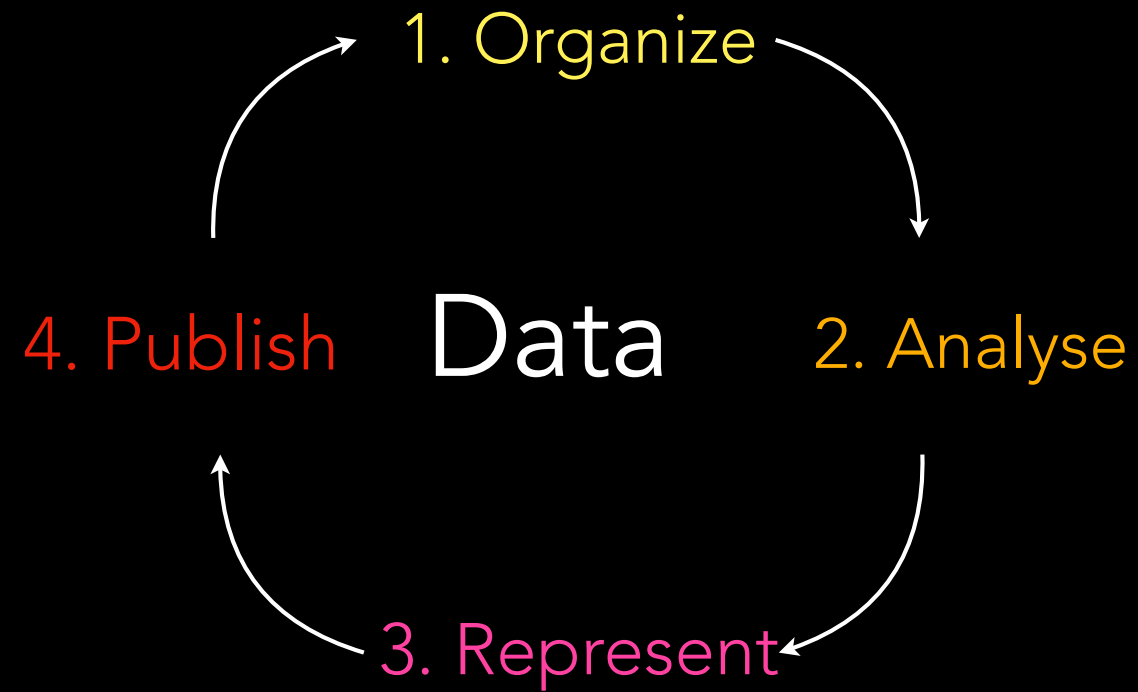


Make data meaningful

Why bother with data organization, analysis and representation?

Thomas Pucadyil
IISER Pune





**Why
organise
data?**

Prerequisite for analysis

- Can I trust what my data tells me?
- How many repeats do I have?
- Is this statistically significant?

**Why
organise
data?**

Why organise data?

Prerequisite for analysis

- Can I trust what my data tells me?
- How many repeats do I have?
- Is this statistically significant?

Allows you to recollect

- Any data is only as good as your ability to know that it exists
- Organisation helps convert data from information to knowledge

Why organise data?

Prerequisite for analysis

- Can I trust what my data tells me?
- How many repeats do I have?
- Is this statistically significant?

Allows you to recollect

- Any data is only as good as your ability to know that it exists
- Organisation helps convert data from information to knowledge

Self-awareness

- Where am I with work?
- What have I done so far?
- How much more do I need to do?

Why organise data?

Prerequisite for analysis

- Can I trust what my data tells me?
- How many repeats do I have?
- Is this statistically significant?

Brings perspective

- Facilitates model building
- If I see this, this, this....then I can assume the following

Allows you to recollect

- Any data is only as good as your ability to know that it exists
- Organisation helps convert data from information to knowledge

Self-awareness

- Where am I with work?
- What have I done so far?
- How much more do I need to do?

Why organise data?

Prerequisite for analysis

- Can I trust what my data tells me?
- How many repeats do I have?
- Is this statistically significant?

Brings perspective

- Facilitates model building
- If I see this, this, this....then I can assume the following

Allows you to recollect

- Any data is only as good as your ability to know that it exists
- Organisation helps convert data from information to knowledge

