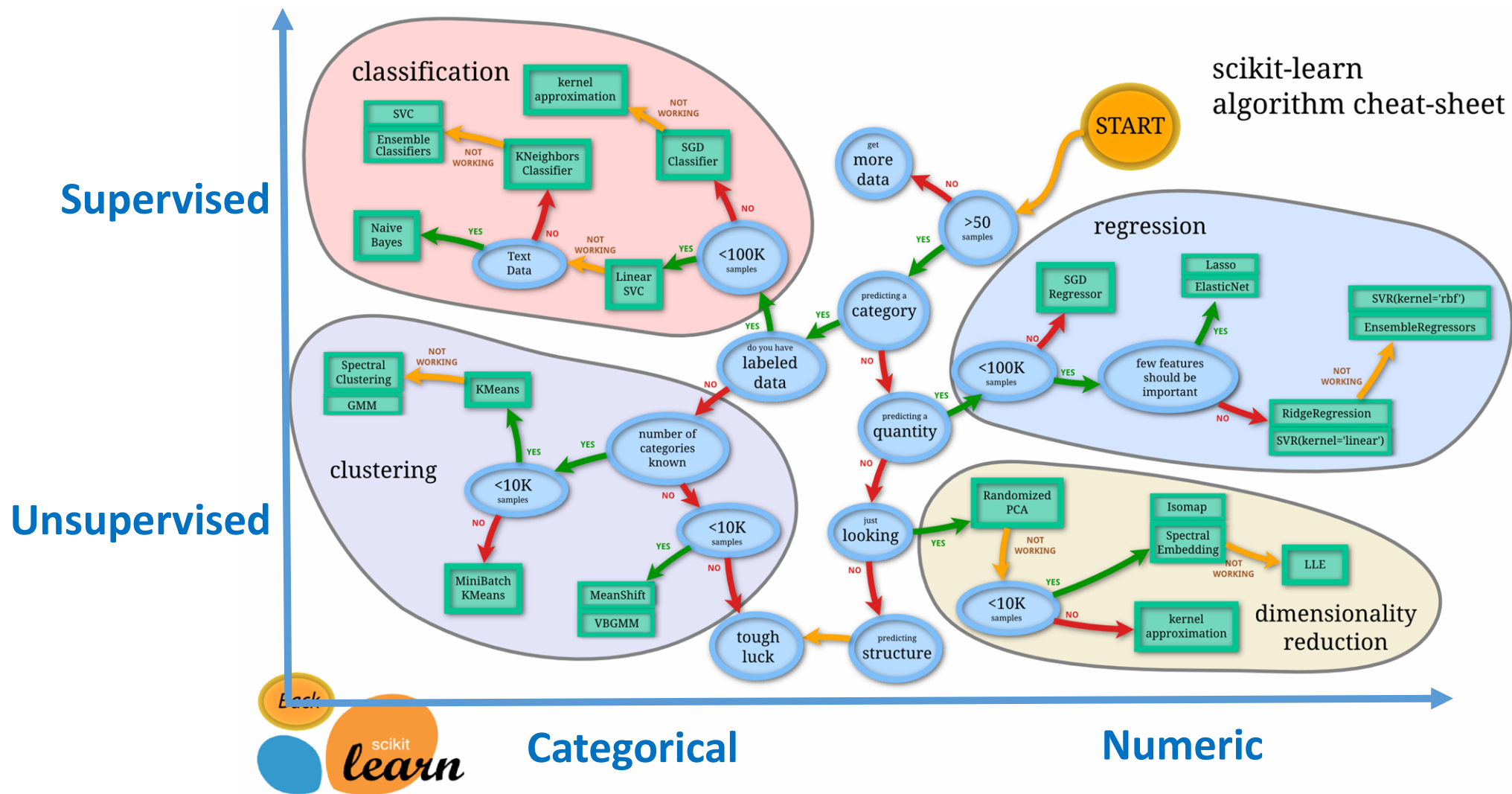
- ✓ Drawing
- ✓ Place
- ✓ Path
- ✓ Location
- ✓ Placing

