
Story Points
Versus
#NoEstimates

{

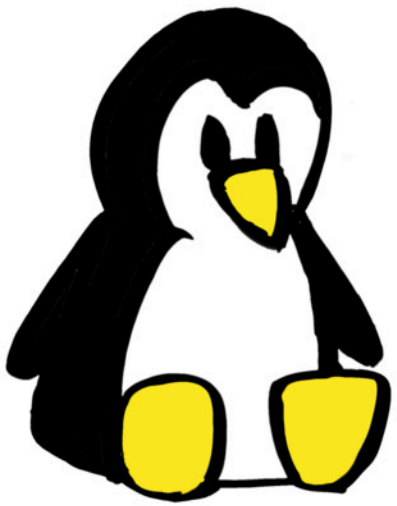
NAME: VICTOR LAMBRET

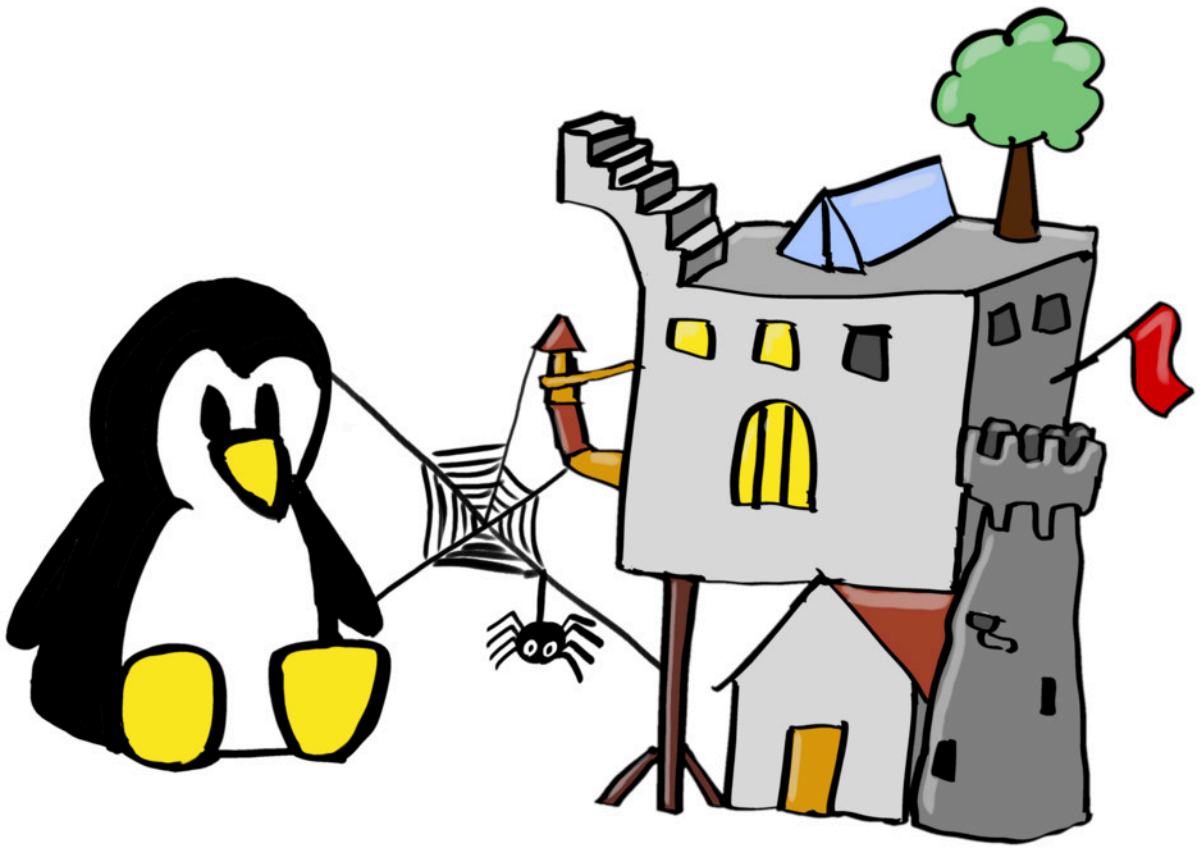
AGE: 0X28

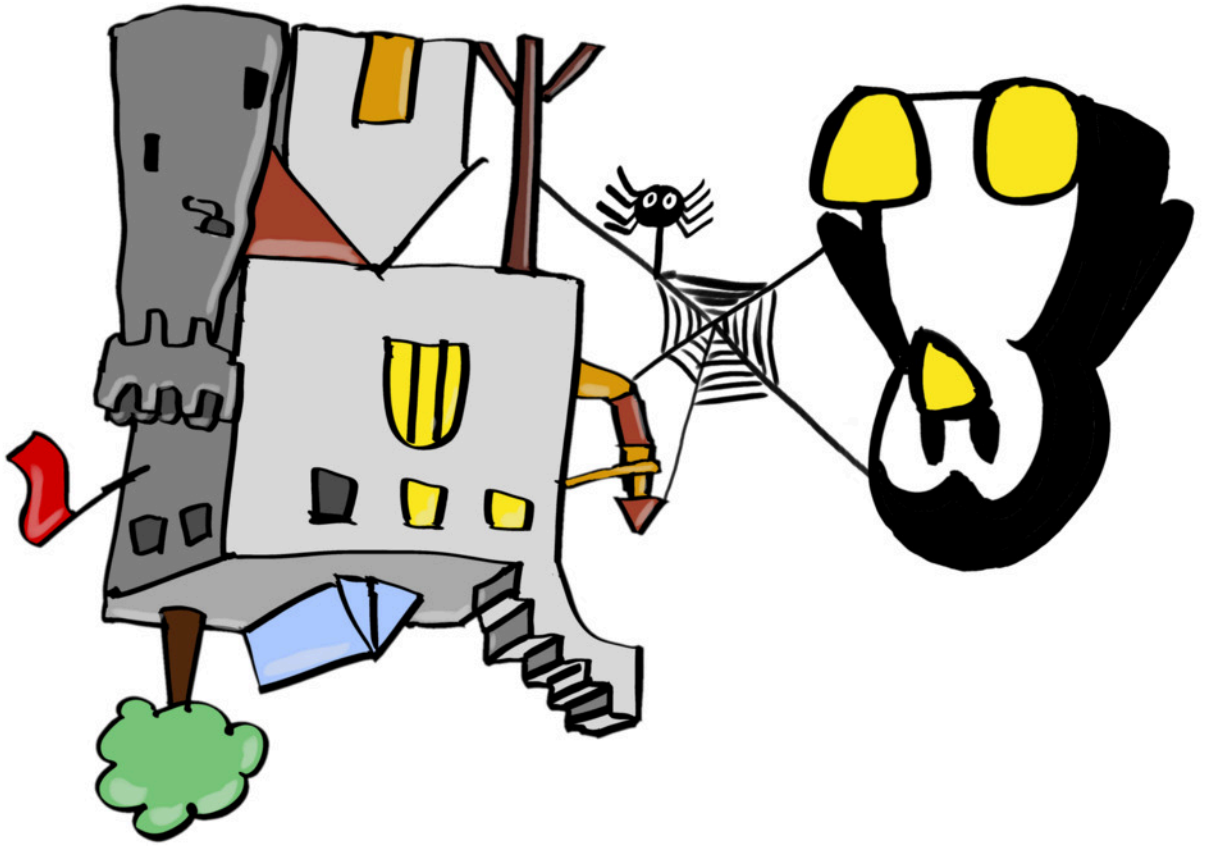
JOB: BACKEND DEVELOPER

COMPANY: AGICAP

}





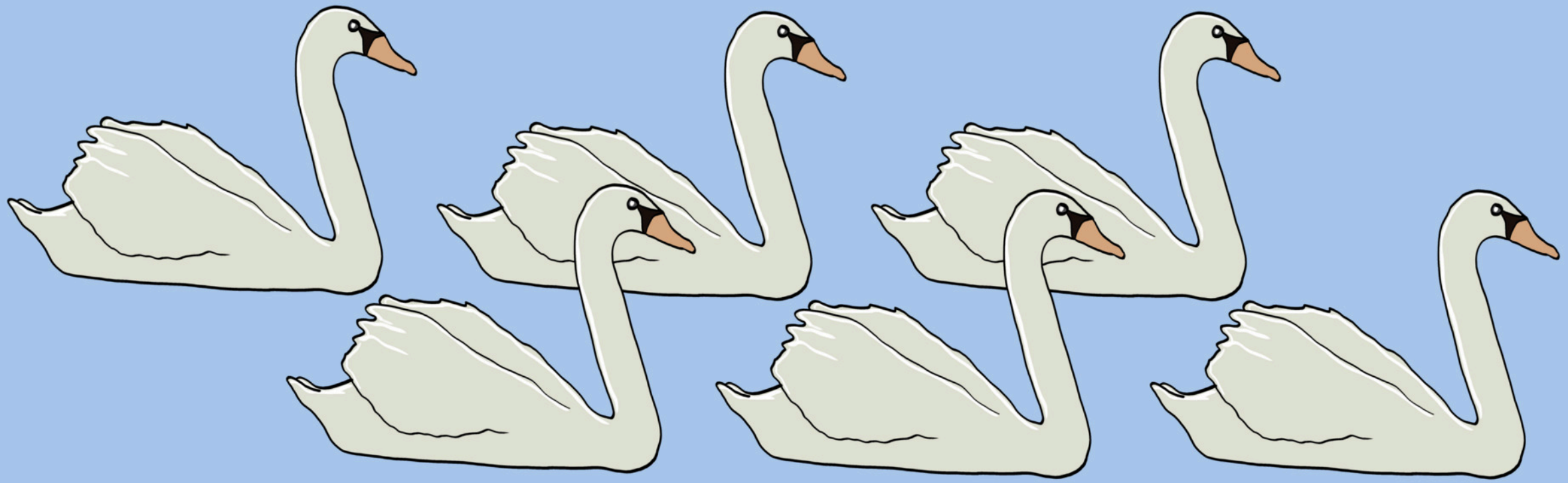


WHY YOUR IT PROJECT MAY BE RISKIER THAN YOU THINK

Flyvbjerg, Bent and Budzier, Alexander

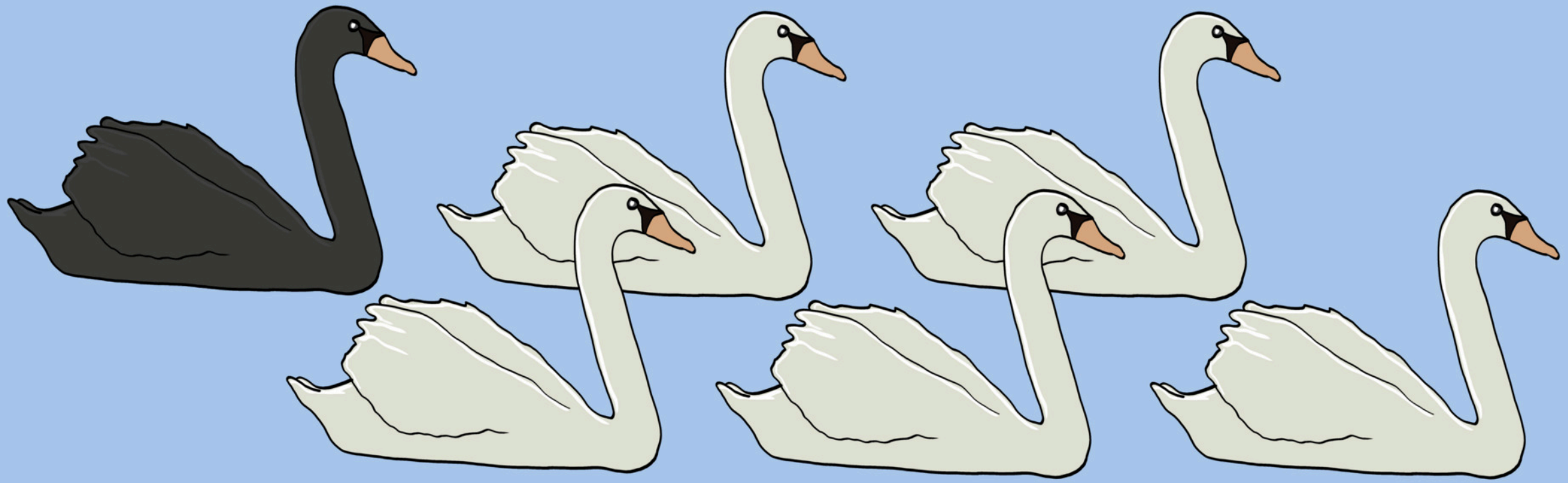
2011 - Harvard business review

AVERAGE +27%

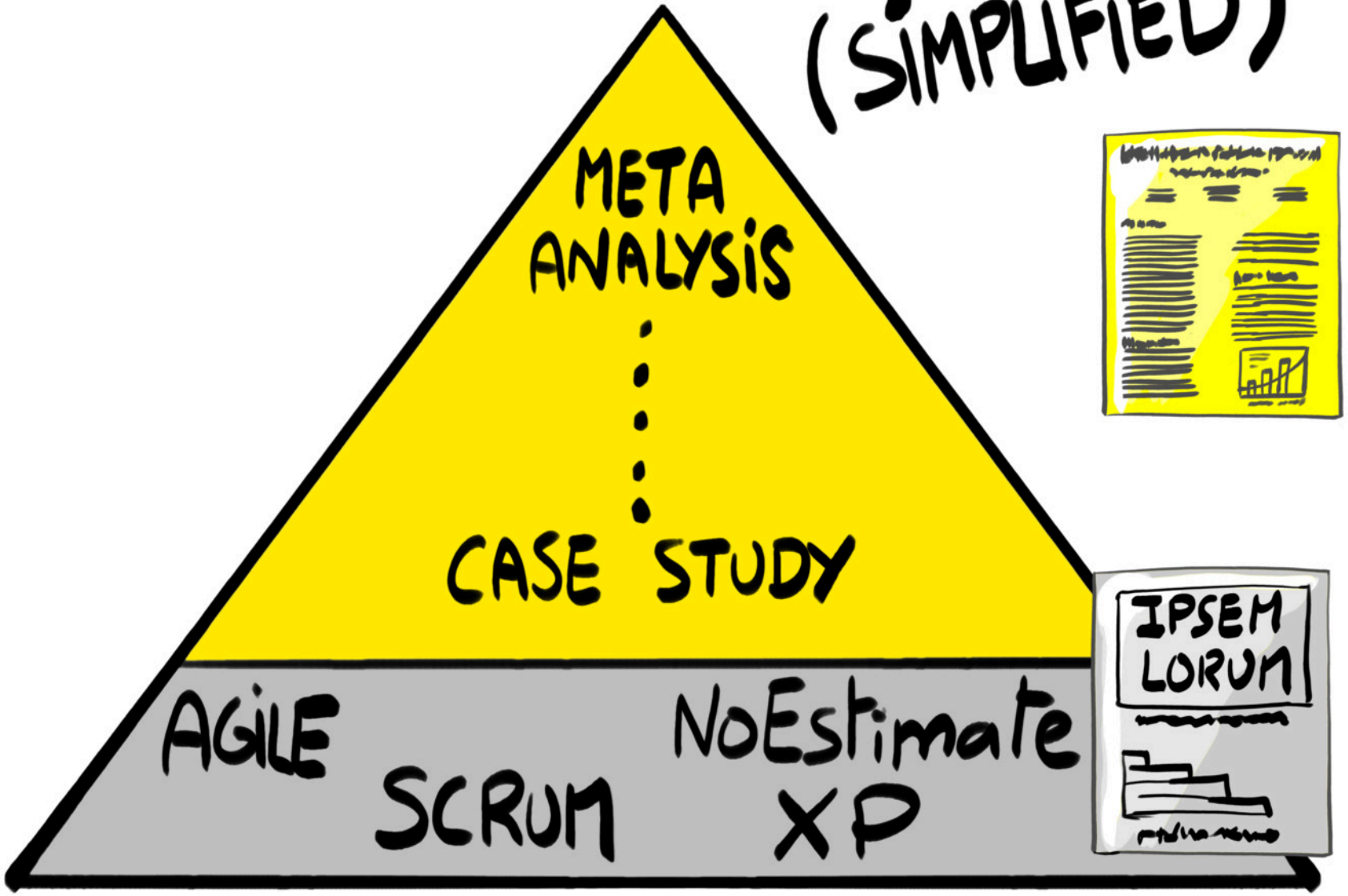


AVERAGE +200%

+70%



LEVELS OF EVIDENCE (SIMPLIFIED)



ACADEMIC RESEARCH

(PEER-REVIEWED)

NON-ACADEMIC MATERIAL

(WIDER RANGE OF TOPICS)