Invites collaborations

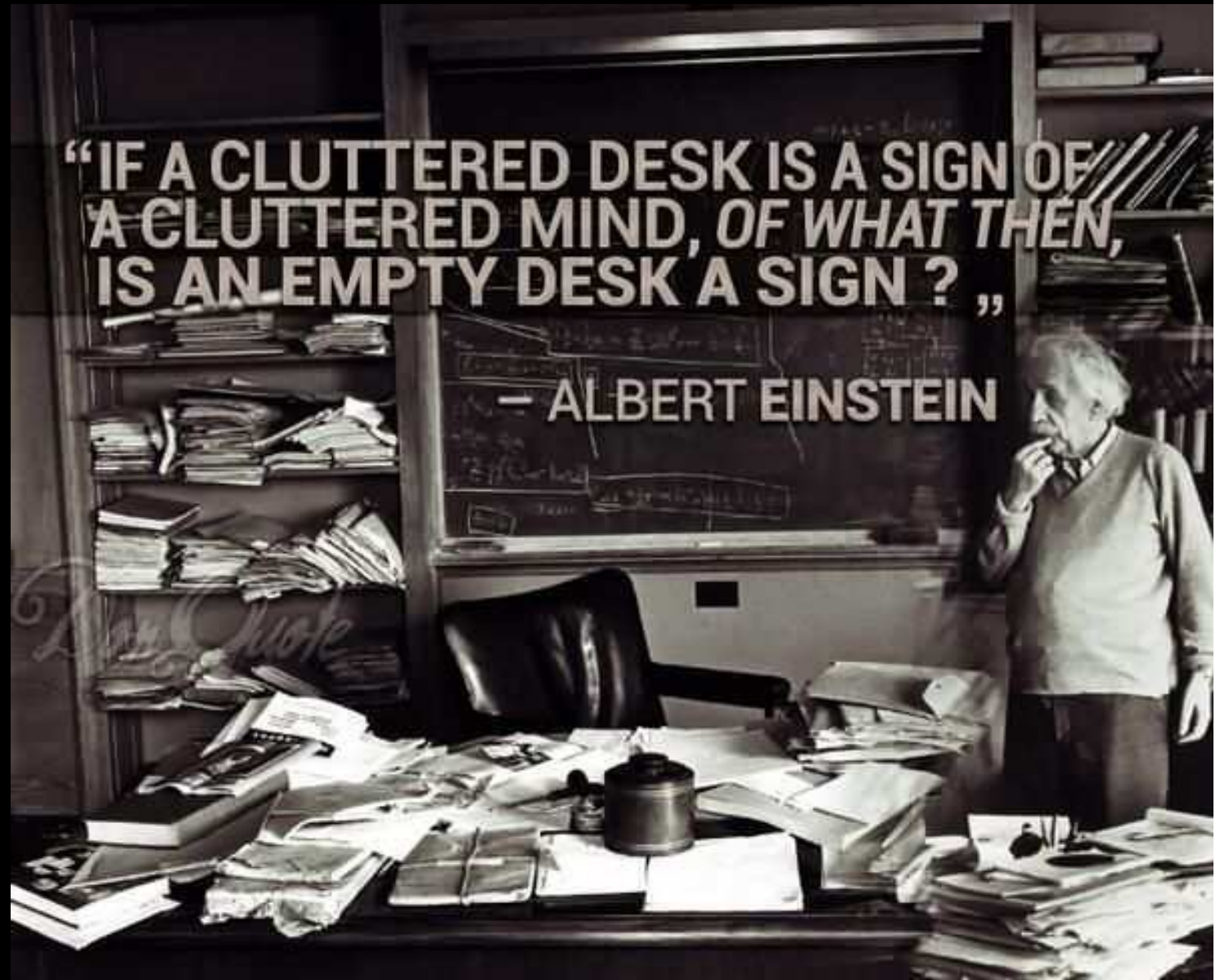
- Organised data draws a better crowd than disorganised data
- Why should someone else care if you don't?
- Sloppy organisation indicates lack of respect for your data

Self-awareness

- Where am I with work?
- What have I done so far?
- How much more do I need to do?