Scikit-learn's "butterfly map"

scikit-learn
algorithm cheat-sheet



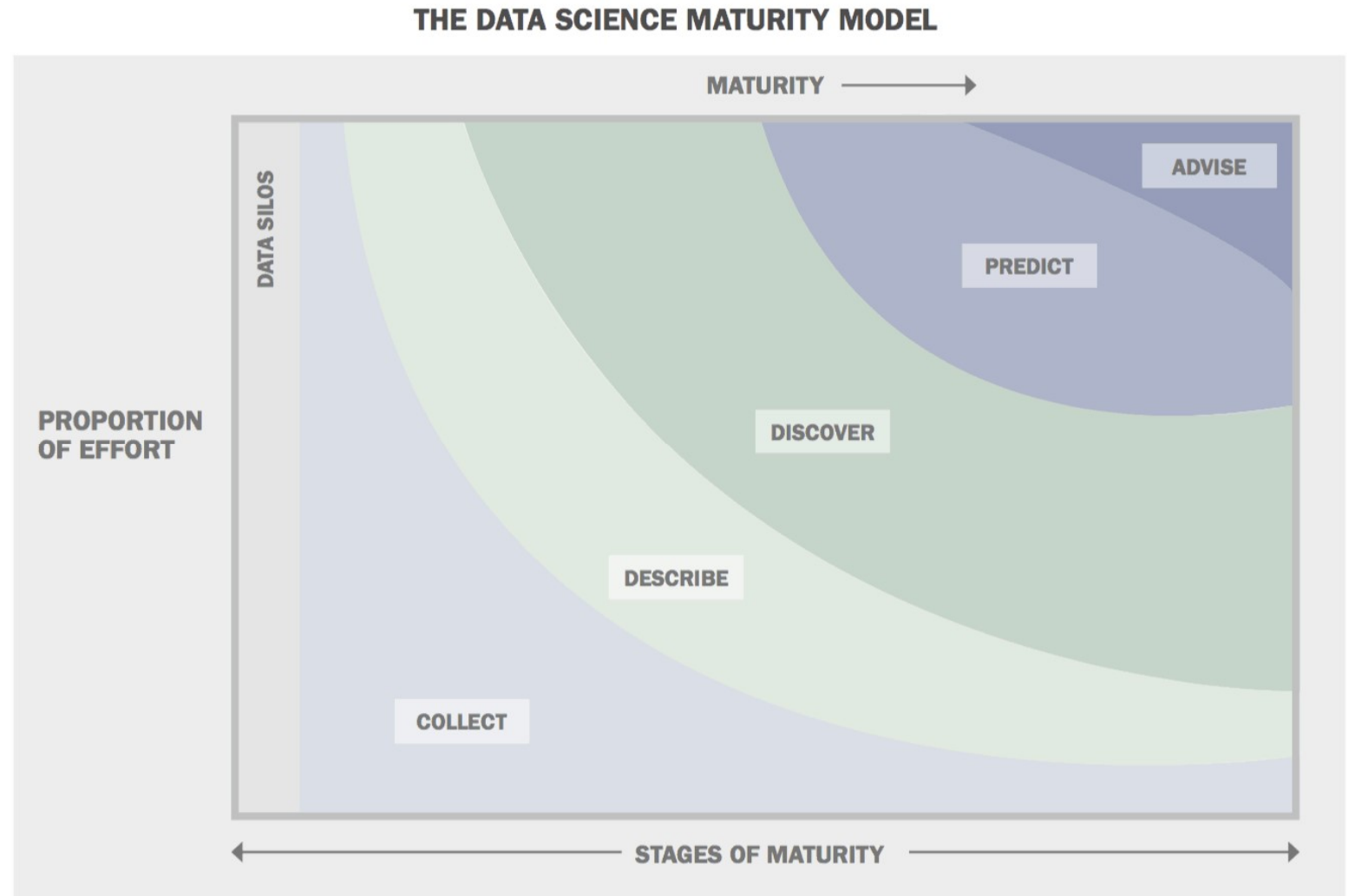
Scikit-learn's "butterfly map" – with origin and very loosely-defined basis



Booz Allen Hamilton's (BAH's) [Field Guide to Data Science](#)

Where are you?