I CONTEXT

II

III

IV

V

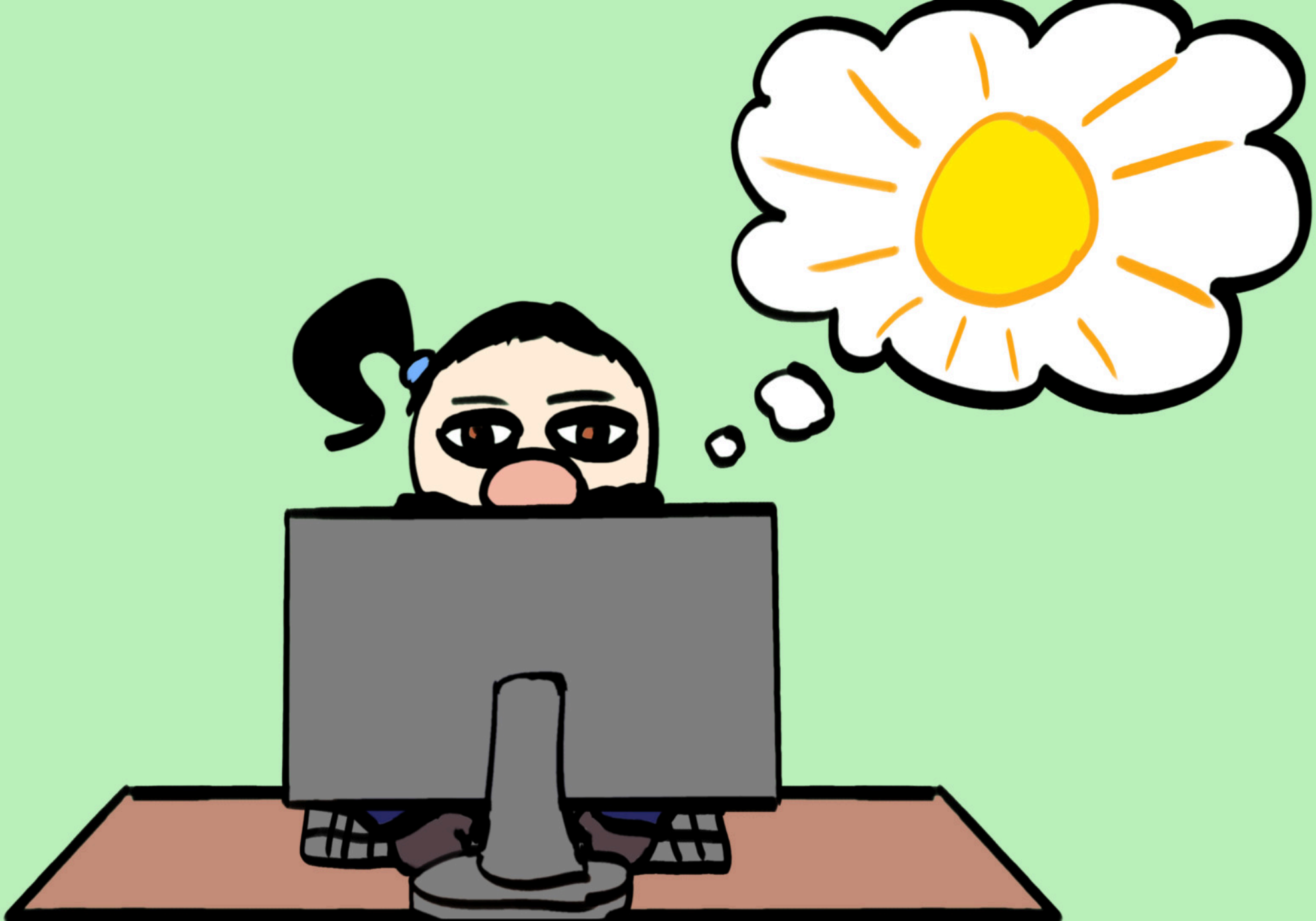
VI

VII


STORY
POINTS

EXTREME PROGRAMMING

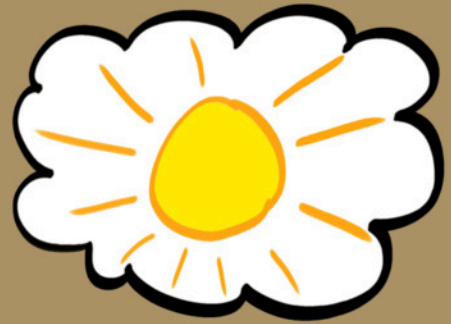
IDEAL ENGINEERING DAY




SORRY, BUT THE PRODUCT HAS CRASHED AGAIN...
IS IT DONE?
Would you like help ?
Yes
No