**"IF A CLUTTERED DESK IS A SIGN OF
A CLUTTERED MIND, OF WHAT THEN,
IS AN EMPTY DESK A SIGN ? ,,"**

— ALBERT EINSTEIN





Reader's Digest

A Trusted Friend in a Complicated World

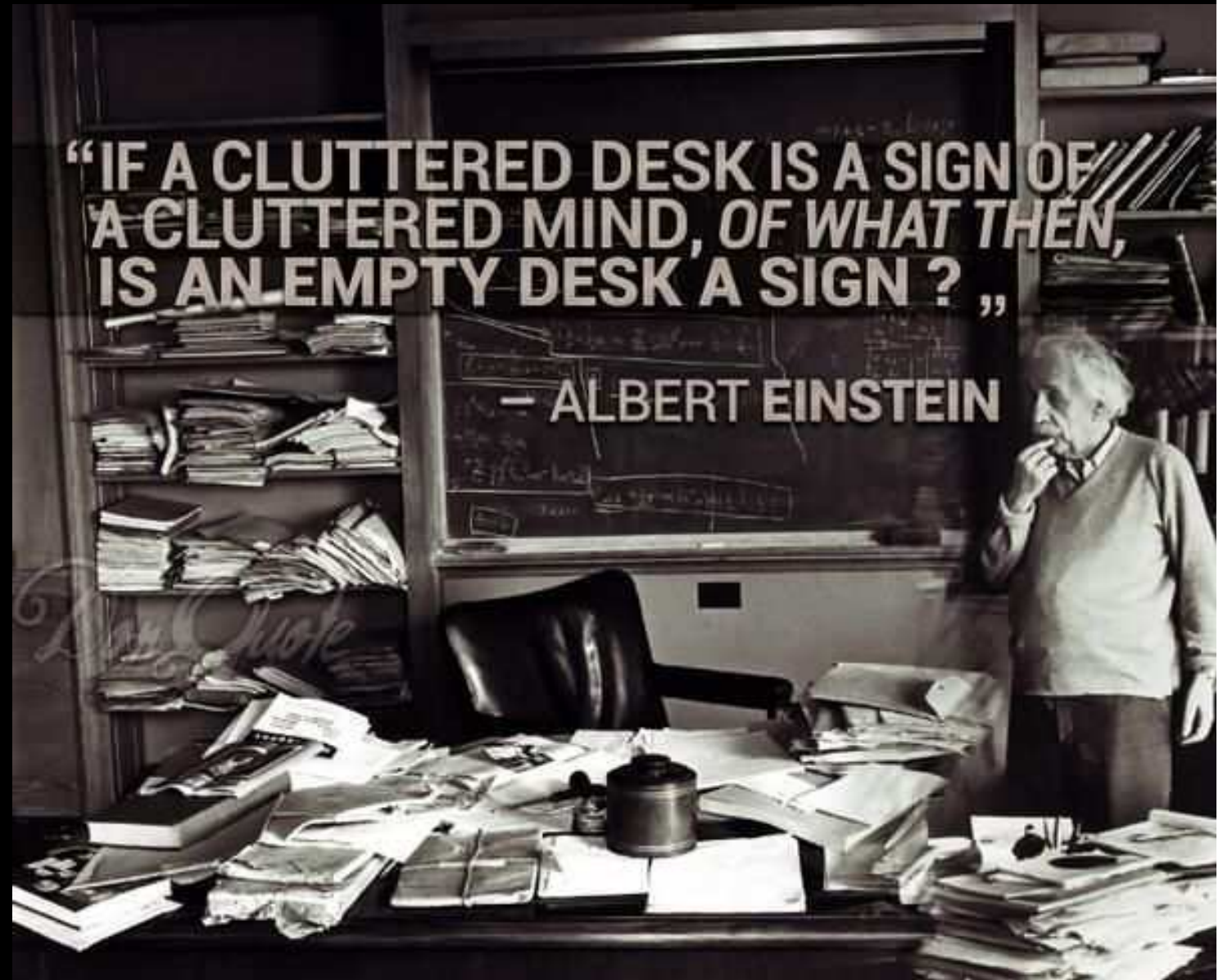
RD.COM ~ Knowledge ~ Psychology

If Your Desk Is a Mess, You're Probably a Genius



Brandon Specktor

Updated: Jun. 22, 2022





Reader's Digest

A Trusted Friend in a Complicated World

RD.COM ~ Knowledge ~ Psychology

If Your Desk Is a Mess, You're Probably a Genius



Brandon Specktor

Updated: Jun. 22, 2022

Smithsonian
MAGAZINE

INNOVATION

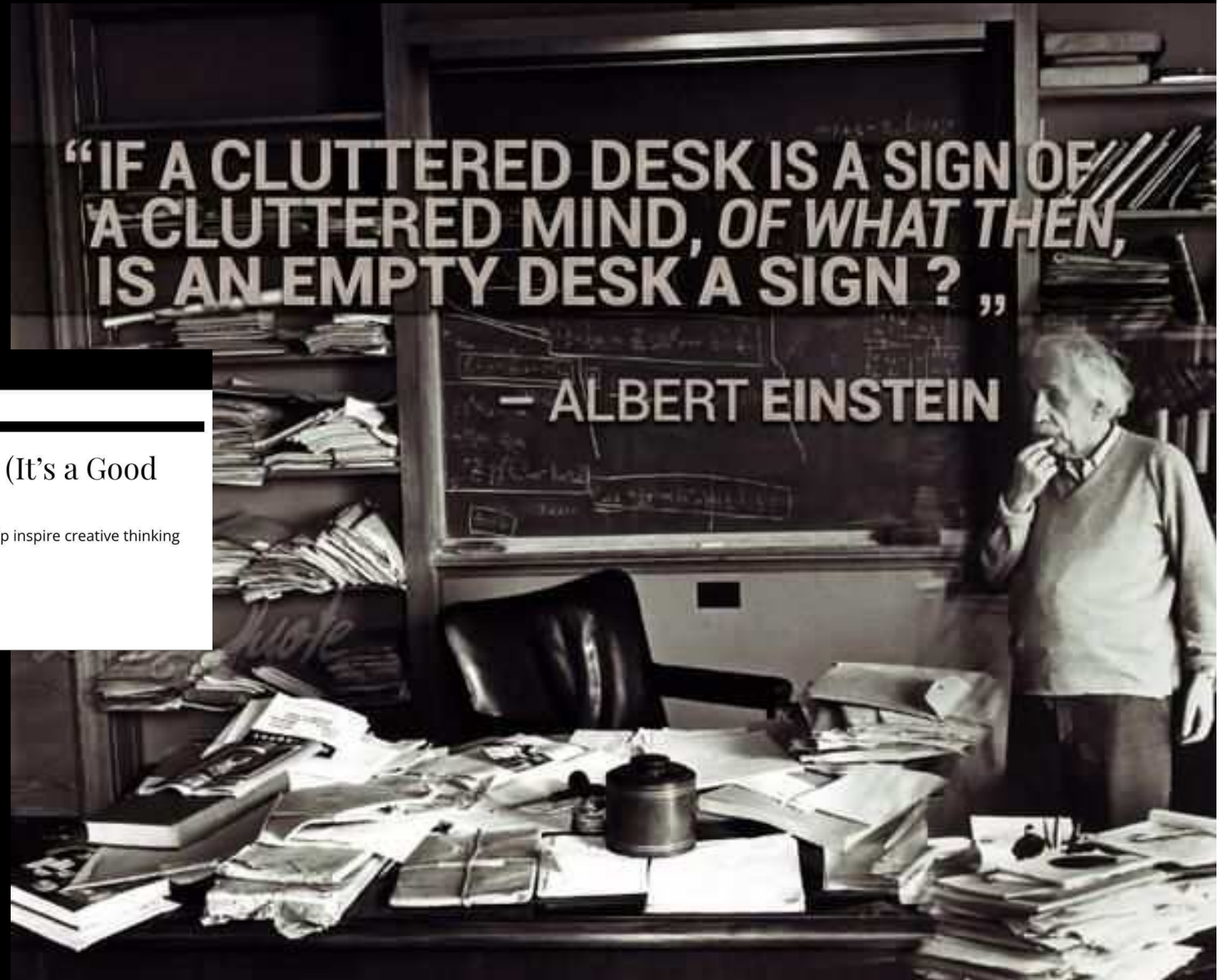
What Your Messy Desk Says About You (It's a Good Thing)