- Individual
- Organization
- Per Project

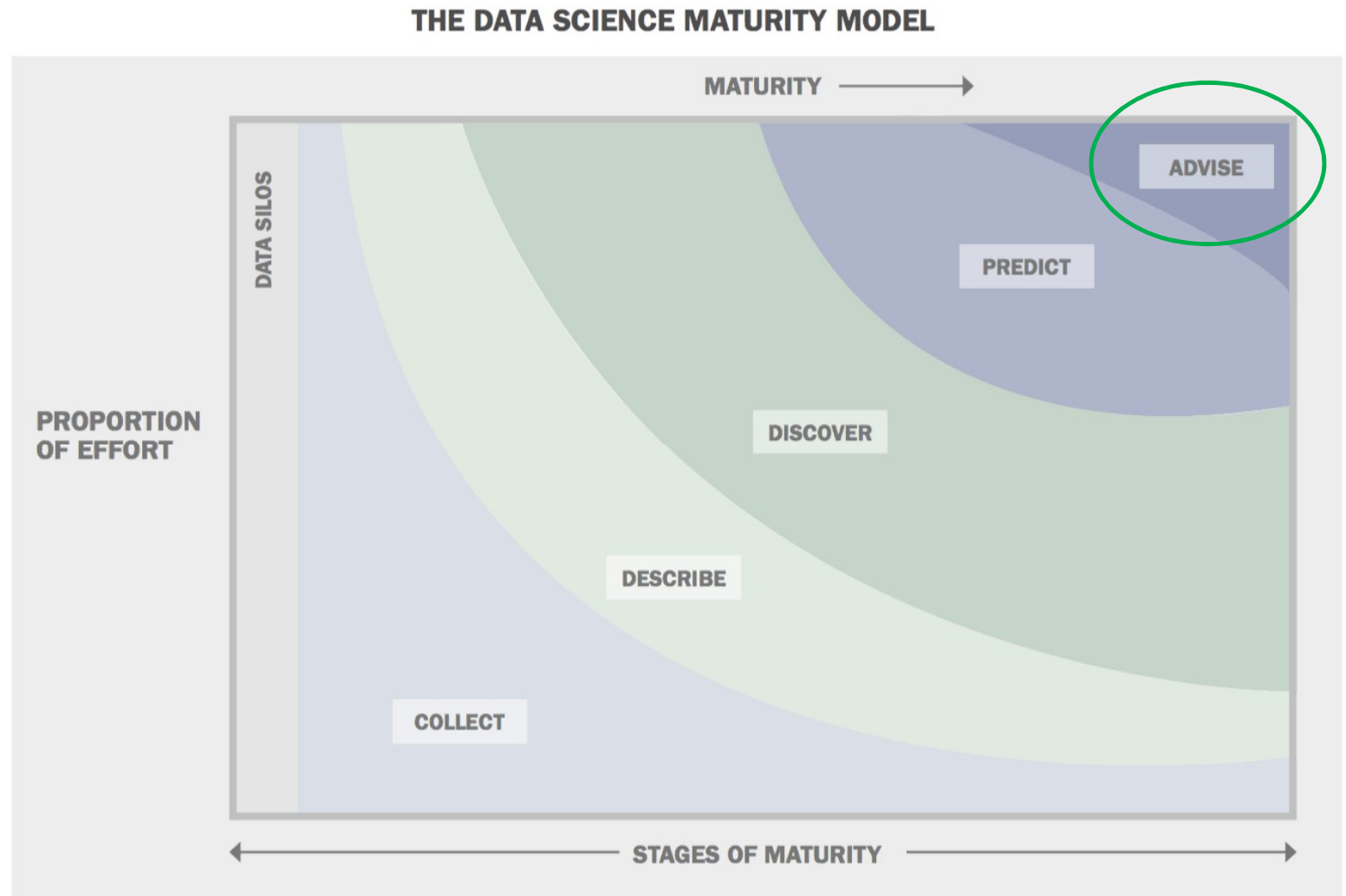
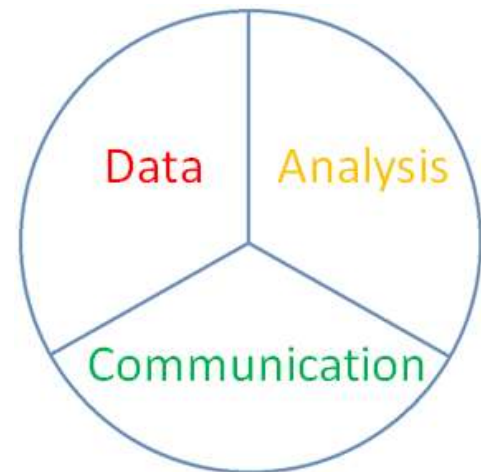


Booz Allen Hamilton's (BAH's) [Field Guide to Data Science](#)

Where are you?