Load Factor

=

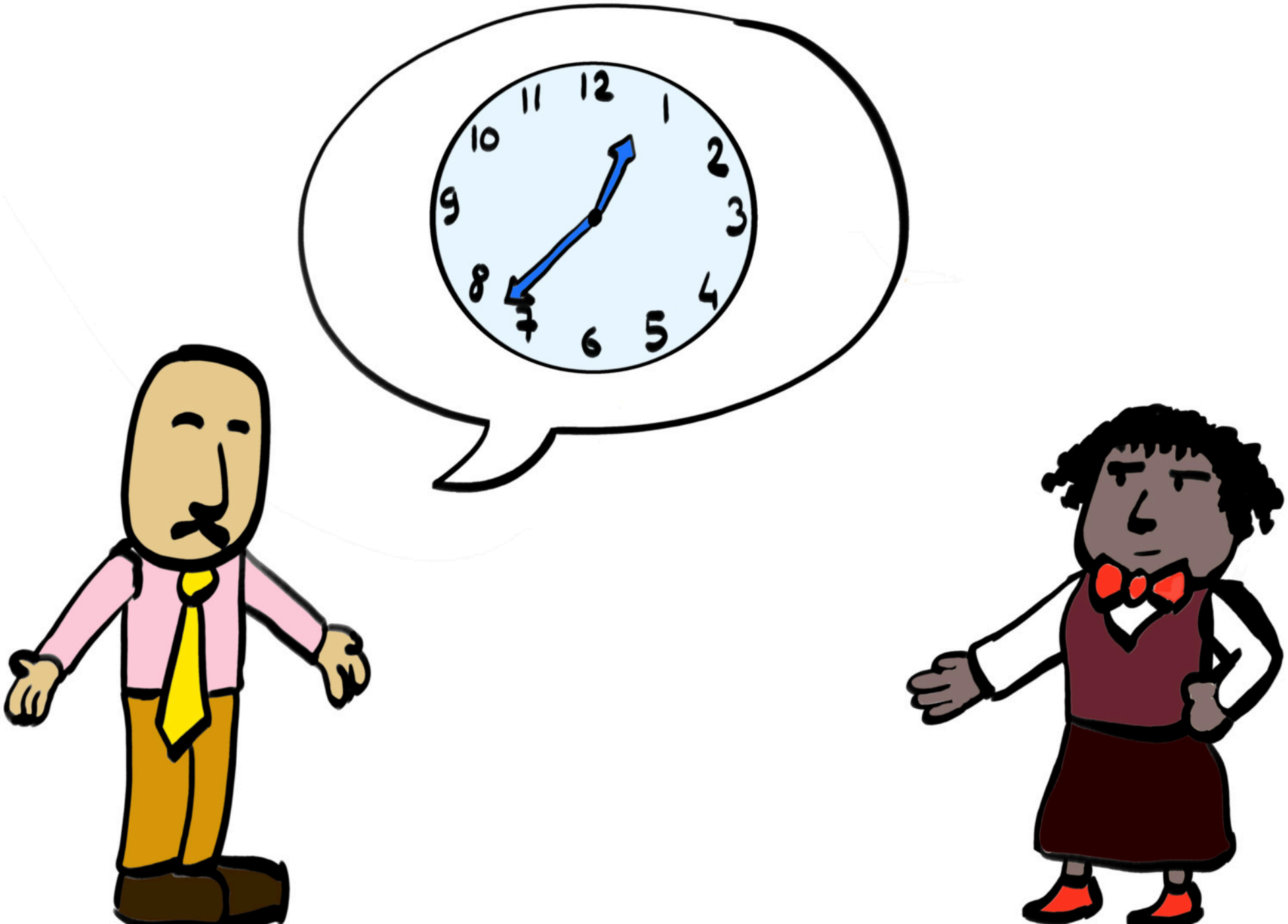


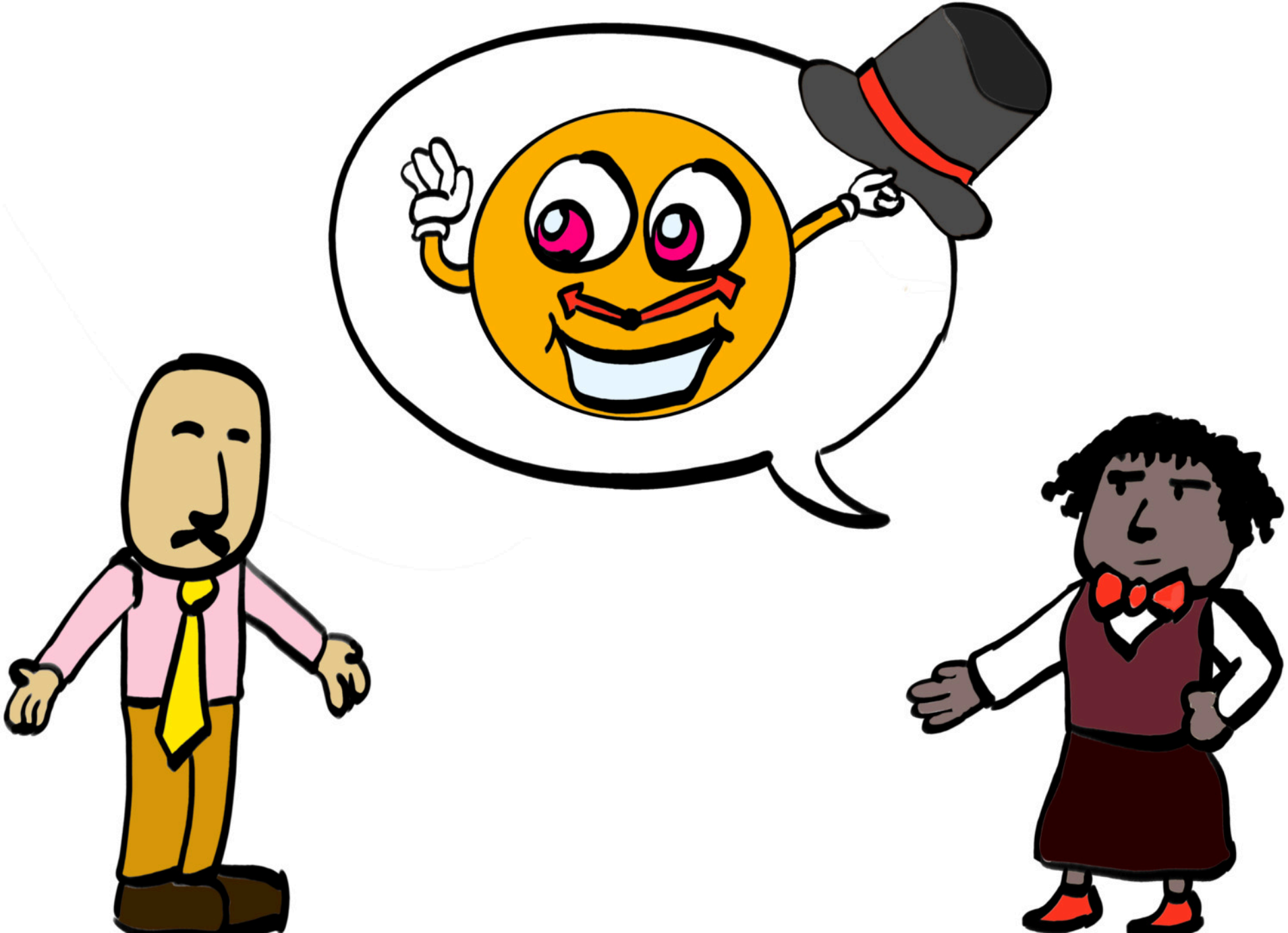
IDEAL ENGINEERING DAY

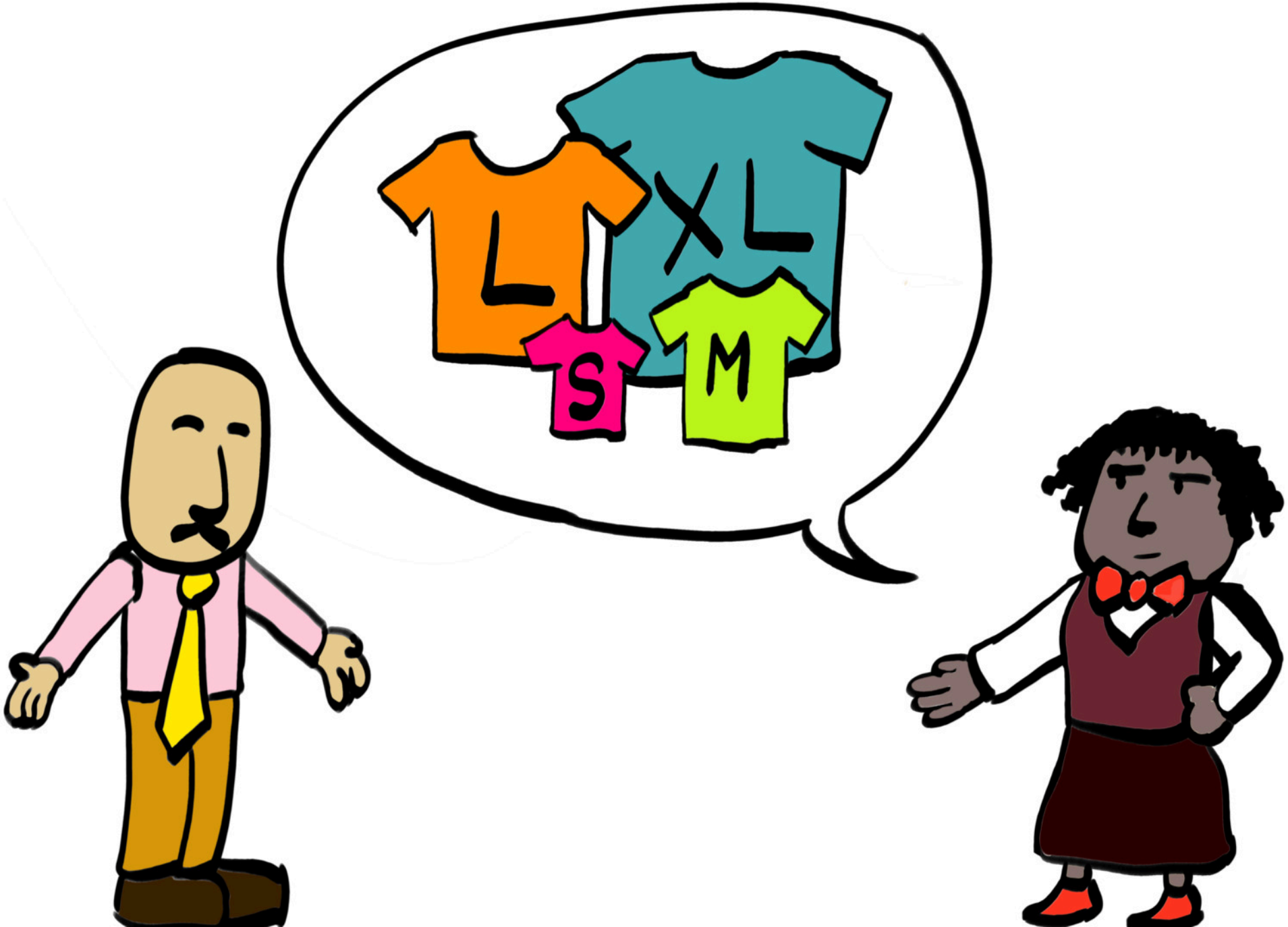
IDEAL
IGINEERING
DAY

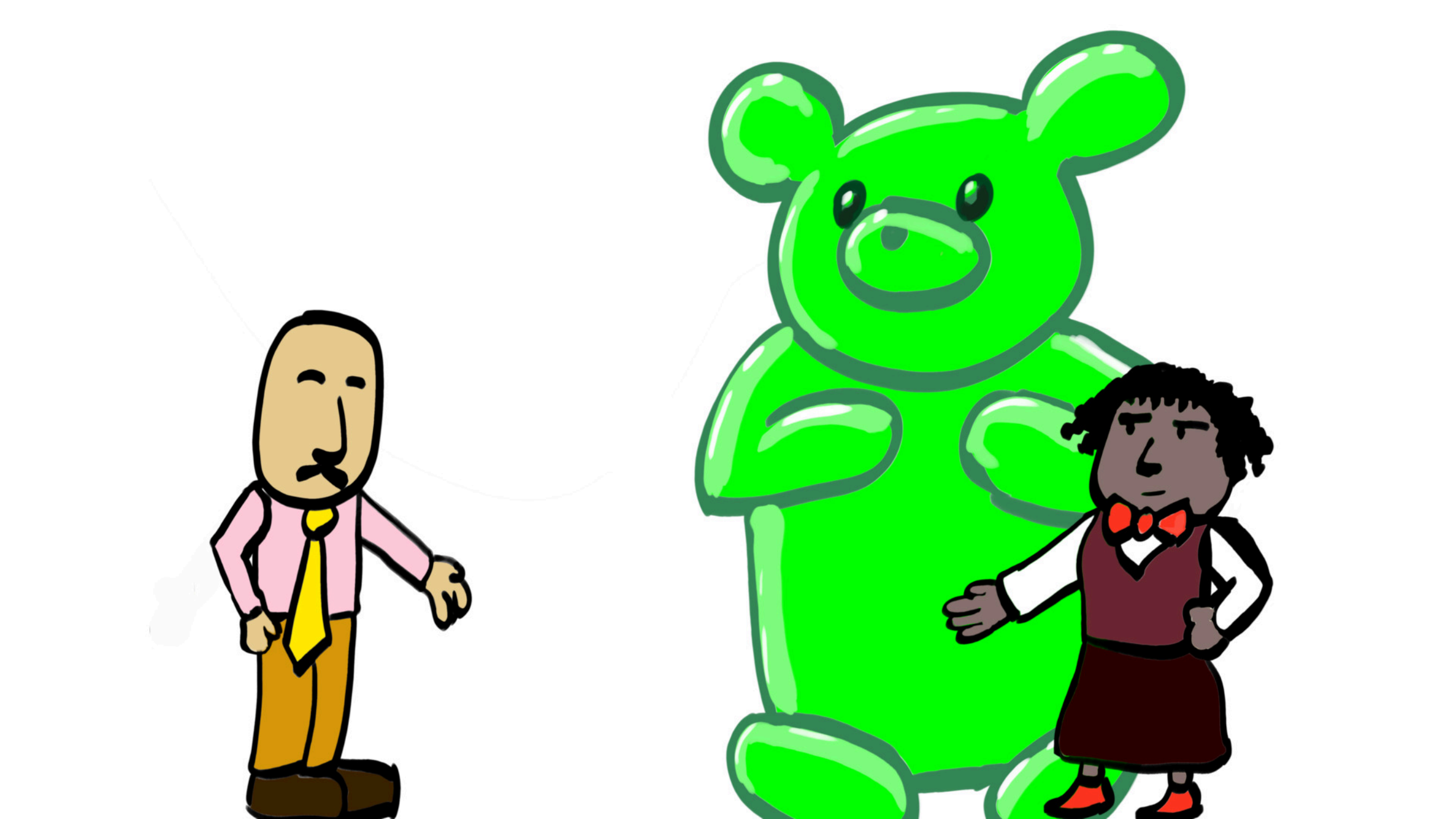
IDEAL ENGINEERING

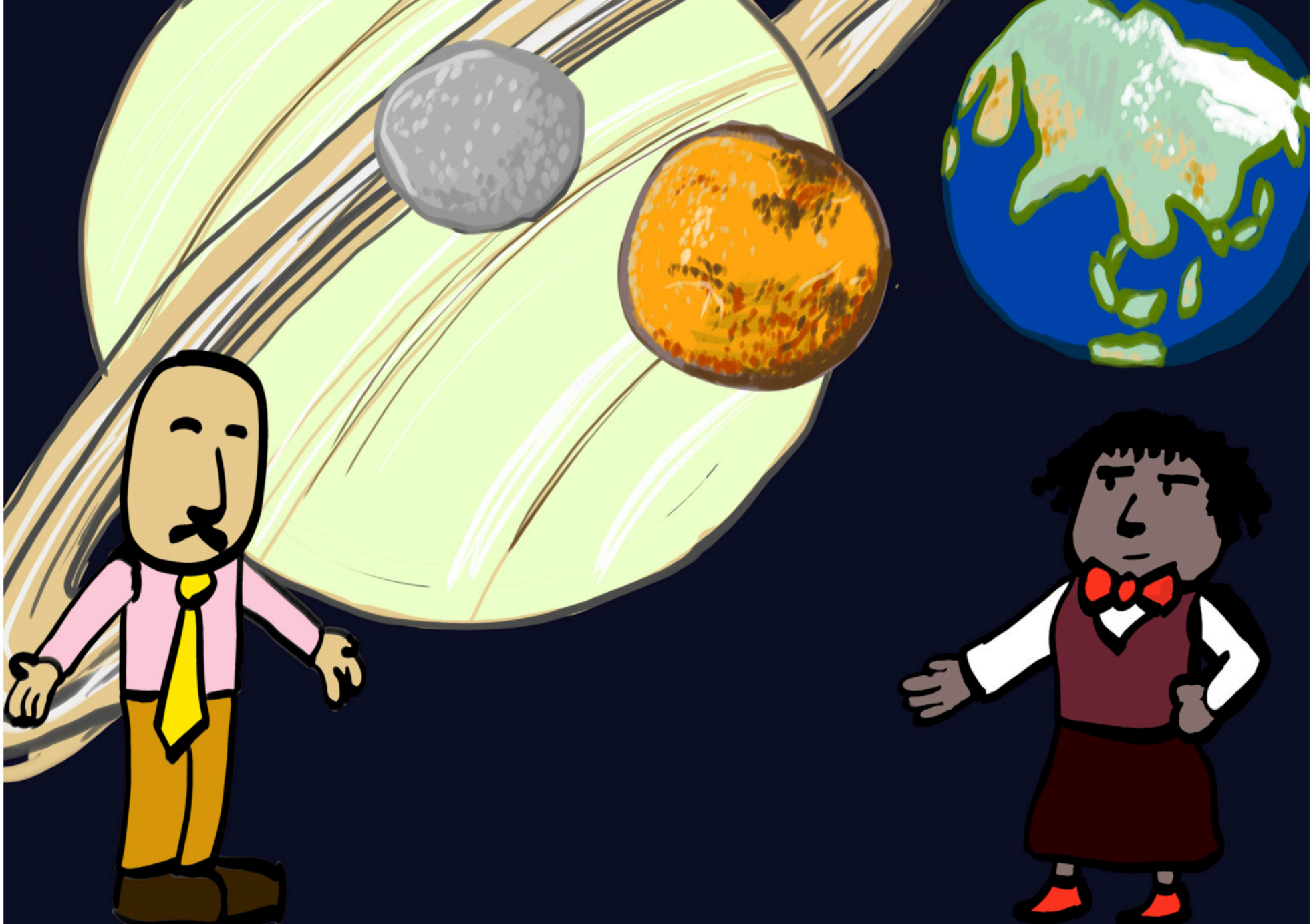
DAYDLINES

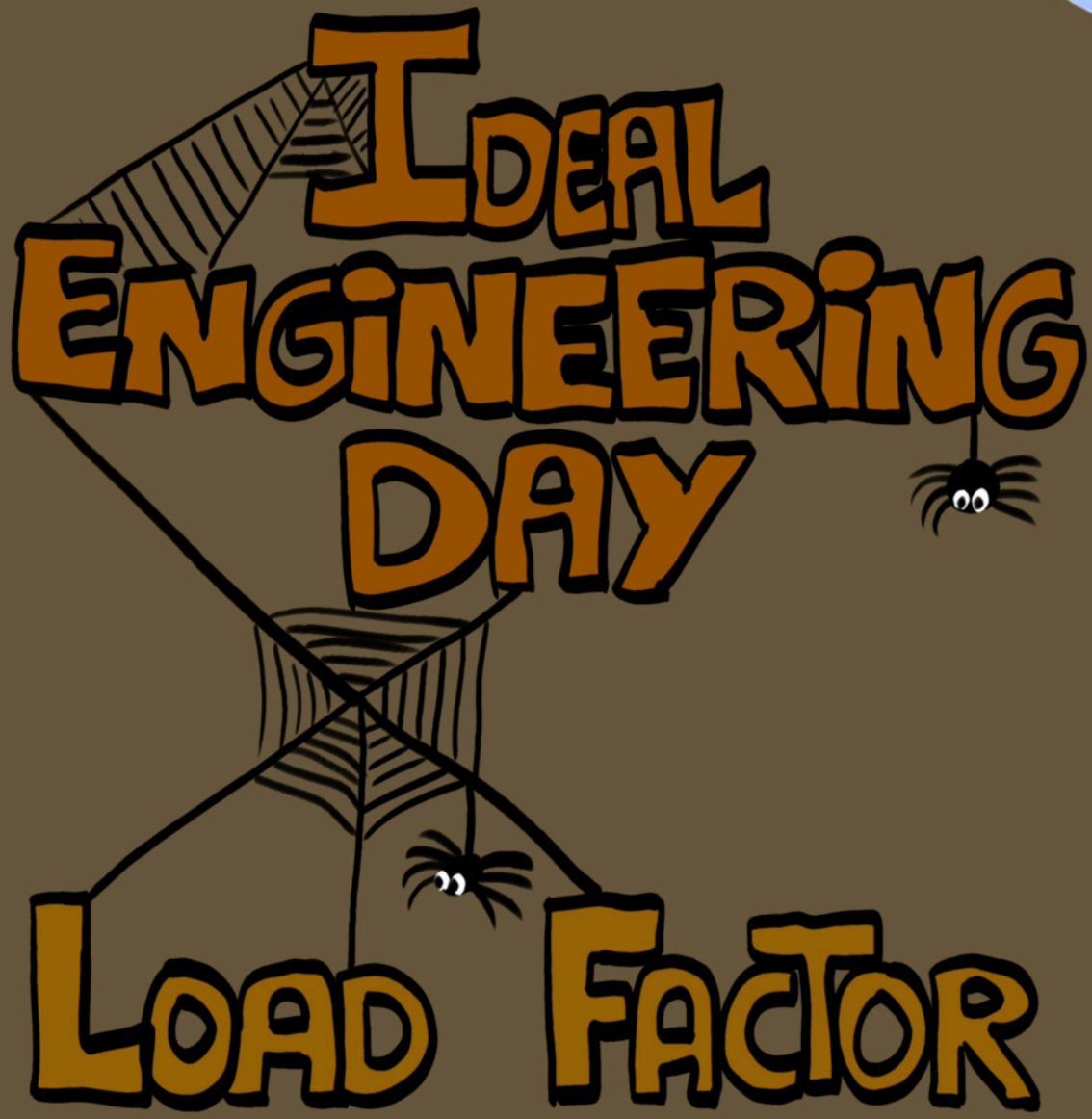

12
1
2
3
4
5
6
7
8
9
10
11




L
XL
S
M