Recent research suggests that working in a sloppy setting may actually help inspire creative thinking



Randy Rieland

September 25, 2013



If Your Desk Is a Mess, You're Probably a Genius

 **Brandon Specktor**
Updated: Jun. 22, 2022

Smithsonian
MAGAZINE

INNOVATION

What Your Messy Desk Says About You (It's a Good Thing)

Recent research suggests that working in a sloppy setting may actually help inspire creative thinking



Randy Rieland

September 25, 2013

The Telegraph

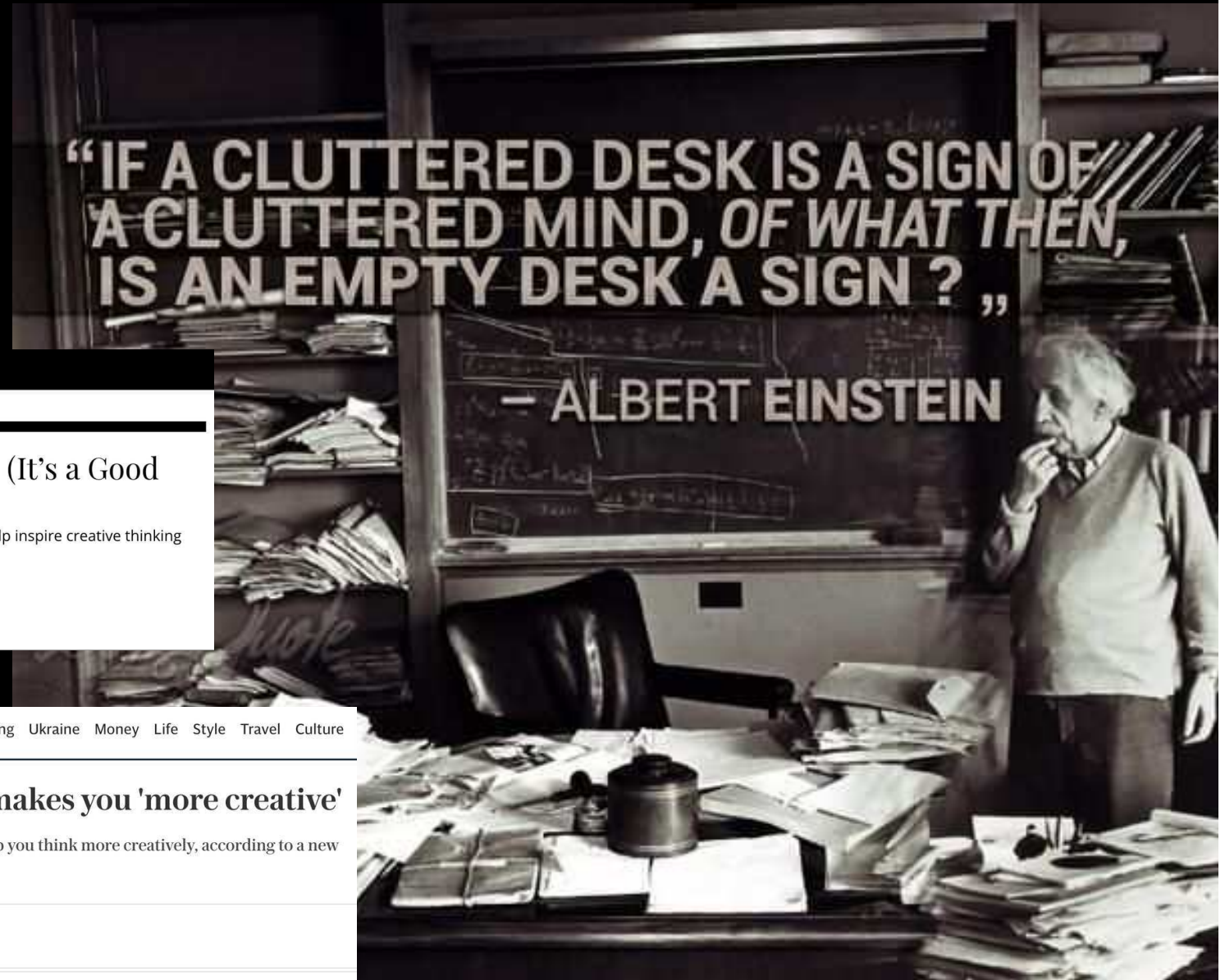
News Sport Business Opinion Cost of living Ukraine Money Life Style Travel Culture

Having a messy desk makes you 'more creative'

Working at a messy desk may actually help you think more creatively, according to a new scientific study.