- Individual
- Organization
- Per Project

Goal: **Advice**

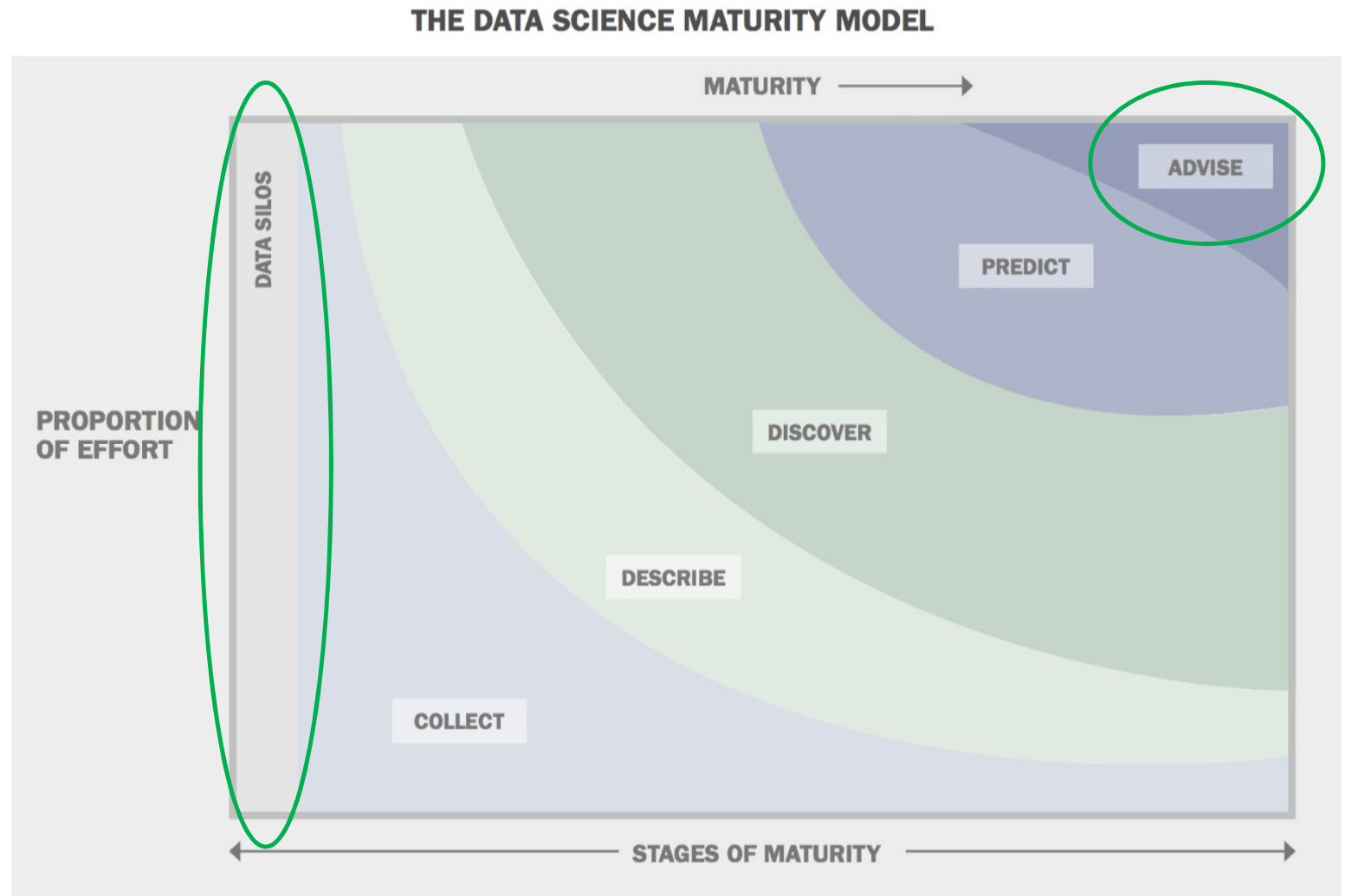
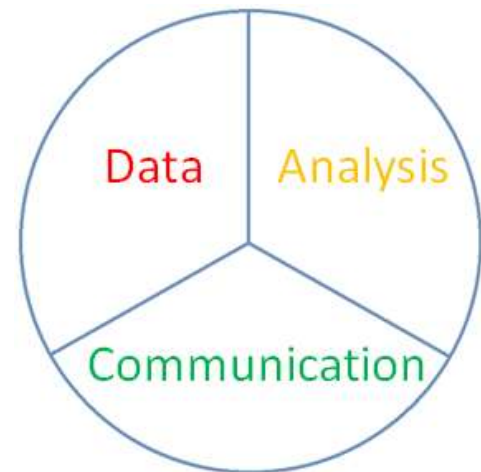