Story
Points

Velocity

EFFORT ESTIMATION IN AGILE SOFTWARE DEVELOPMENT: A SURVEY ON THE STATE OF THE PRACTICE

Usman, Muhammad and Mendes, Emilia and Börstler, Jürgen

2015

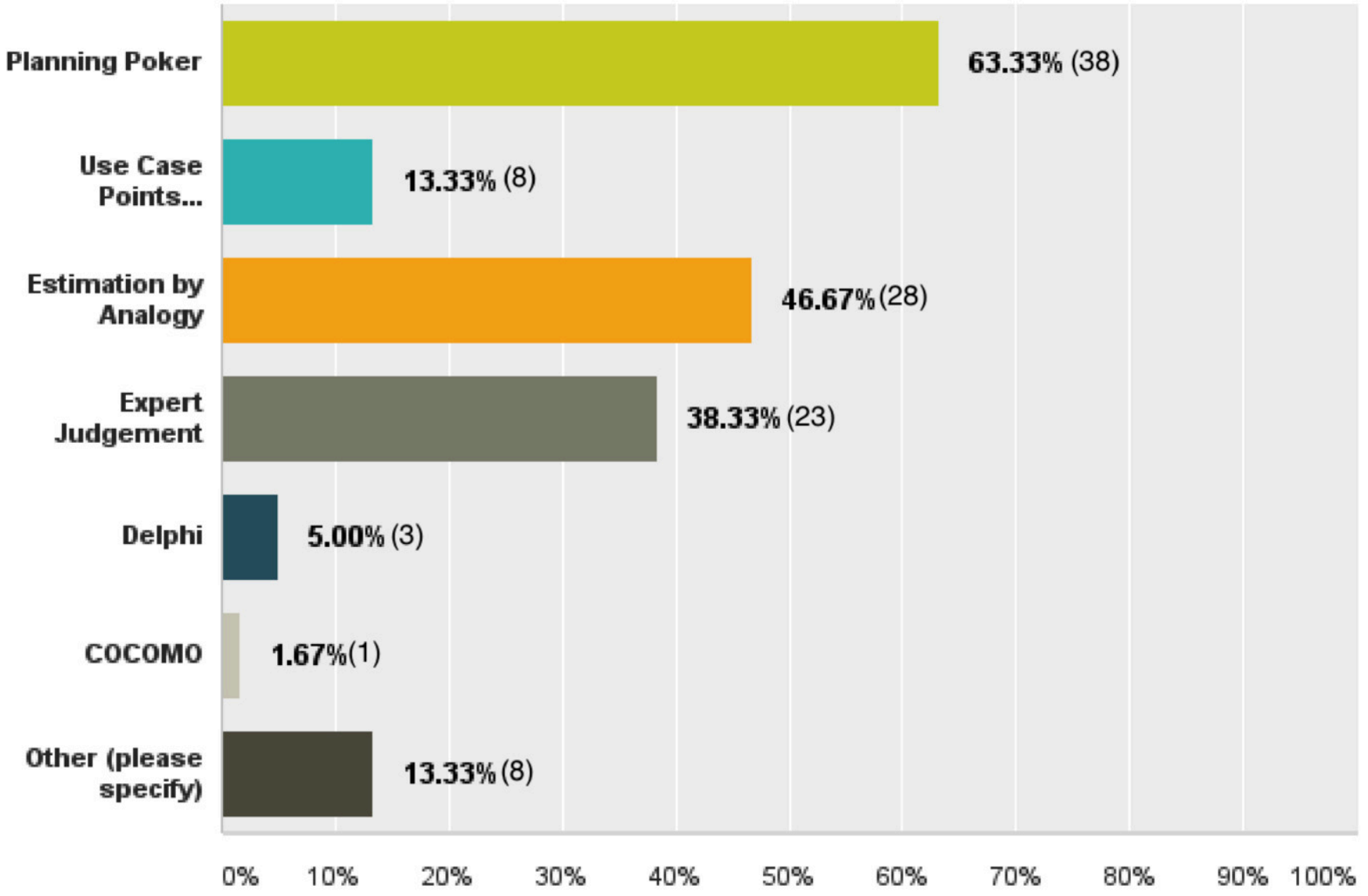


Figure 3: Effort Estimation Techniques

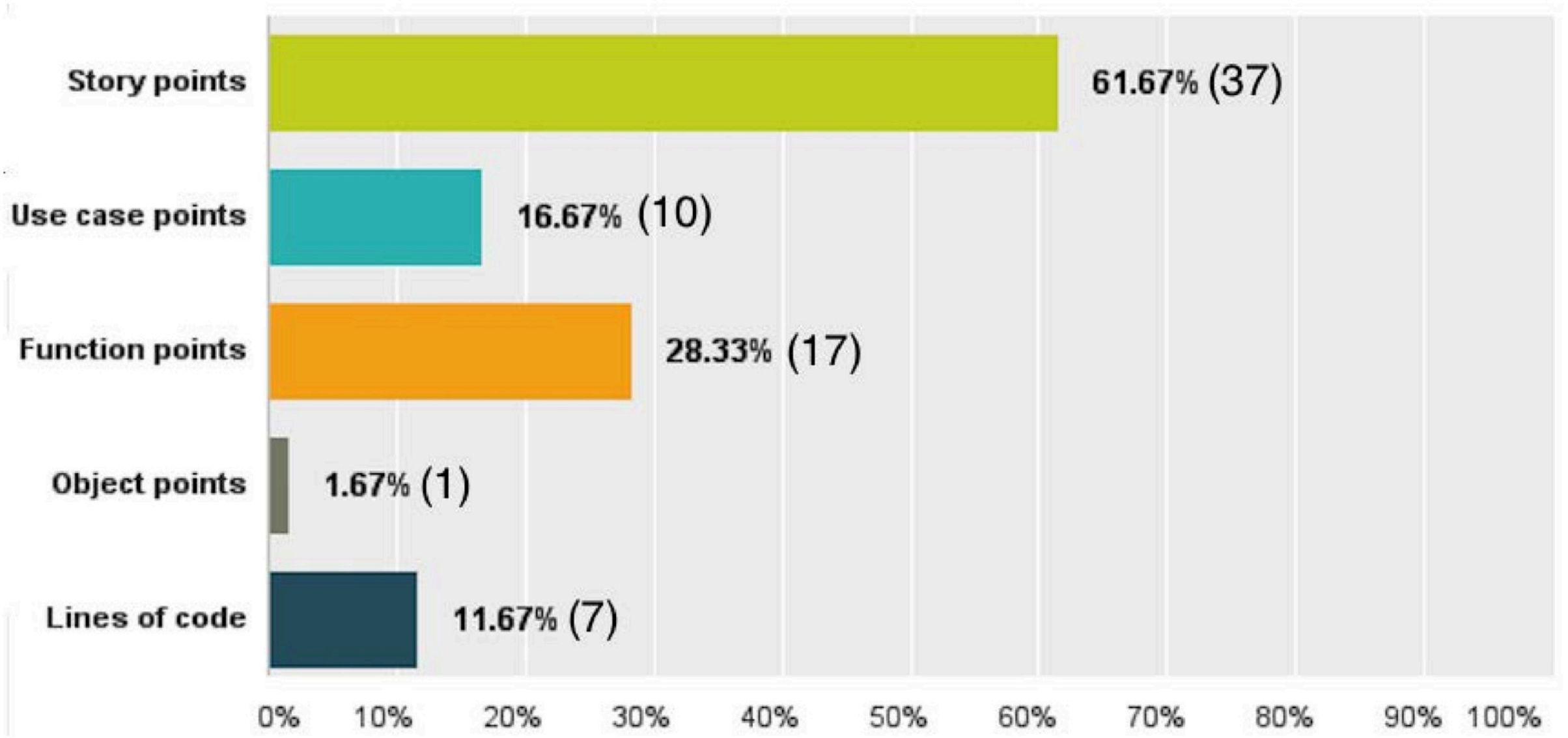
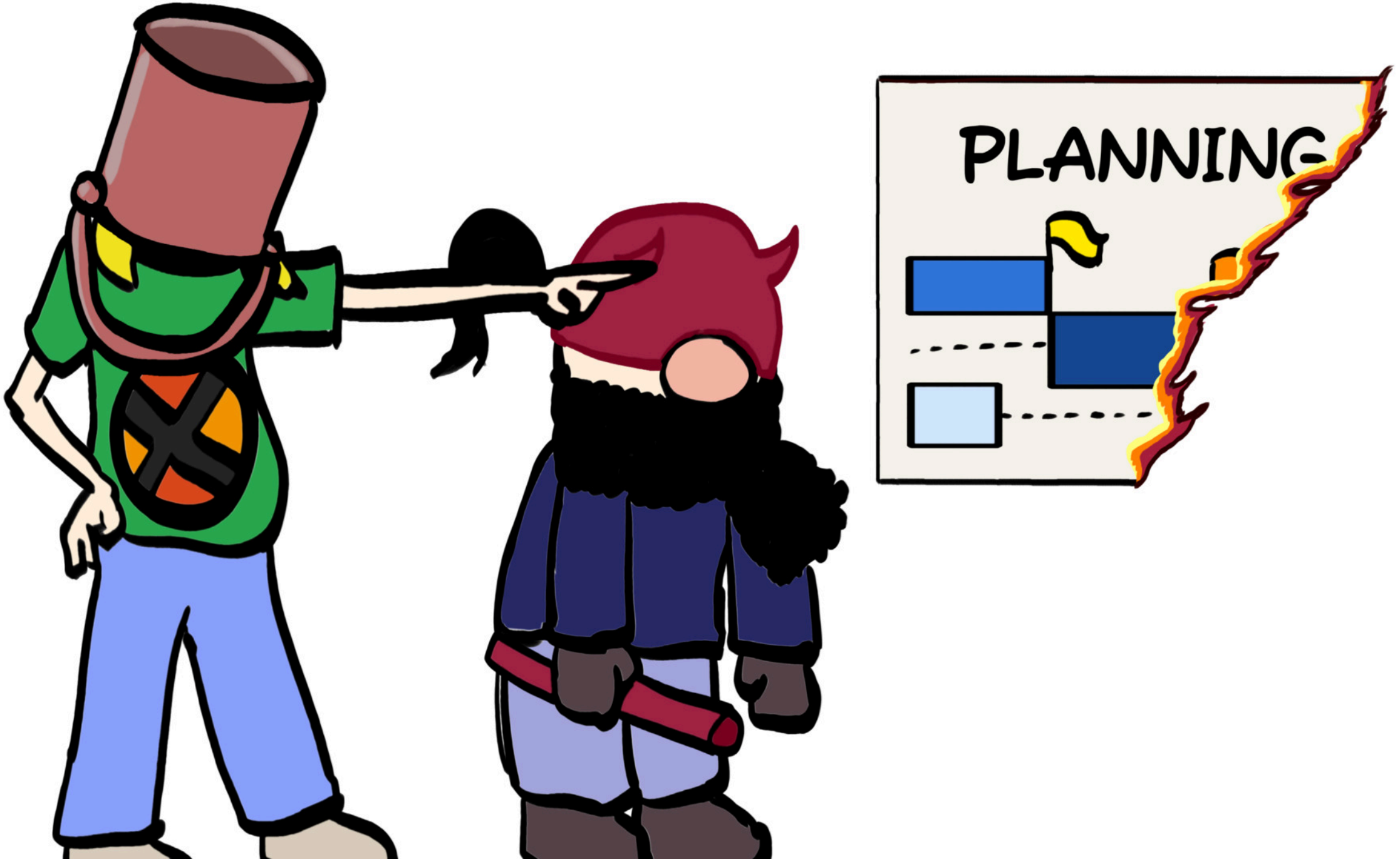
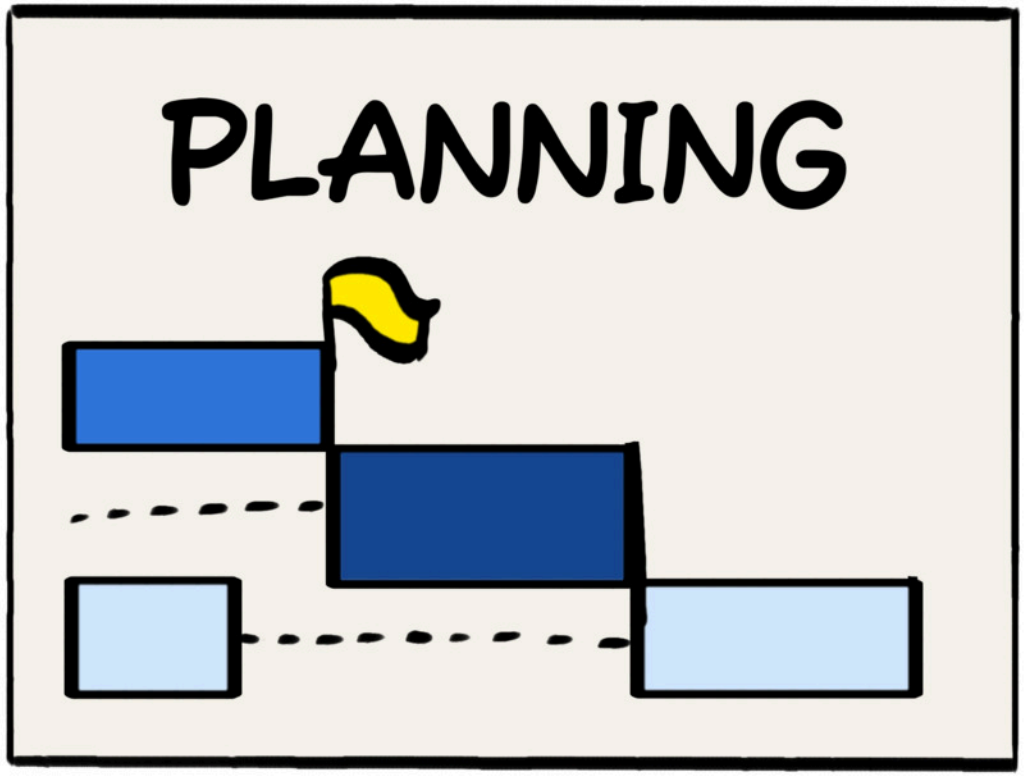