By News Agencies

05 August 2013 - 11:22pm



If Your Desk Probably

Brandon Spivey
Updated: Jun. 22

- Einstein was probably misunderstood - operative term is 'empty' not 'disorganised'

- Geniuses are exceptions - organisation works for the majority so try it on for size

- You can't expect to become one by intentionally cluttering your workspace

INNOVATION

What Your Messy Desk Means (and What It Doesn't)

Recent research suggests that messy desks are not necessarily a sign of genius.



Randy Rieland

September 25, 2013

The Telegraph

scientific study.

By News Agencies

05 August 2013 - 11:22pm



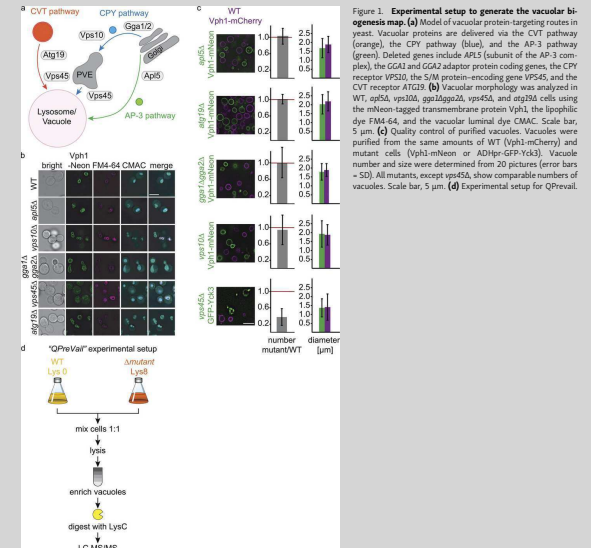
Understand the flow of information

Source

Organization
Analysis

Representation

Finished product



knockouts and highlights clusters of proteins that are potentially transported to the vacuole via different trafficking routes. Importantly, known model cargoes of each pathway were identified in the corresponding clusters. For example, a potential AP-3-dependent cluster (inset 1) contained the known AP-3 cargoes Pho8, Nyv1, and Yck3. Knockout of *ATG9* affected the known CVT cargoes *Apel* and *Apel4* (inset 4). CPY was affected in *vps30Δ* cells, in the GGA double deletion as well as *VPS45* knockout cells (inset 2). The vacuole biogenesis map also highlights proteins that are more abundant in mutant vacuoles, e.g., the cluster of plasma membrane proteins in *gga1gga2Δ* cells (inset 3). As expected, all identified subunits of the V1 domain of the V-ATPase were only affected in *vps45Δ* cells (inset 5). Several proteins only show very small changes in their vacuolar abundance in all analyzed mutants (inset 6 and Table S1). These were mainly peripheral membrane proteins such as subunits of the homotypic fusion and protein sorting (HOPS) complex (Rieder and Emr, 1997) and the target of rapamycin (TOR) complex 1. Besides minimal changes in the SILAC ratios of these proteins, subunits of each complex formed distinct clusters. Together, this suggests that for HOPS and TORC1 all subunits are affected similarly, even though changes in the ratios are only minimal.

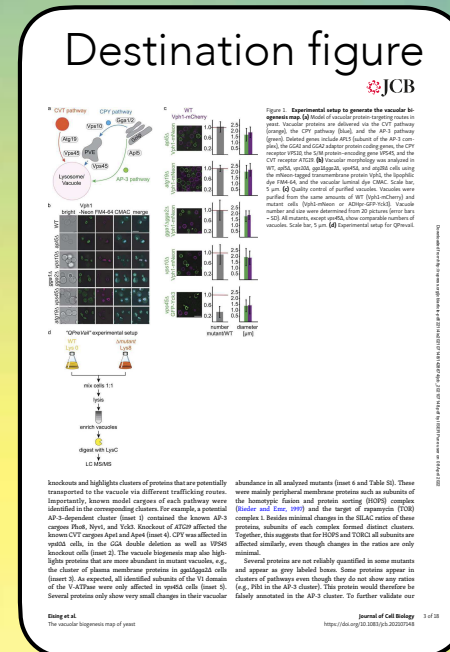
Several proteins are not reliably quantified in some mutants and appear as grey labeled boxes. Some proteins appear in clusters of pathways even though they do not show any ratios (e.g., *Pib1* in the AP-3 cluster). This protein would therefore be falsely annotated in the AP-3 cluster. To further validate our

Data gathering may
give you a
temporary sense of
purpose.....