Booz Allen Hamilton's (BAH's) [Field Guide to Data Science](#)

Where are you?

- Individual
- Organization
- Per Project

Goal: **Advice**

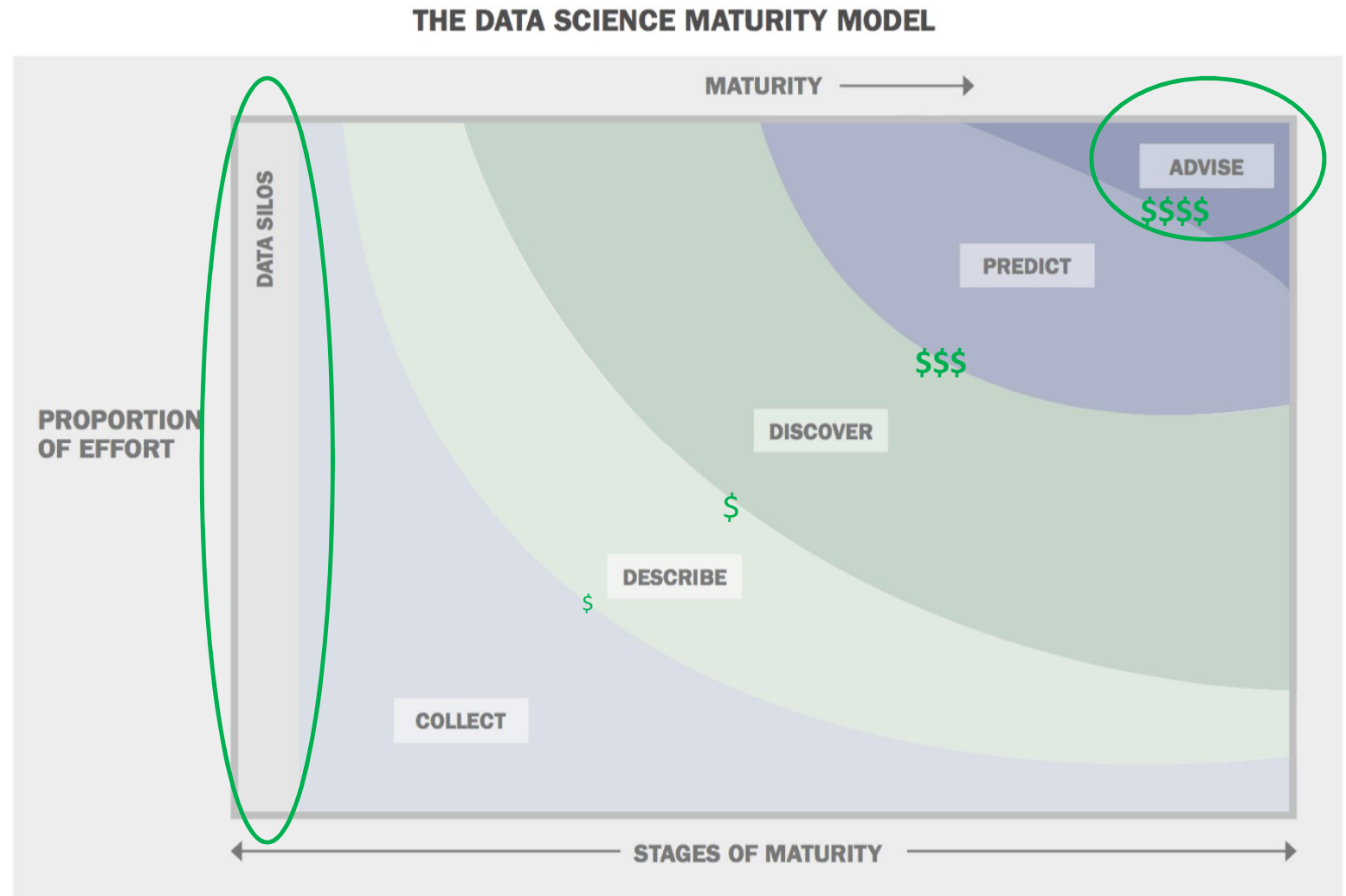
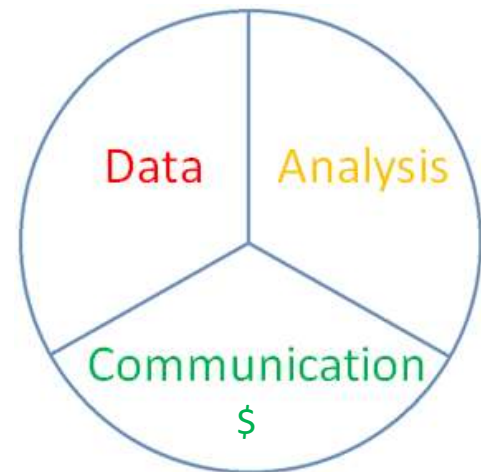