Figure 4: Size Metrics

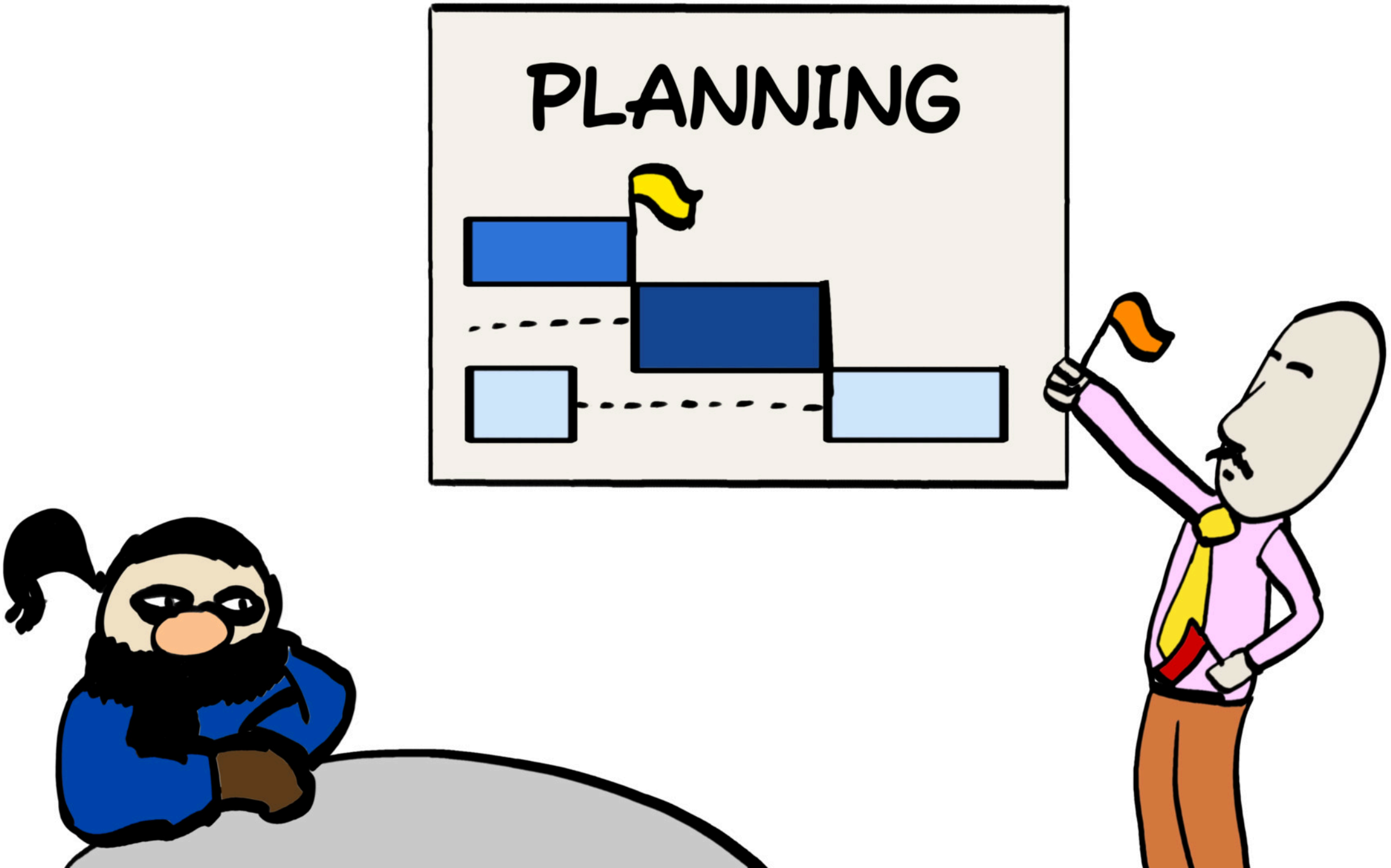

#NOEstimates


NO ESTIMATES

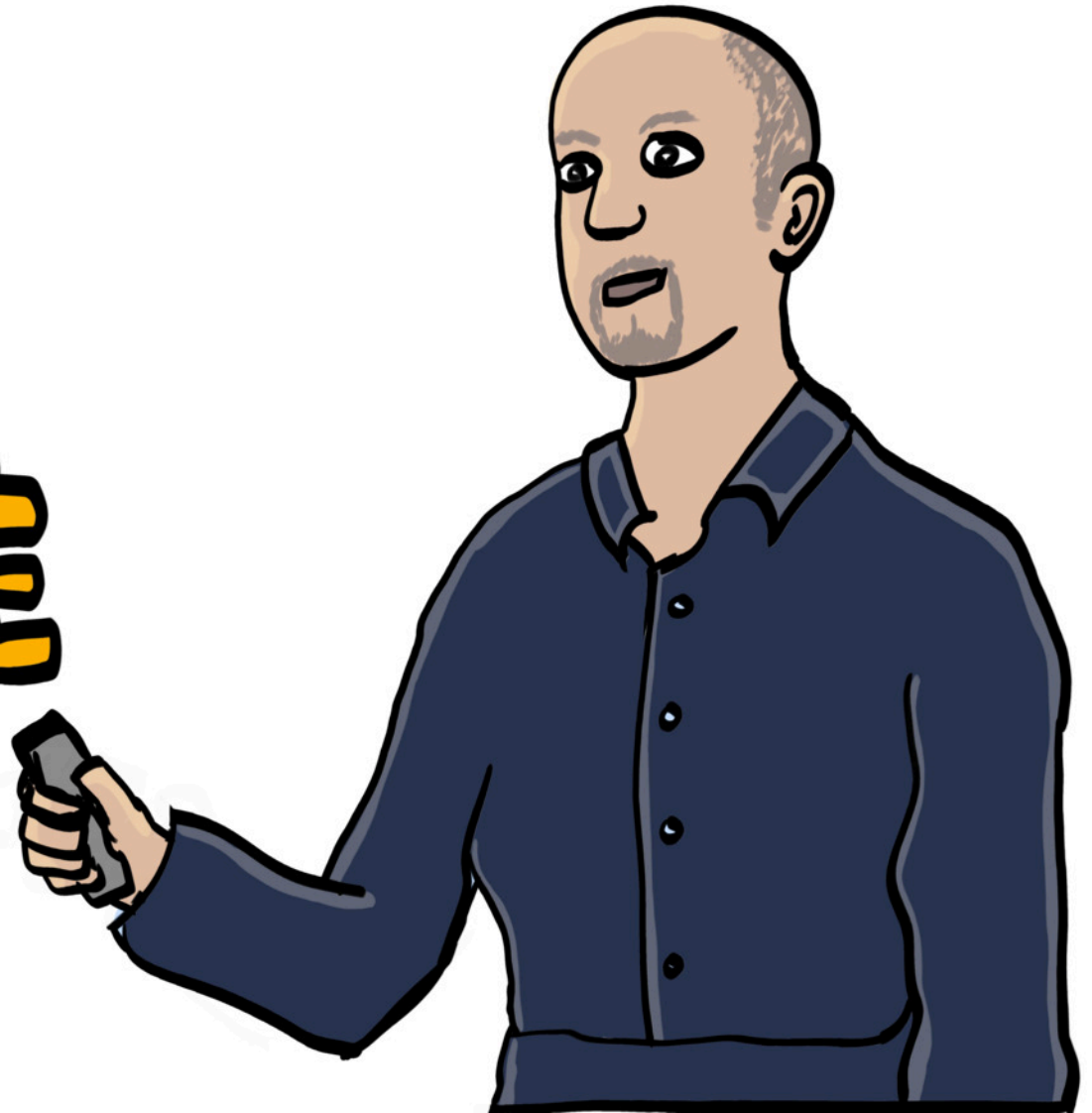

PLANNING


PLANNING

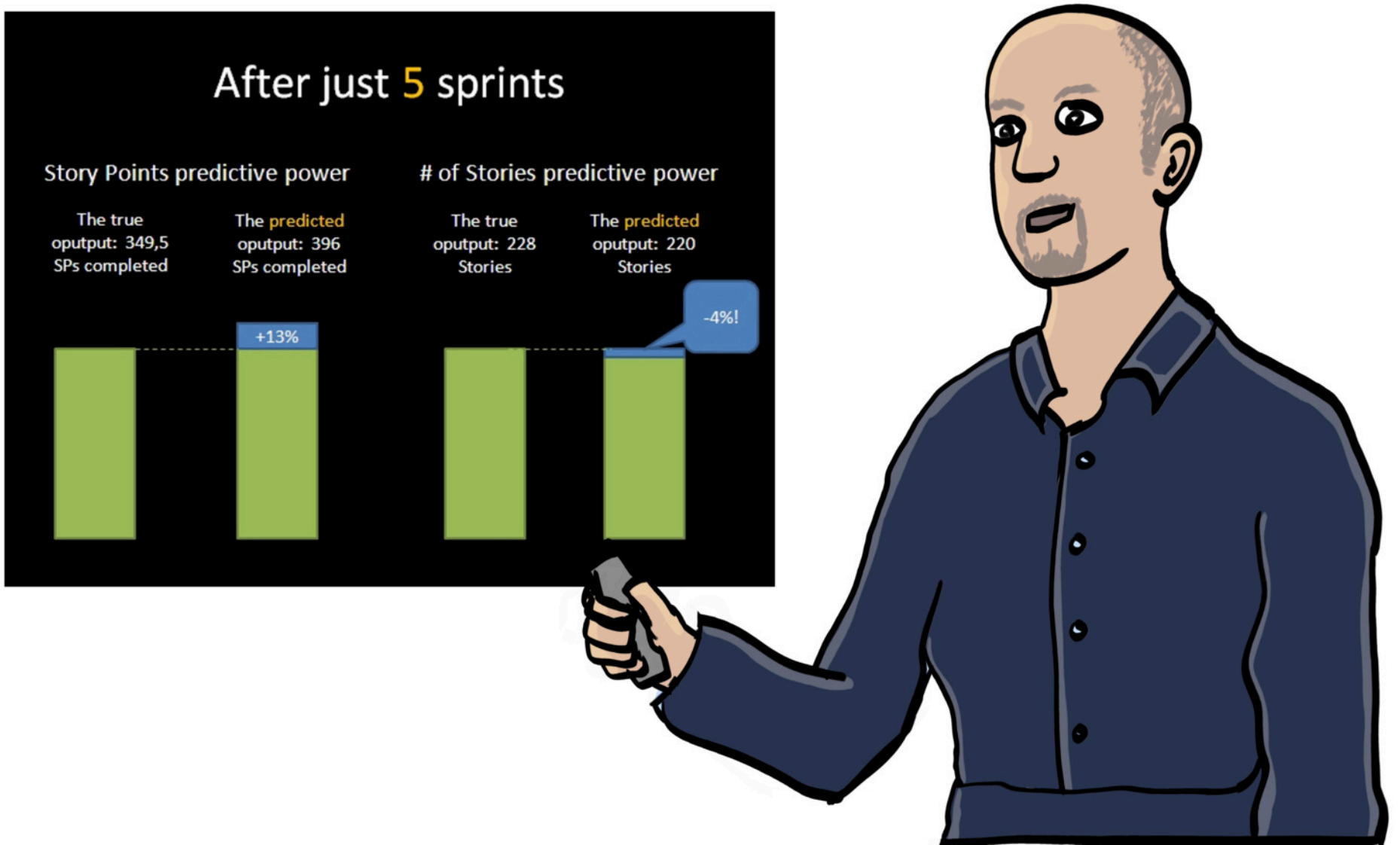

NNING
3

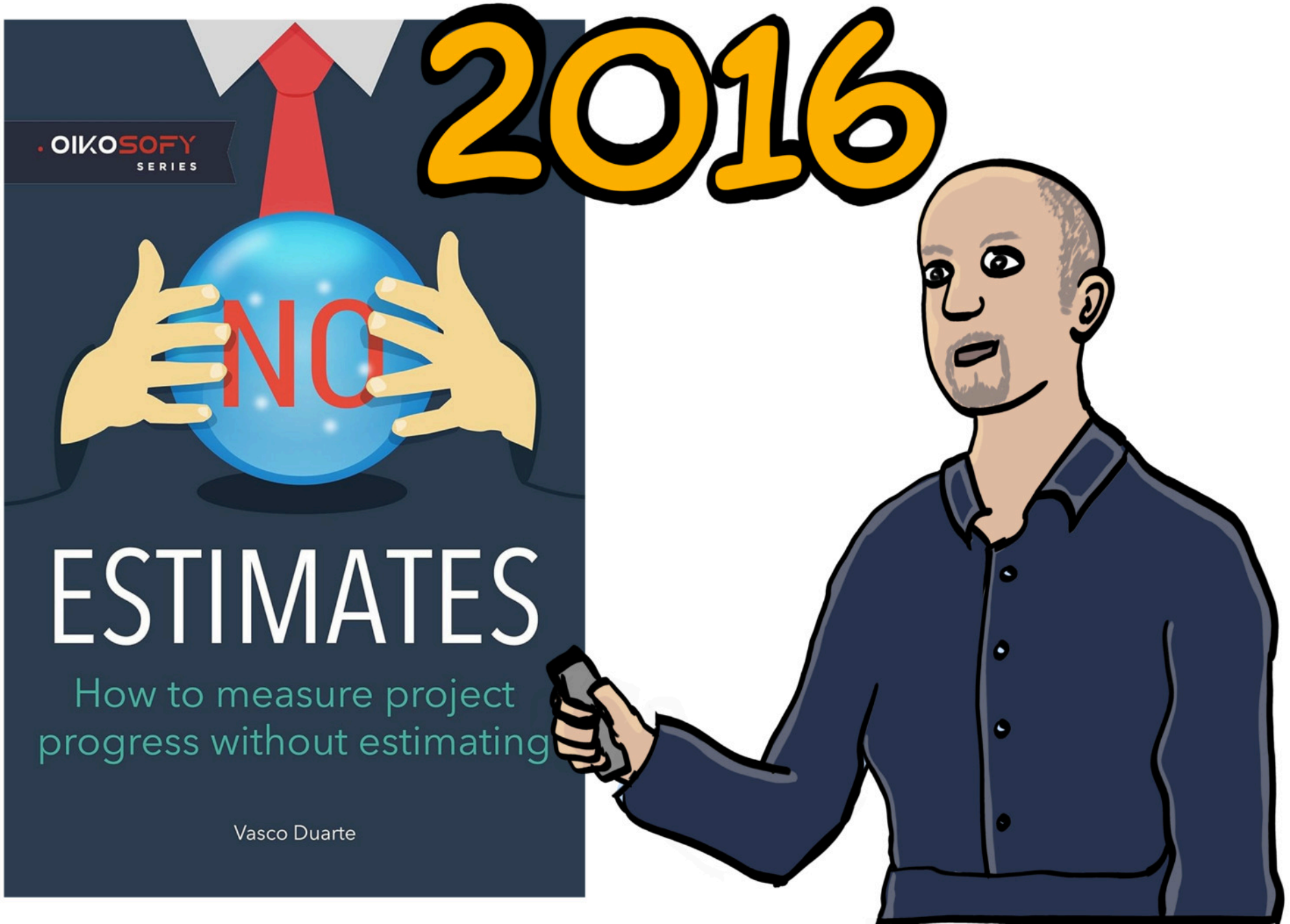

PLANNING

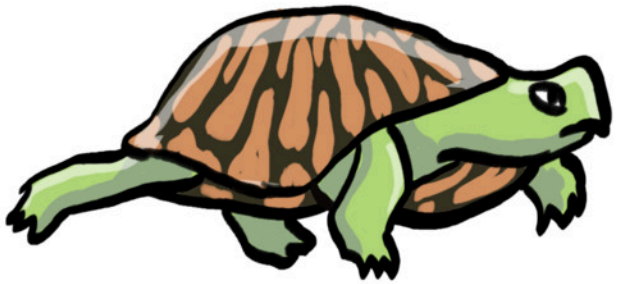
Vasco Duarte

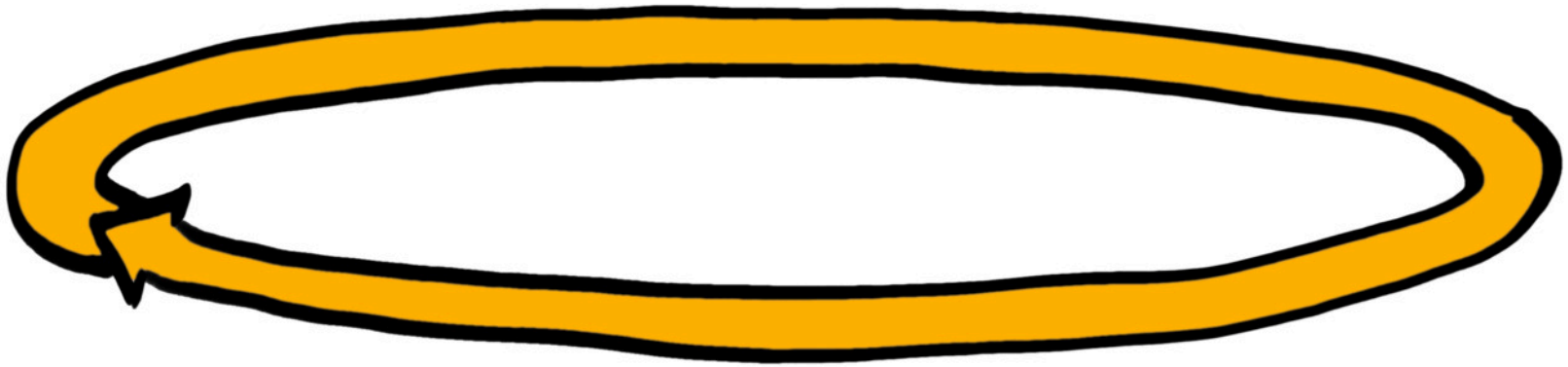


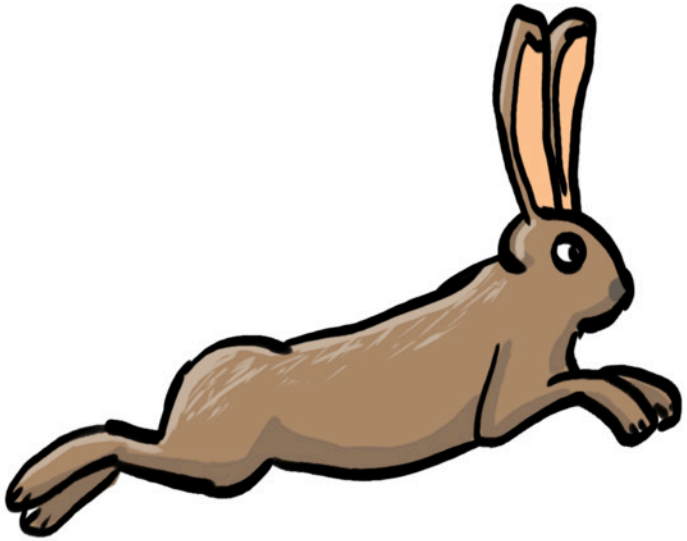
2012 A BETTER WAY TO PREDICT PROJECT RELEASE DATE!