.... but if done
aimlessly, it is quite
futile



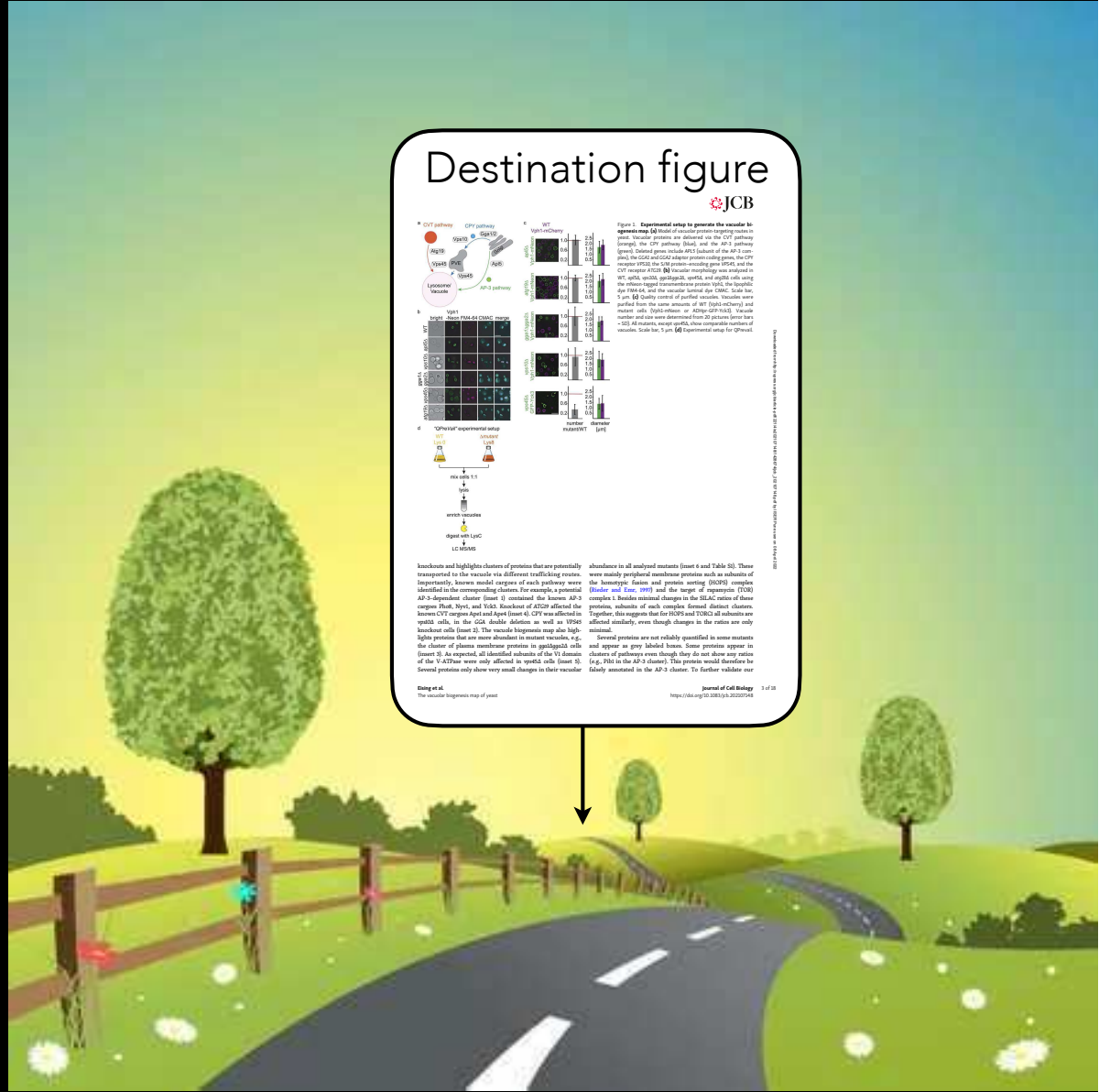
Fortunately, in science, you can imagine where you need to go



Imagine/picture how
your data would look
in a figure of a
manuscript

Follow an intuitive logic

Work your way backwards



Source



- Diverse, proprietary and somewhat ugly formats
- Not catered for your 'finished product'
- Data analysis typically done via a secondary interface. For e.g., Excel, Graphpad Prism, ImageJ
- Go through the process of 'reformatting' your data
- Why not get to doing this from the start