Booz Allen Hamilton's (BAH's) [Field Guide to Data Science](#)

Where are you?

- Individual
- Organization
- Per Project

Goal: **Advice**



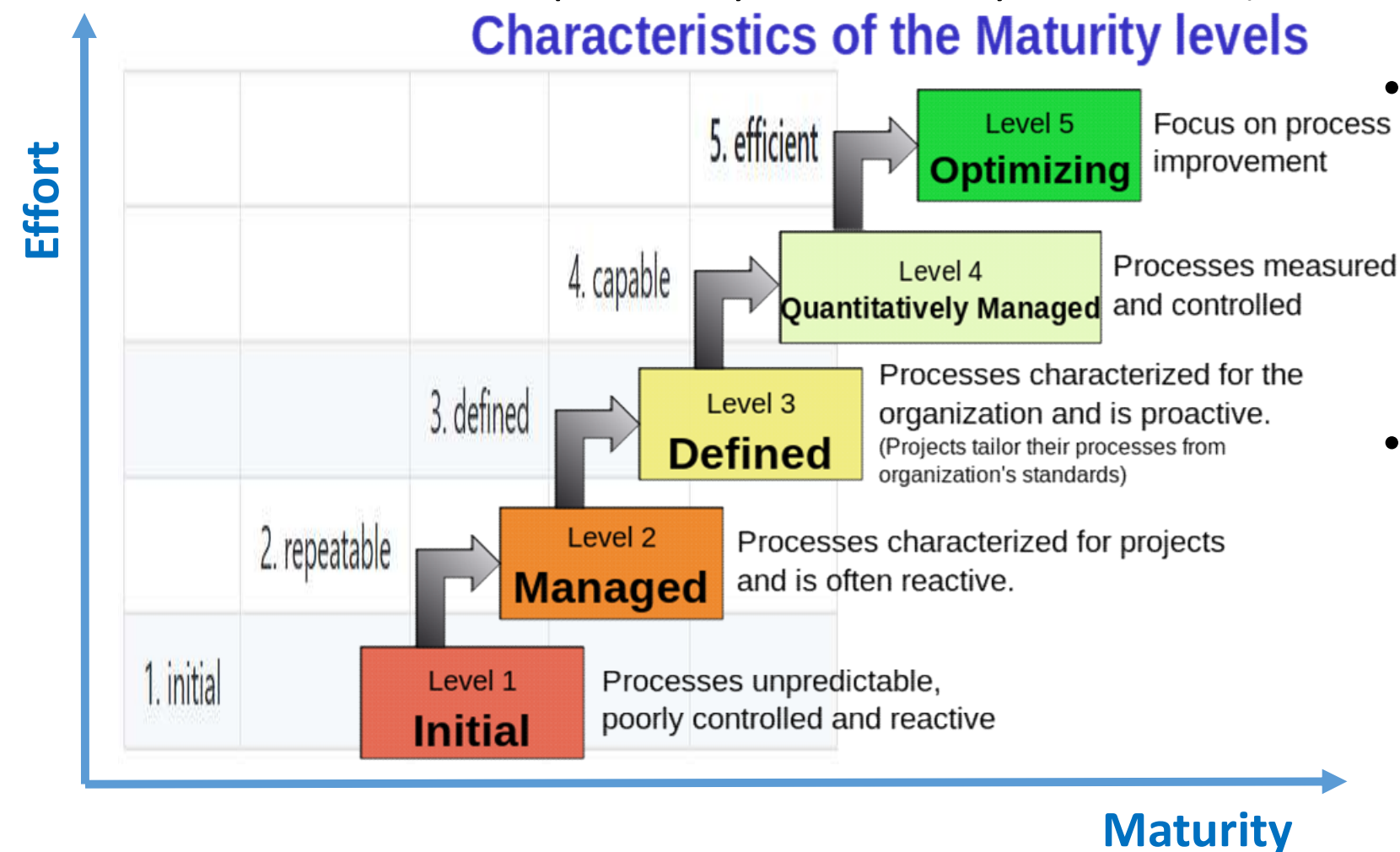
BAH Field Guide:

Data Science is the art of turning data into actions:

- Tradecraft *creates **data products***
 - Drive actions – **advice** – possibility and danger
- Supports and encourages shifting between:
 - **Deductive** (hypothesis-based) reasoning
 - **Inductive** (pattern-based) reasoning
- Necessary for companies to **compete**
- Capabilities **can be built** over time
- Different kind of team sport
 - Company-wide *open culture* – commitment to insights from data

Vs. DoD Process Capability Maturity Model (CMM) Levels?

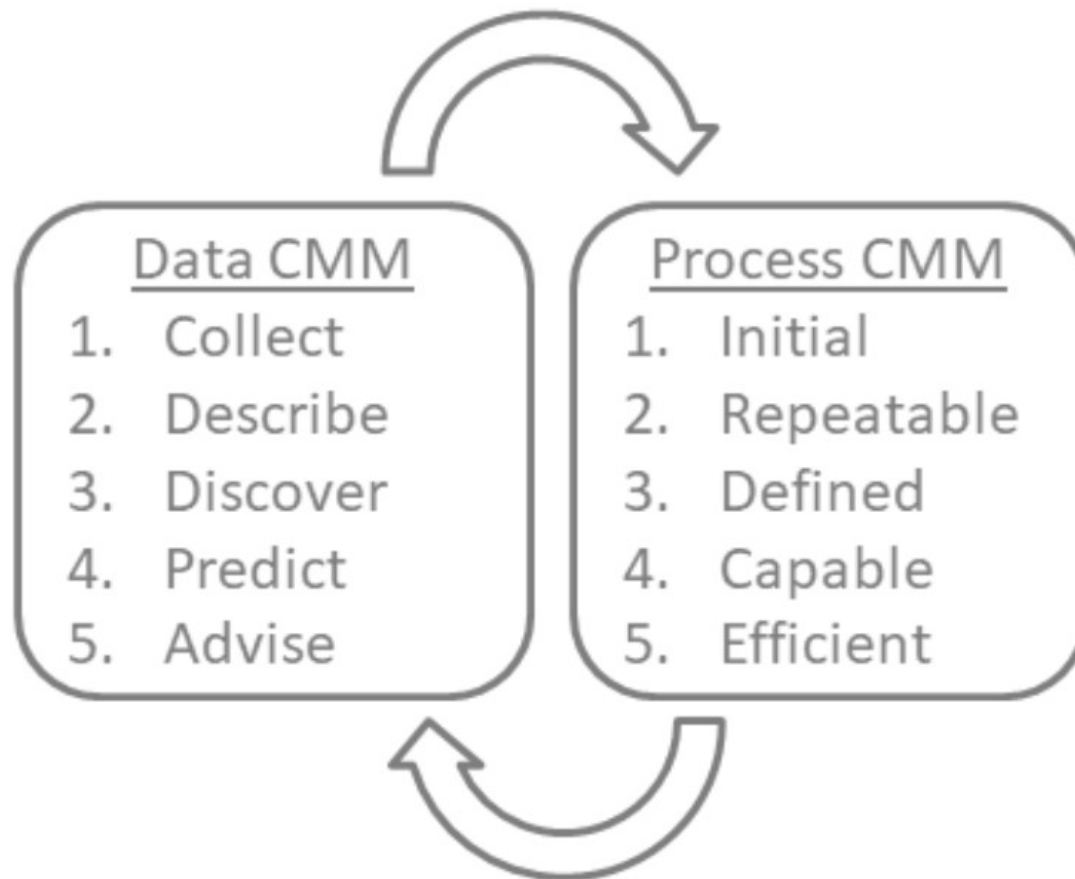
Characteristics of the Maturity levels



- Earned Value Management = value earned from work being done dau.mil
- Plan your work
work your plan

source: wikipedia

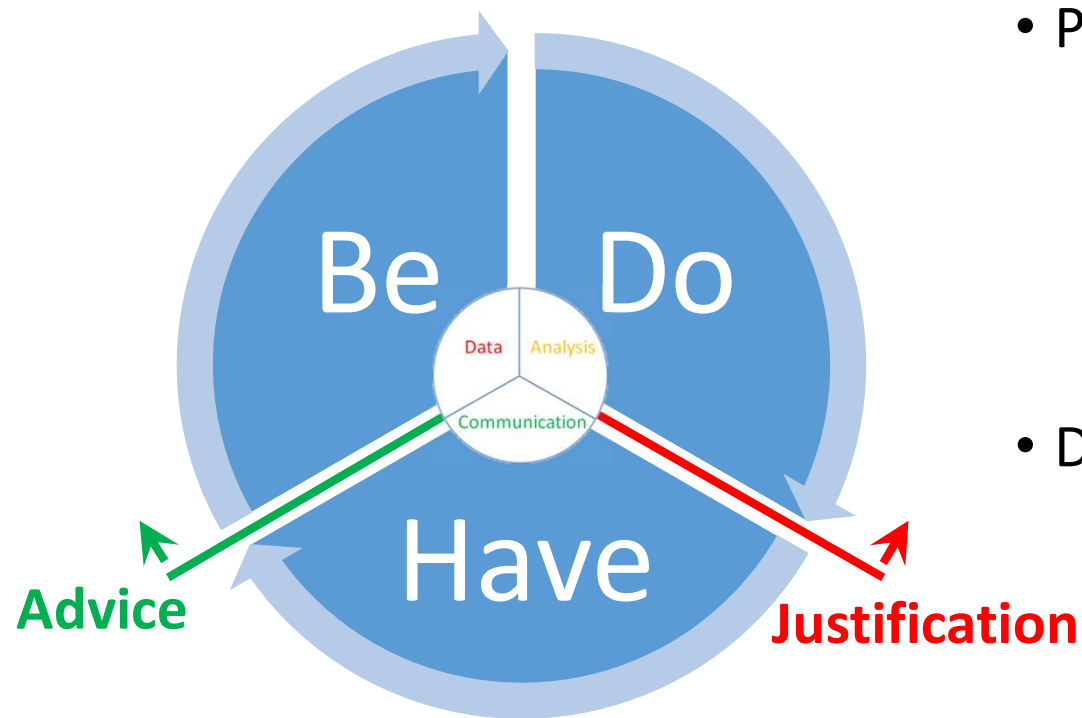
BAH Data CMM and DOD Process CMM Complement Each Other



“Noun-ified” Deliverables for Data Science Efforts

	Step	"Noun-ified"
1.	Collect	Collection
2.	Describe	Description
3.	Discover	Discovery
4.	Predict	Prediction
5.	Advise	Advice

Limits of Data Science – States of Being



- Possibility - clockwise (Opportunity):
 - DS provides Clarity, Predictions, **Advice**
 - You declare where to go (what to be) next
 - Data is only a starting point
 - Be -> Do -> Have
 - Empowering
- Danger - counterclockwise (Constraint):
 - DS provides **justification** to constrain being
 - Data limits where you can go next
 - Data becomes beginning and ending
 - Have -> Do -> Be
 - Limiting

Thank you

What we covered:

- What
- How
- Why
- Limits

of Data Science

What did you learn? Any questions?

Ron Neely

<https://github.com/SpeedCoder5/notes>