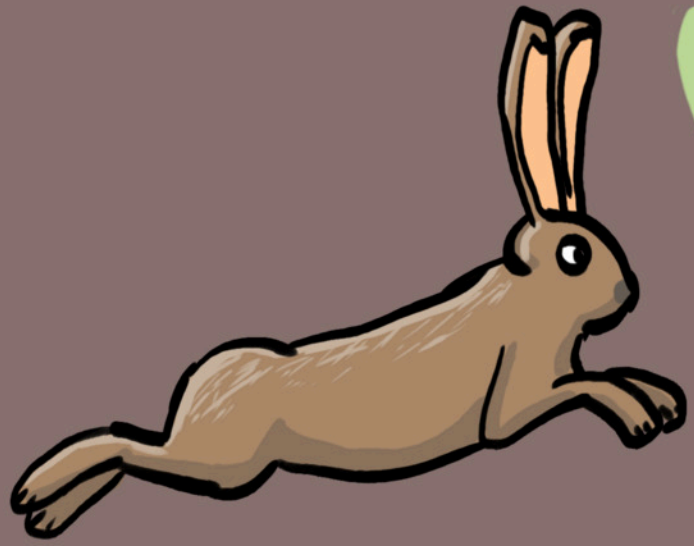


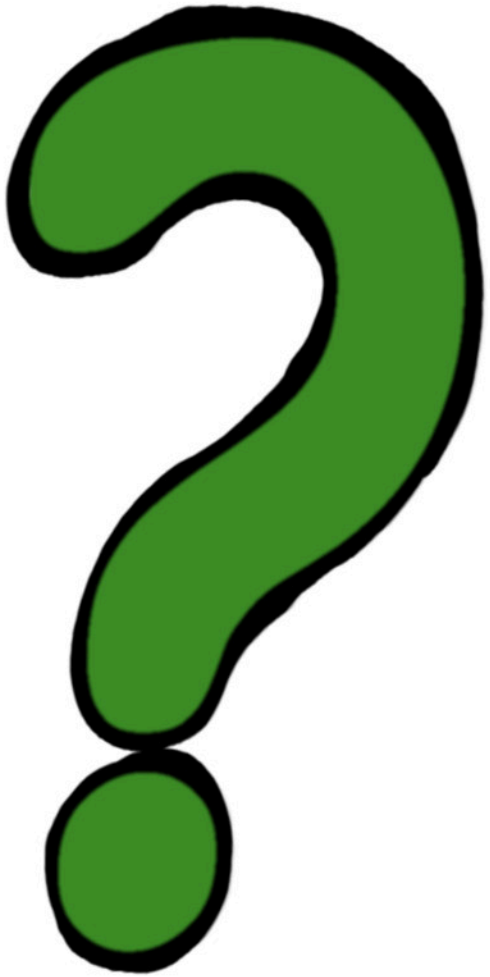

2016
OIKOSOFY
SERIES
NO
ESTIMATES
How to measure project progress without estimating
Vasco Duarte



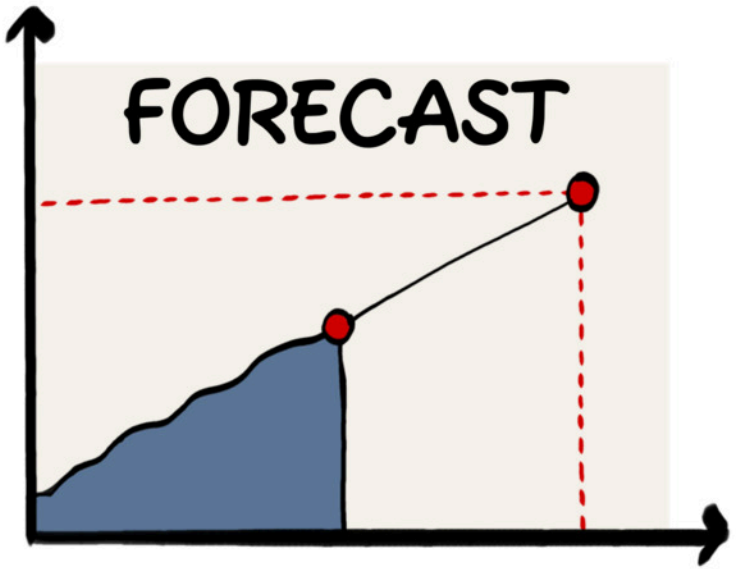






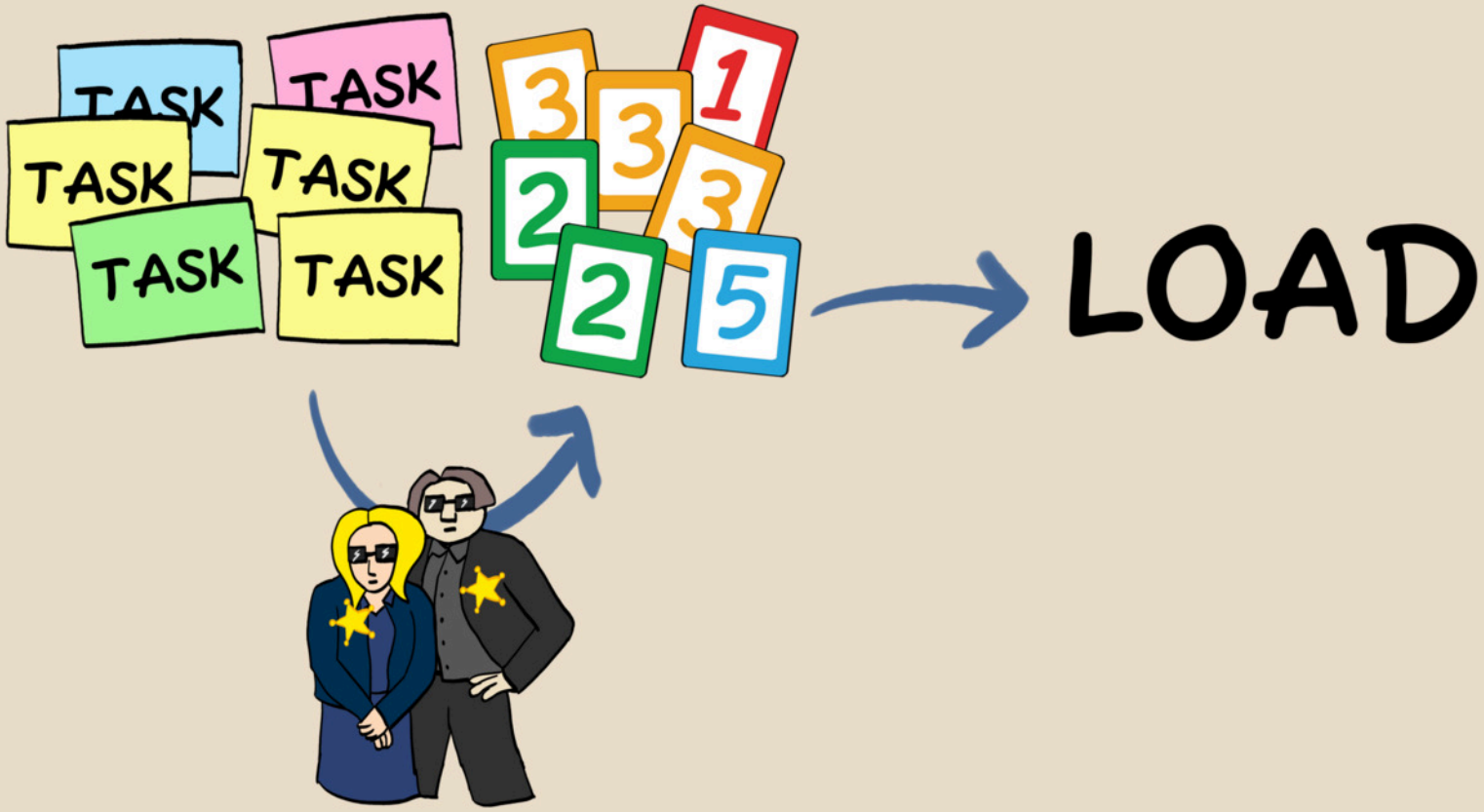


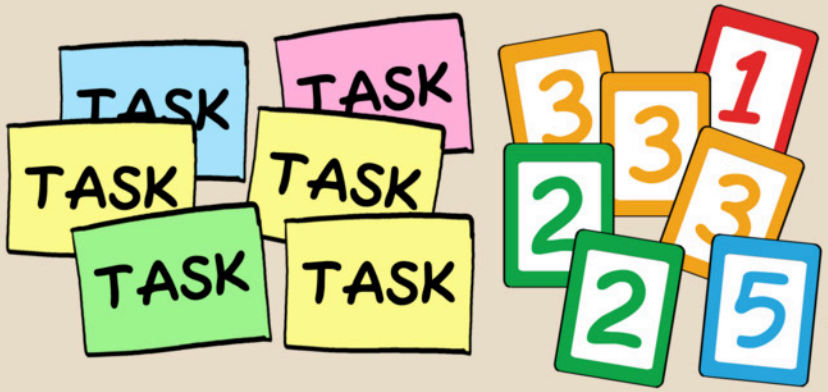
Story Points #NoEstimates



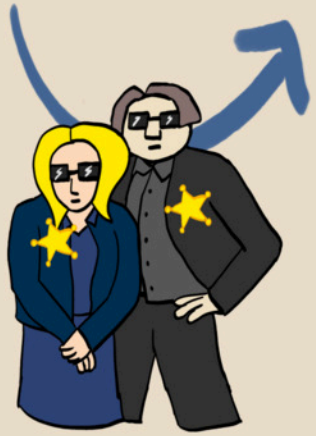

TASK
TASK
TASK
TASK
TASK
TASK

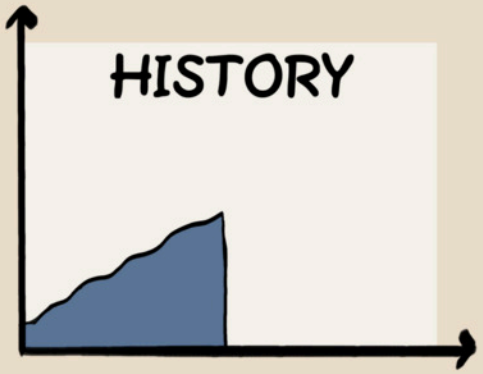

TASK
TASK
TASK
TASK
TASK
TASK
3
3
1
2
3
2
5

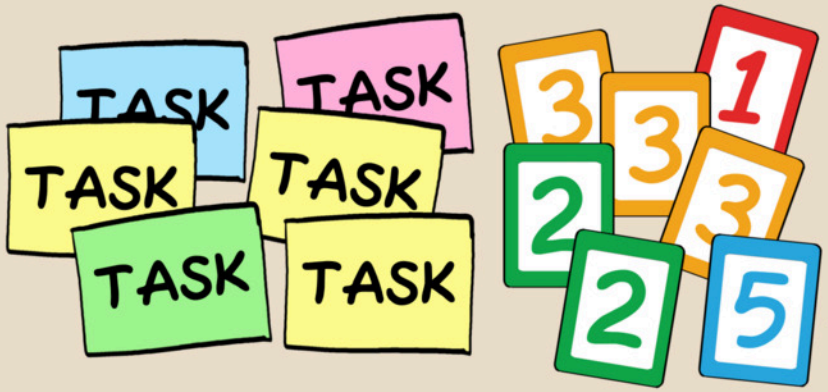

TASK
TASK
TASK
TASK
TASK
TASK
3
3
1
2
3
2
5
LOAD



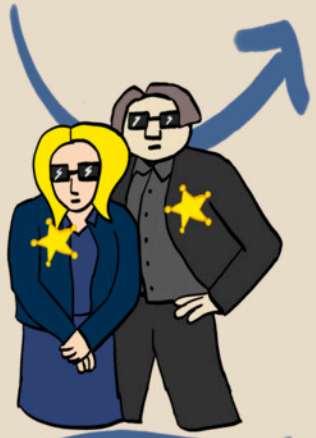
LOAD

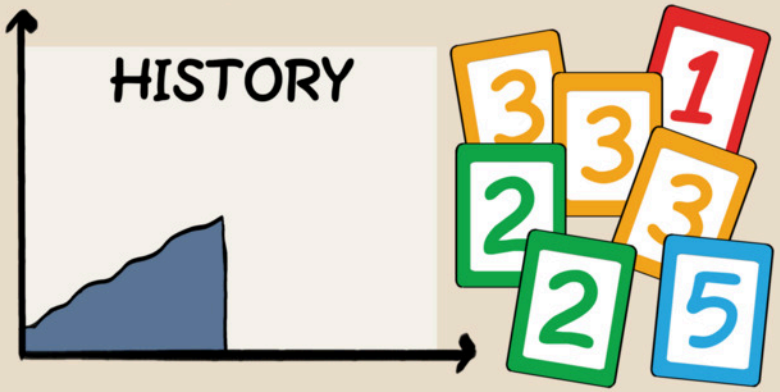


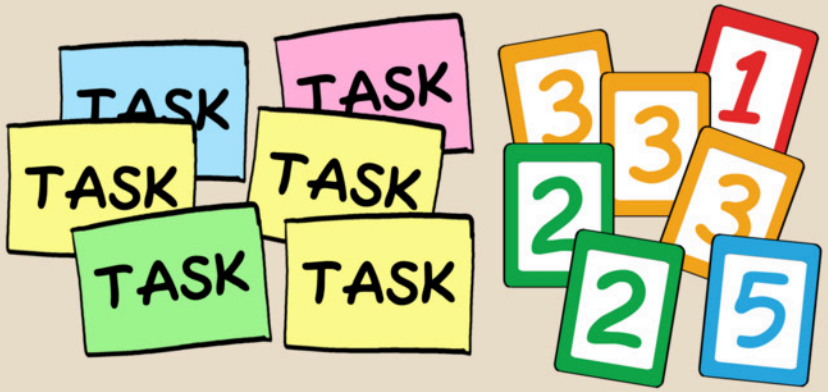




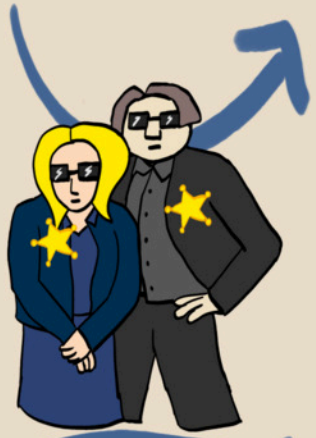
LOAD

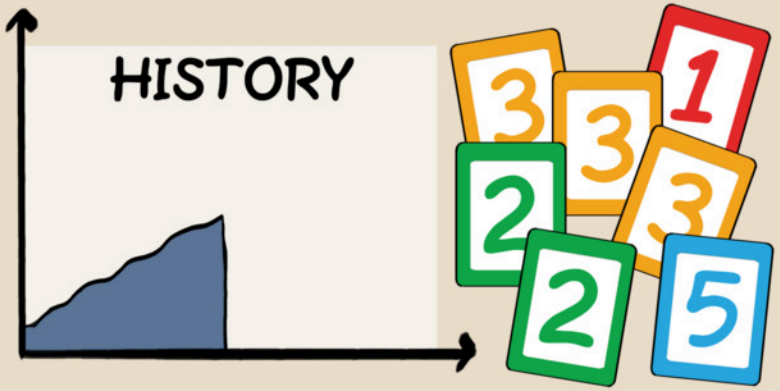




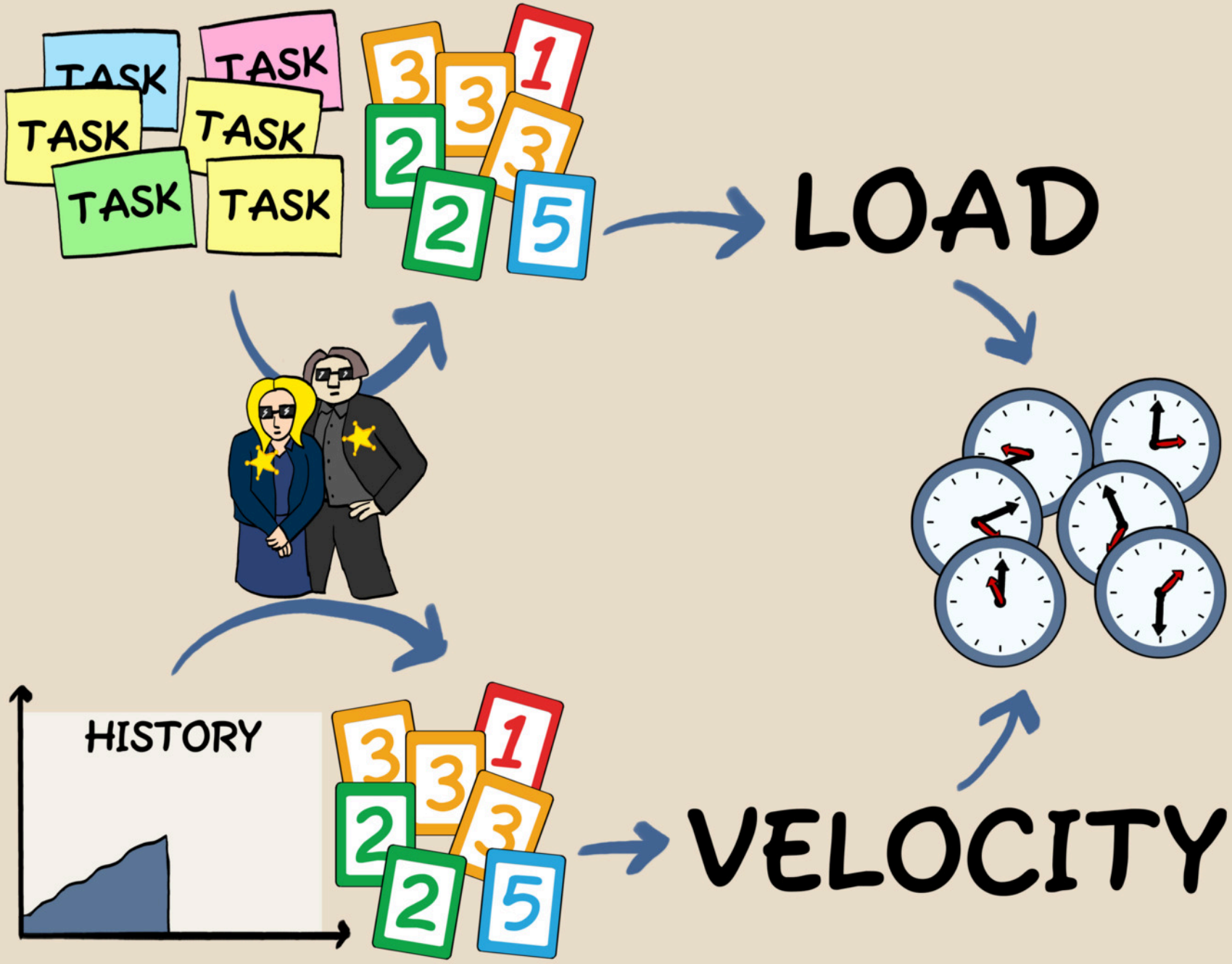


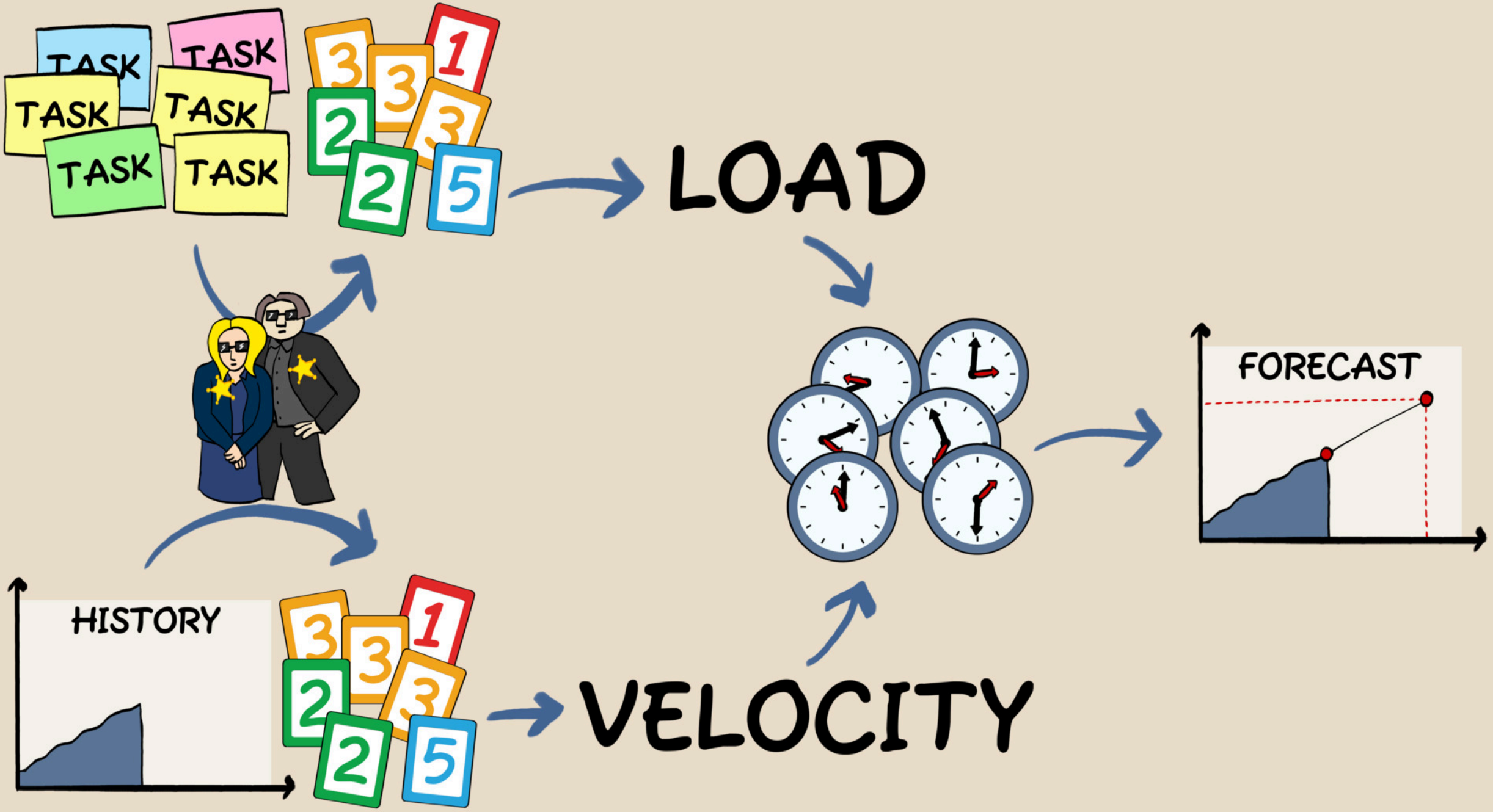
LOAD



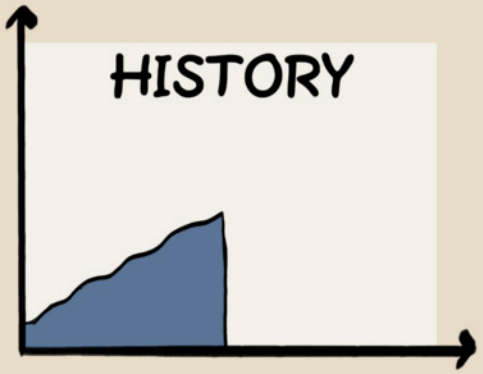


VELOCITY

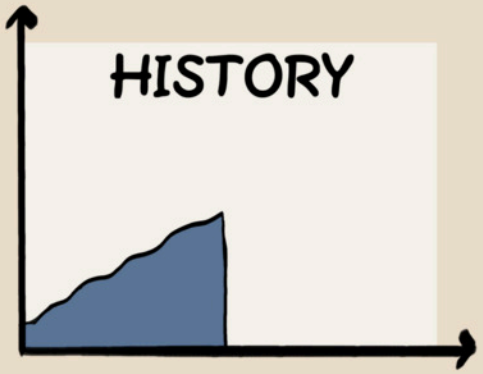

TASK
TASK
TASK
TASK
TASK
TASK
3
3
1
2