Reformatting source data

Source

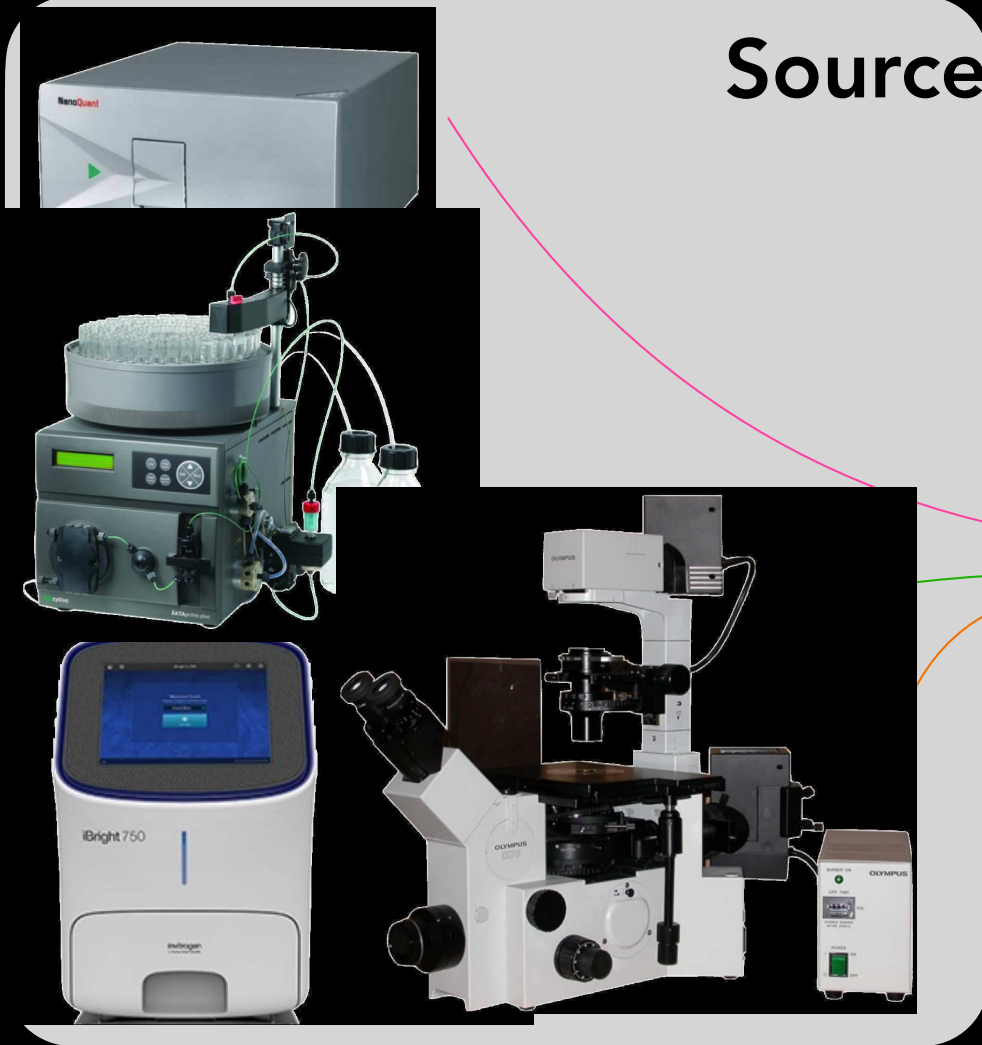
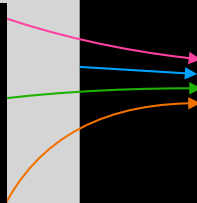
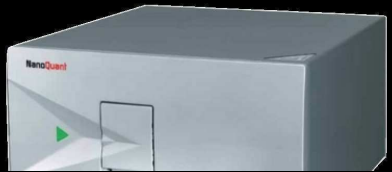


Plate reader data can be converted from 96-well format into an XY table format

Table format: XY		X	Group A	Group B
		Time (min)	OD	Title
	⊗	X	Y	Y
1	Title	0	0.23	
2	Title	10	0.34	
3	Title	20	0.44	
4	Title	30	0.56	
5	Title			

Source

FPLC data can be
exported and replotted in
Graphpad Prism



Images can be saved only
in highest resolution

Source

Reformatted with easily
recognisable and
reliable identifiers

CBB.tif, Fluor.tif

Name ↓



stream flowing in prtoein_MMStack_Pos0.ome.tif



stream flowing in prtoein_MMStack_Pos0_metadata.txt

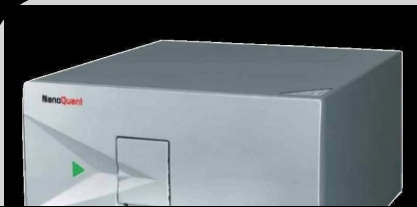


Movie1.tif, image1.tif, still1.tif

Advantages

- Reformatting is tedious but necessary
- First step in organising
- Instils a habit to revisit your data
- You will invariably emerge with a better understanding of data
- Gets you to appreciate the graphing process
- Helps clarify the basis of linear and non-linear regression
- Helps understand data fitting (which fit works best?)
- Helps understand what significant numbers are?
- Helps understand which statistical test to use?
- Helps with data representation
- The more you do this, the better you get at experiment design