3
3
2
5
LOAD
HISTORY
VELOCITY


TASK
TASK
TASK
TASK
TASK
TASK
LOAD
FORECAST
HISTORY
VELOCITY

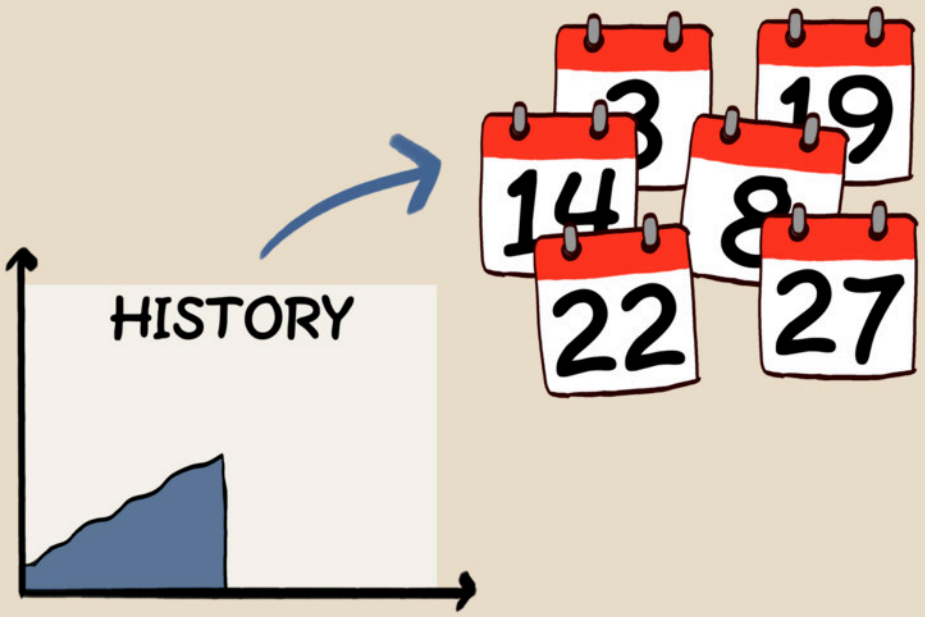

TASK
TASK
TASK
TASK
TASK
TASK


HISTORY

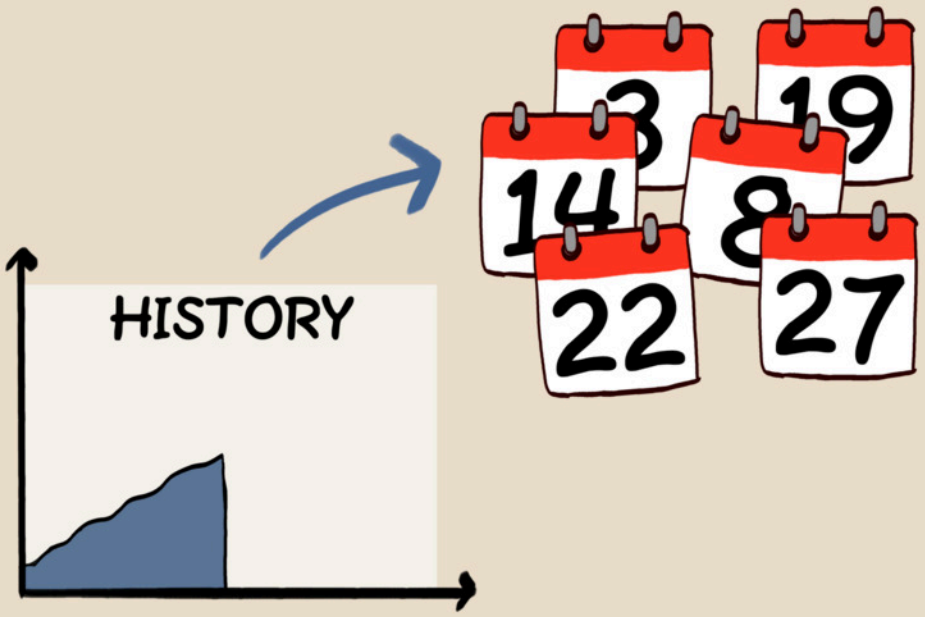
LOAD



LOAD

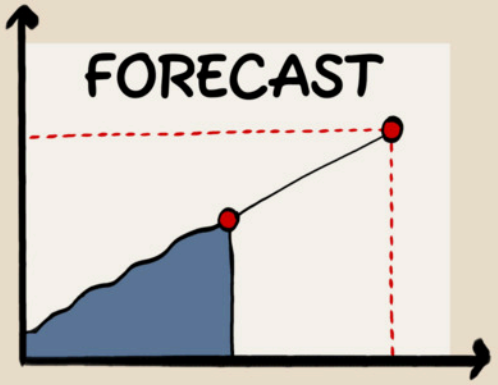


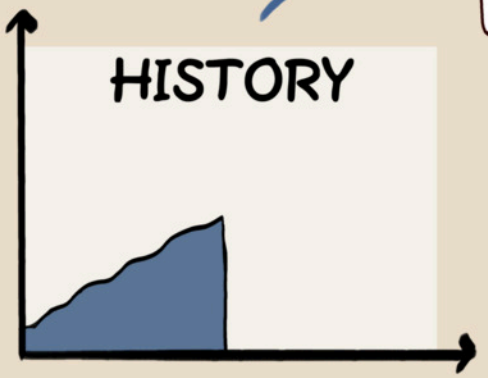
LOAD



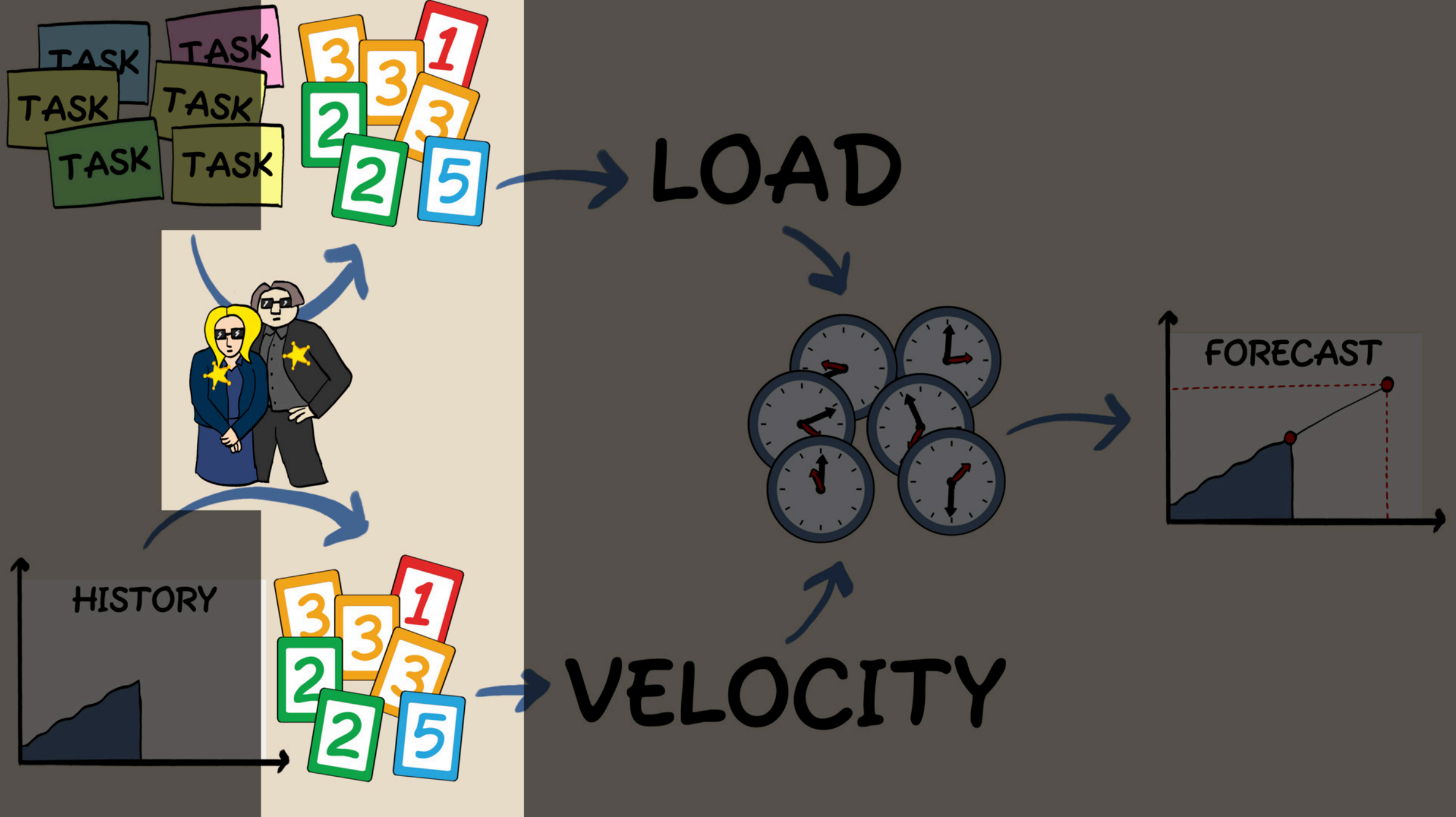
VELOCITY

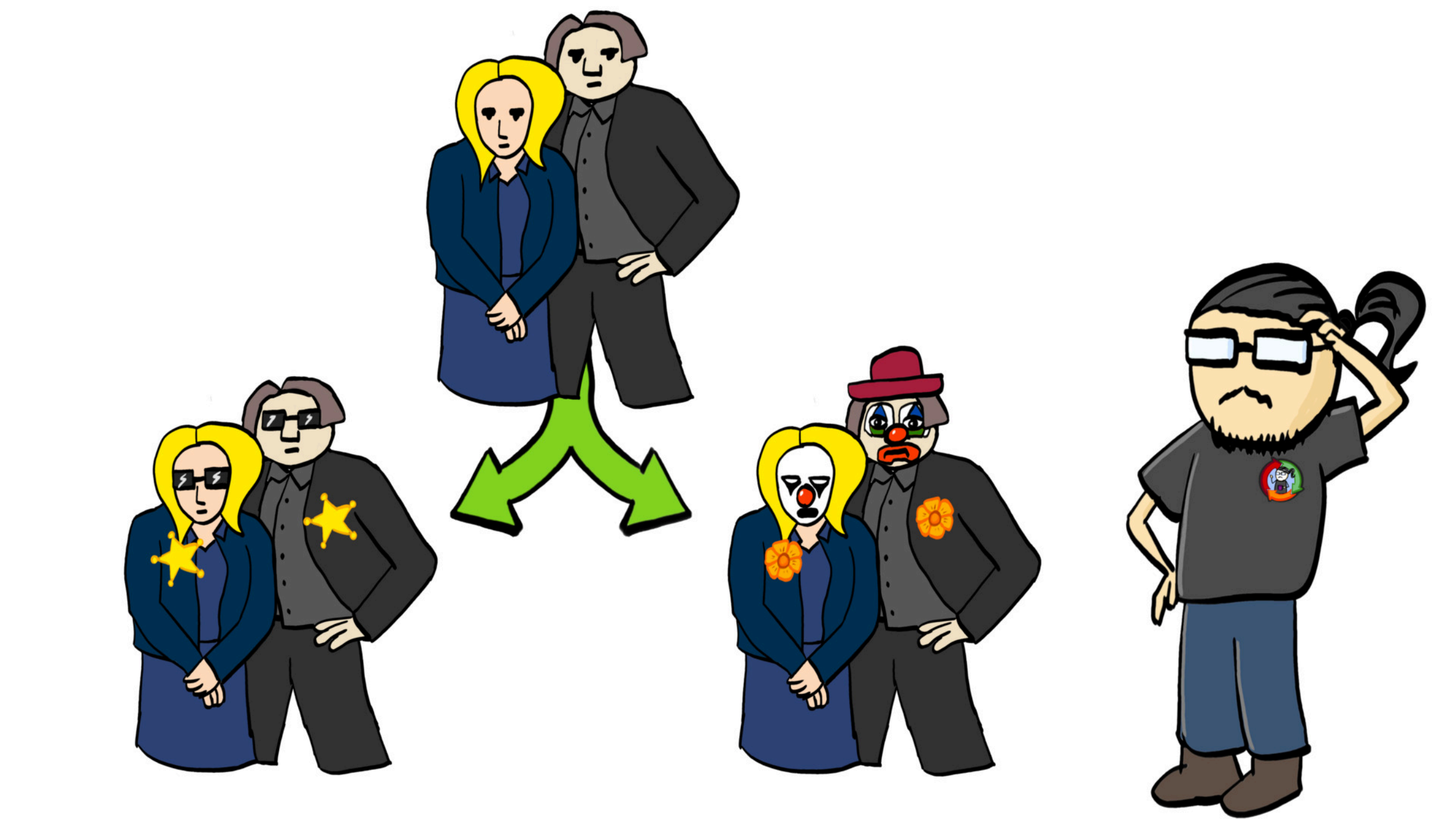
LOAD





VELOCITY


TASK
TASK
TASK
TASK
TASK
TASK
LOAD
HISTORY
VELOCITY
FORECAST



I CONTEXT

II STORY POINTS

III

IV

V

VI

VII

ON THE RELATIONSHIP BETWEEN STORY POINT AND DEVELOPMENT EFFORT IN AGILE OPEN-SOURCE SOFTWARE

Tawosi, Vali and Moussa, Rebecca and Sarro, Federica

2022

TAWOS DATASET

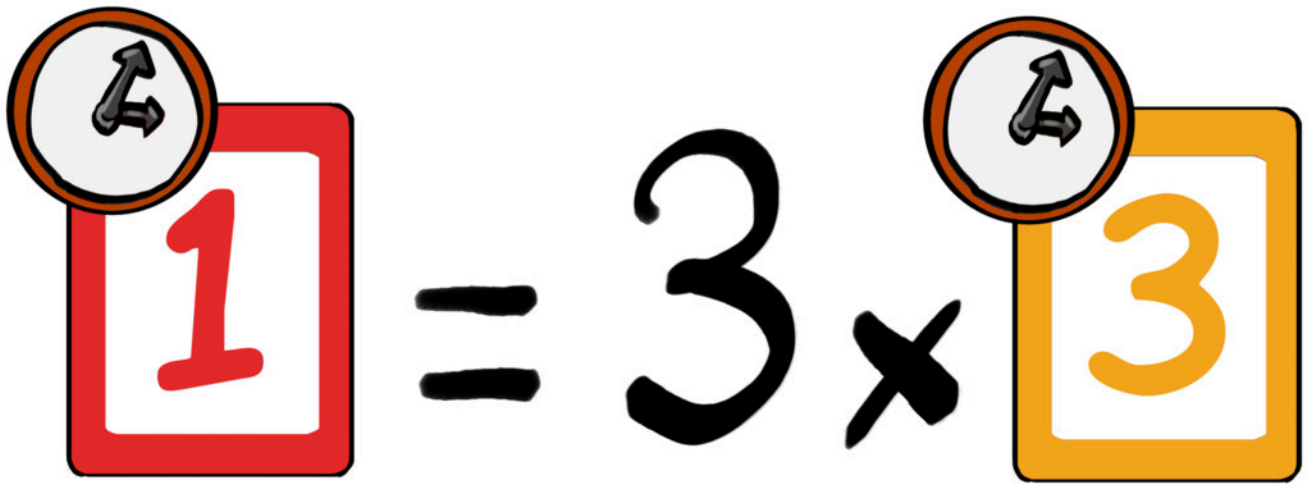
39 PROJECTS
~450 000 ISSUES

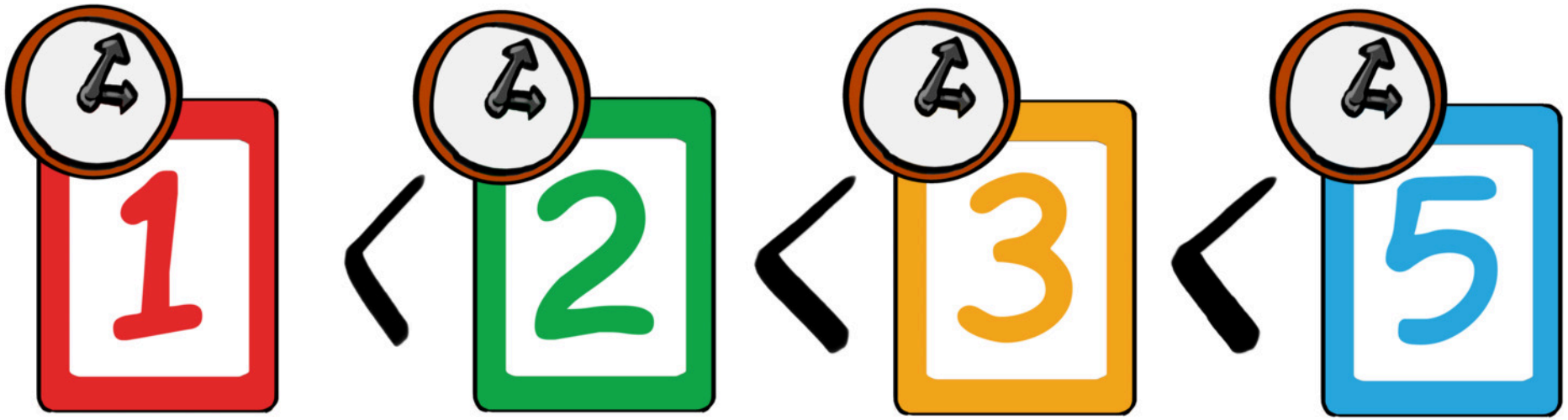
spring

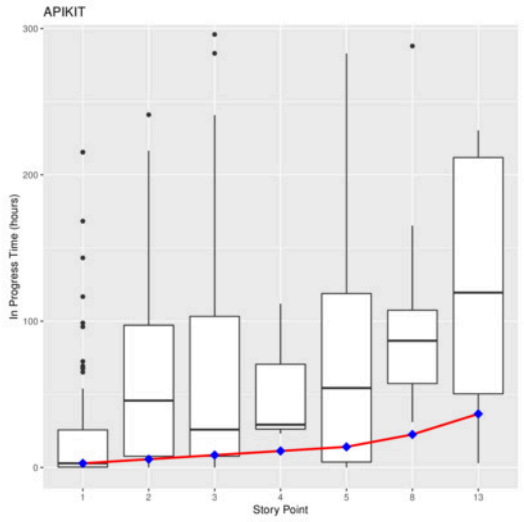
mongo DB

APACHE

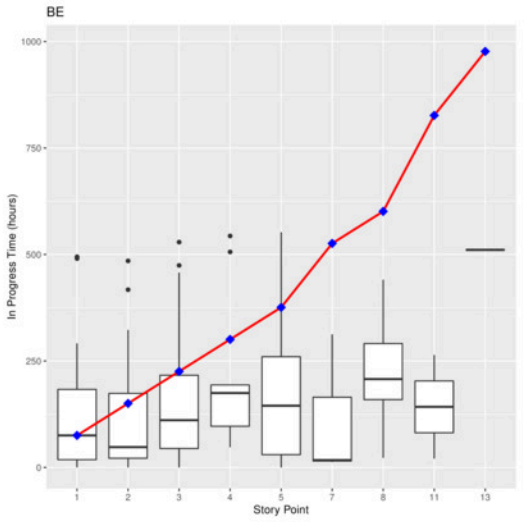
ATLASSIAN


1
=
3
×
3

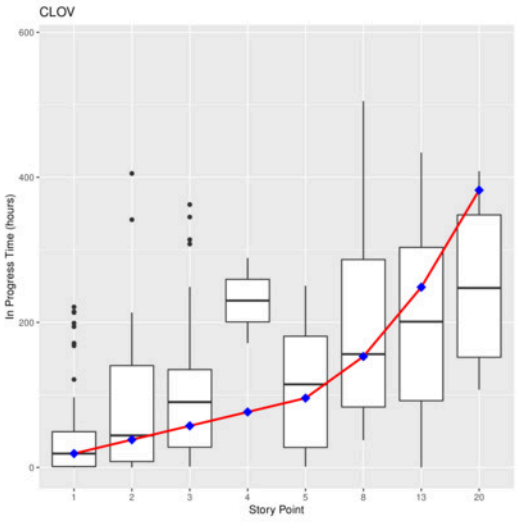

1
<
2
<
3
<
5



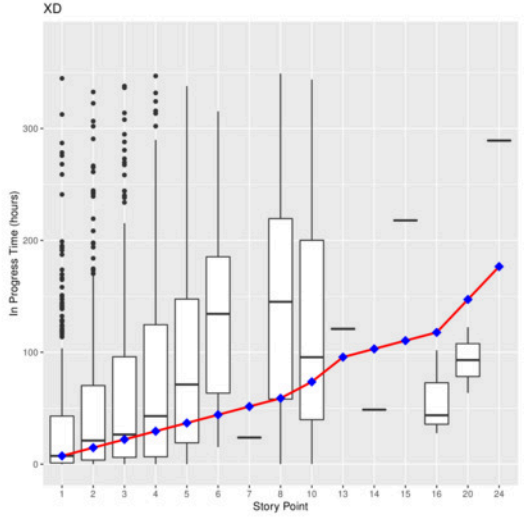