Source

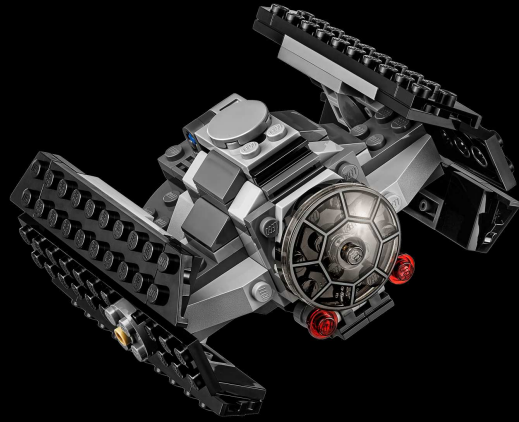


Go to google folder

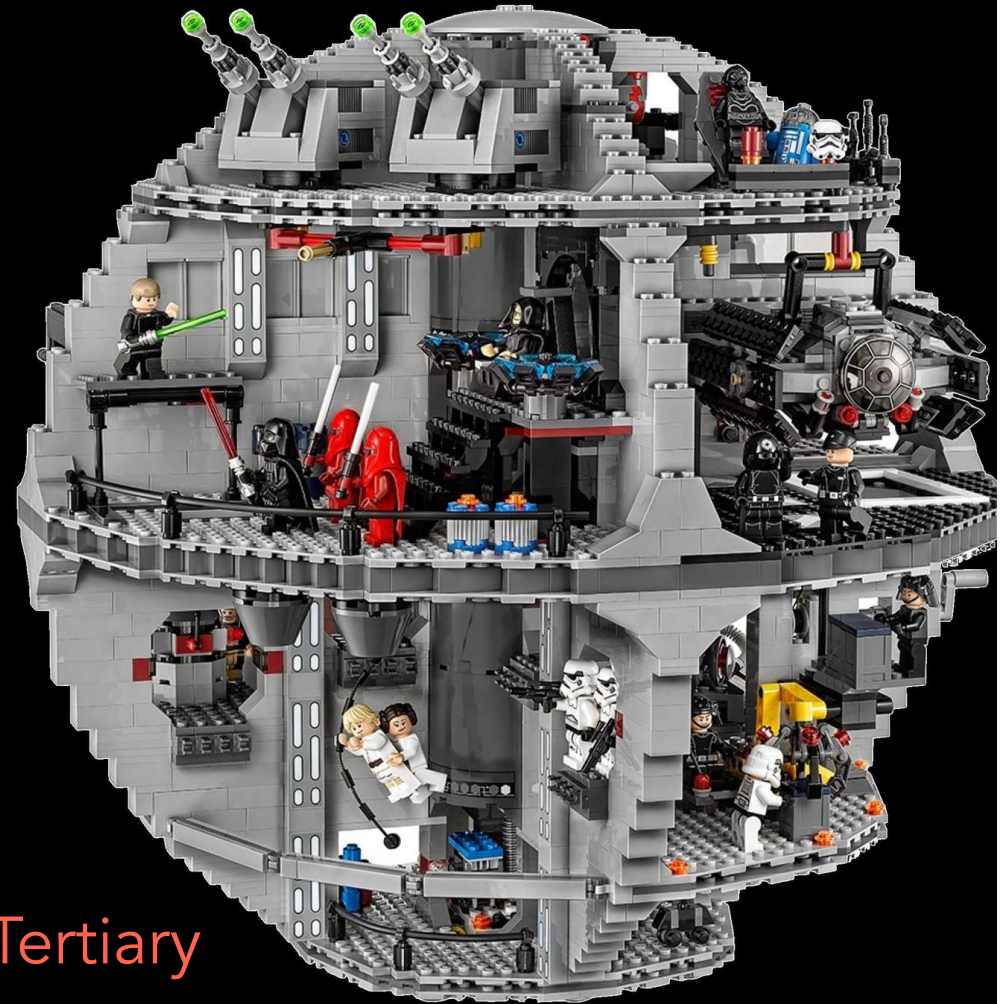
Organised data has multiple tiers



Primary



Secondary



Tertiary

Organised data has multiple tiers

Organised data has multiple tiers

Primary

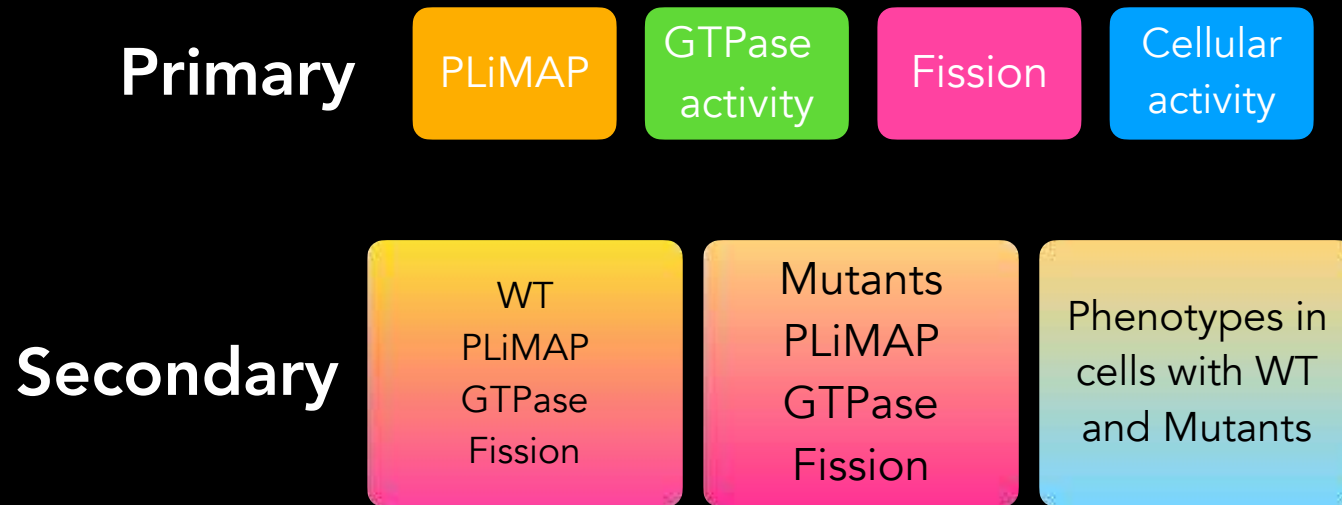
PLiMAP

GTPase
activity

Fission

Cellular
activity

Organised data has multiple tiers



Organised data has multiple tiers

Primary

PLiMAP

GTPase
activity

Fission

Cellular
activity

Start preparing figures

Secondary

WT
PLiMAP
GTPase
Fission

Mutants
PLiMAP
GTPase
Fission

Phenotypes in
cells with WT
and Mutants

Organised data has multiple tiers

Primary

PLiMAP

GTPase
activity

Fission

Cellular
activity

Start preparing figures

Secondary

WT
PLiMAP
GTPase
Fission

Mutants
PLiMAP
GTPase
Fission

Phenotypes in
cells with WT
and Mutants

Tertiary

Report
activity

Establish
and validate
with mutants

Focus
on mechanism

Expand scope