(a) APIKIT



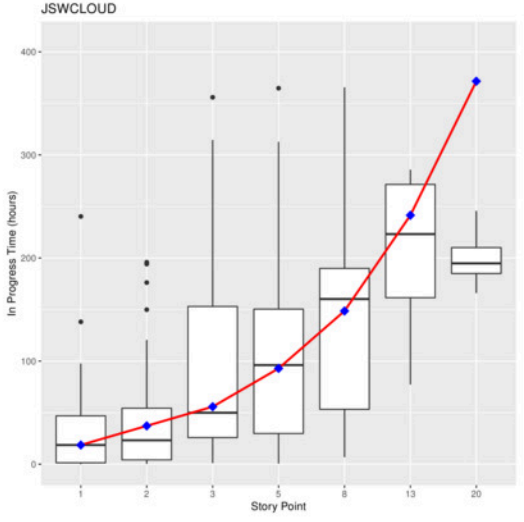
(b) BE



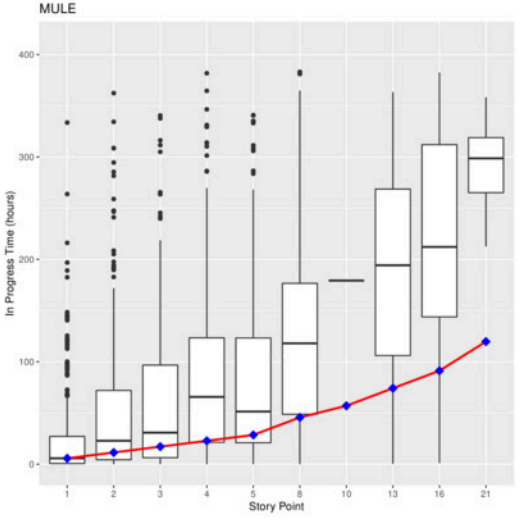
(c) CLOV



(d) XD

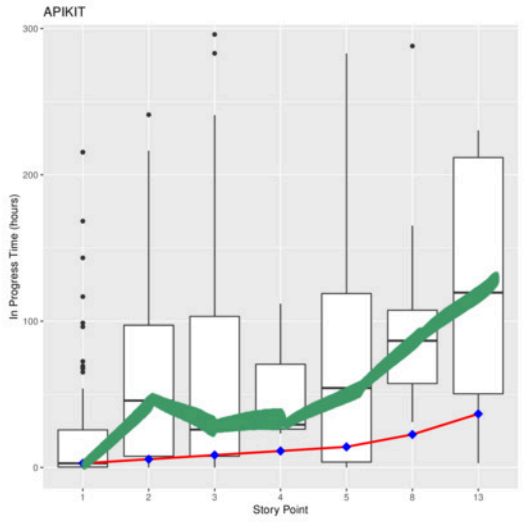


(e) JSWCLOUD

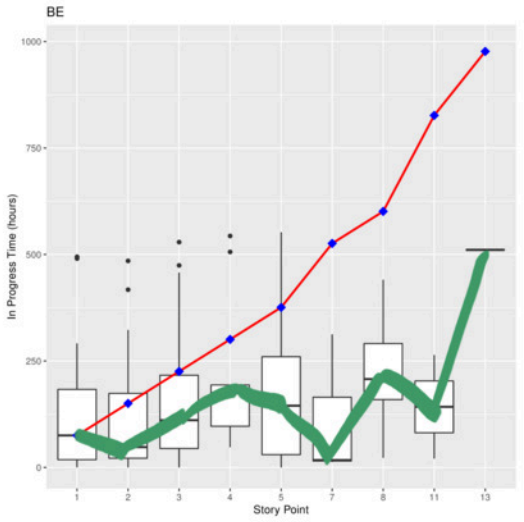


(f) MULE

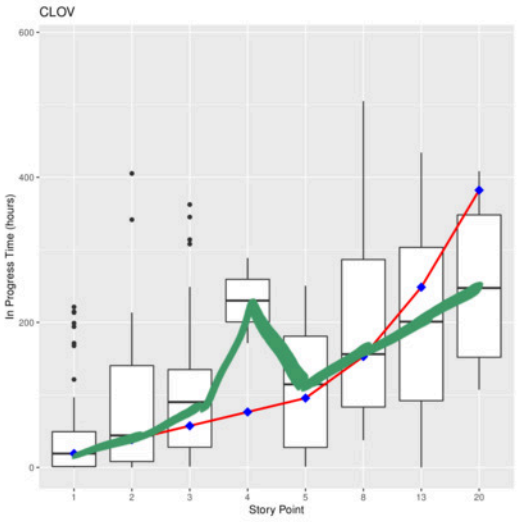
Figure 2: Boxplots of the distribution of development time per SP class for (a) APIKIT, (b) BE, (c) CLOV, (d) XD, (e) JSWCLOUD, (f) MULE. The red line depicts a project-specific baseline, drawn based on the median development time for one SP.



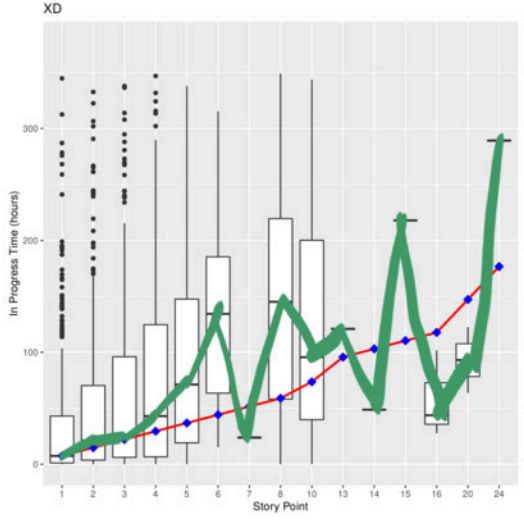
(a) APIKIT



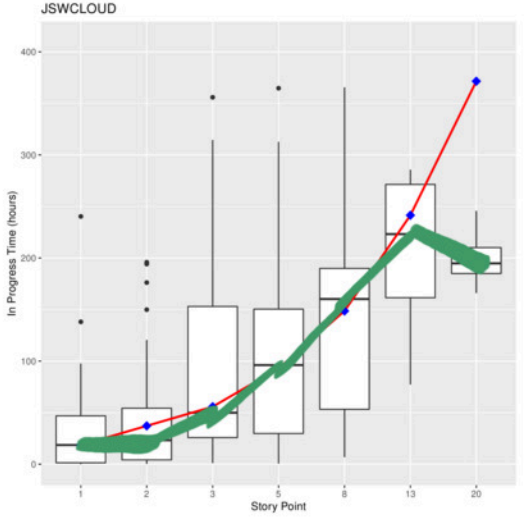
(b) BE



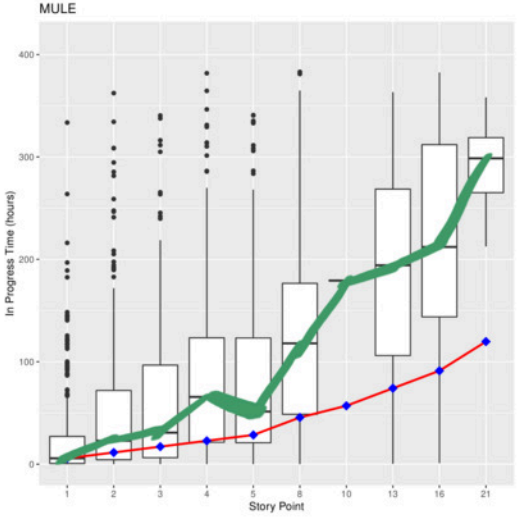
(c) CLOV



(d) XD



(e) JSWCLOUD

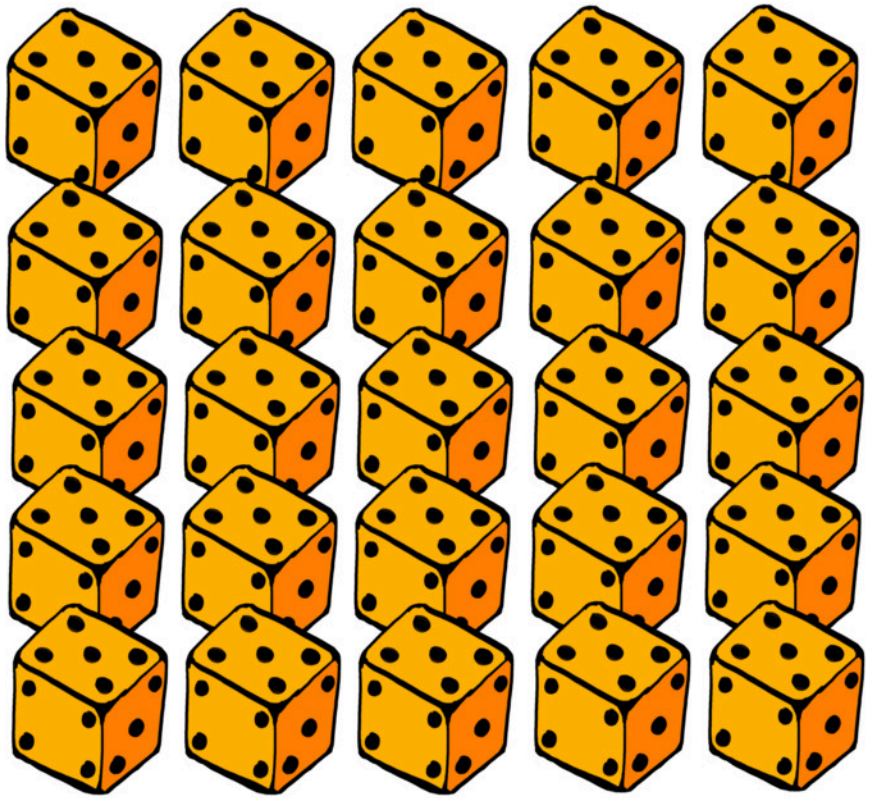


(f) MULE

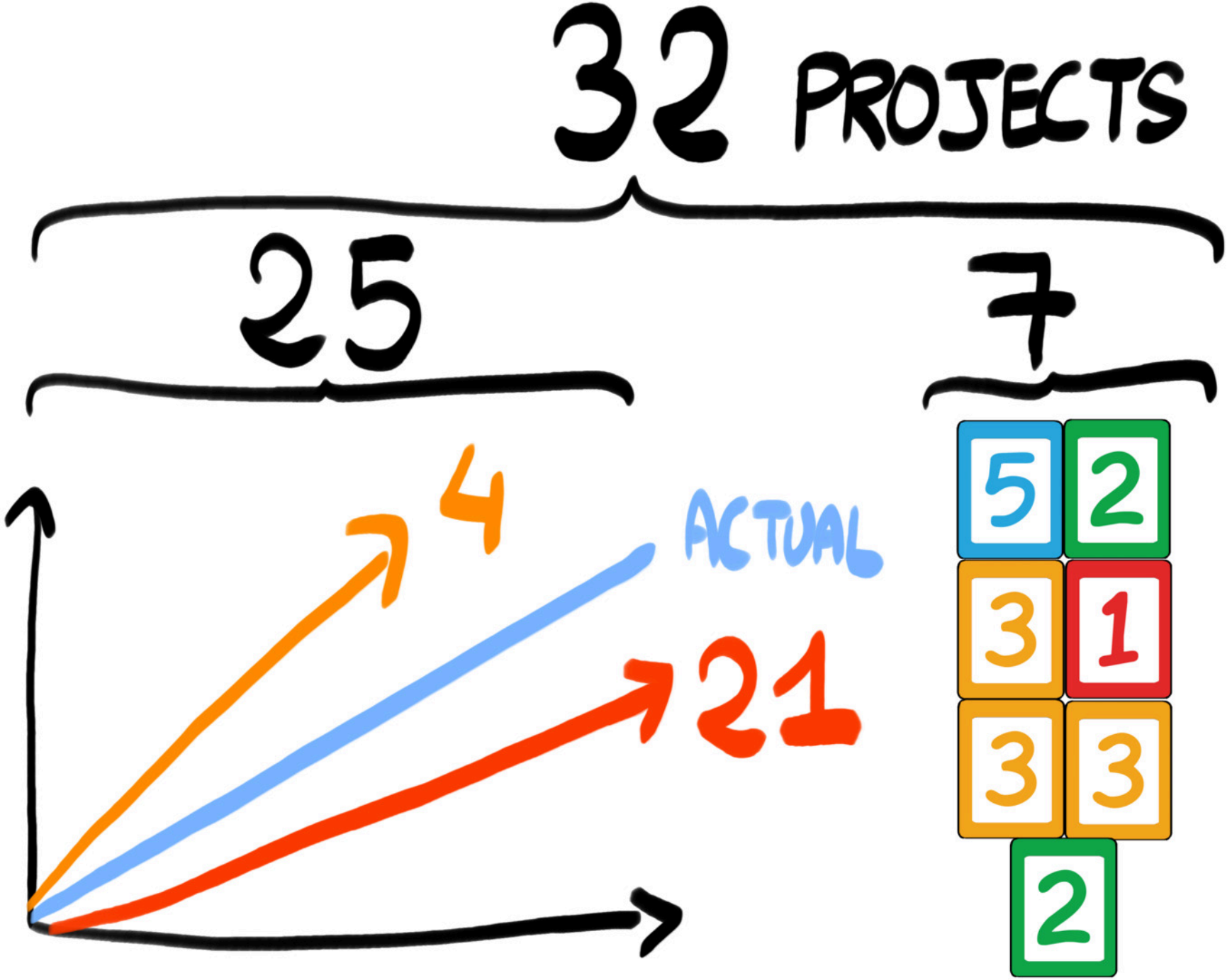
Figure 2: Boxplots of the distribution of development time per SP class for (a) APIKIT, (b) BE, (c) CLOV, (d) XD, (e) JSWCLOUD, (f) MULE. The red line depicts a project-specific baseline, drawn based on the median development time for one SP.

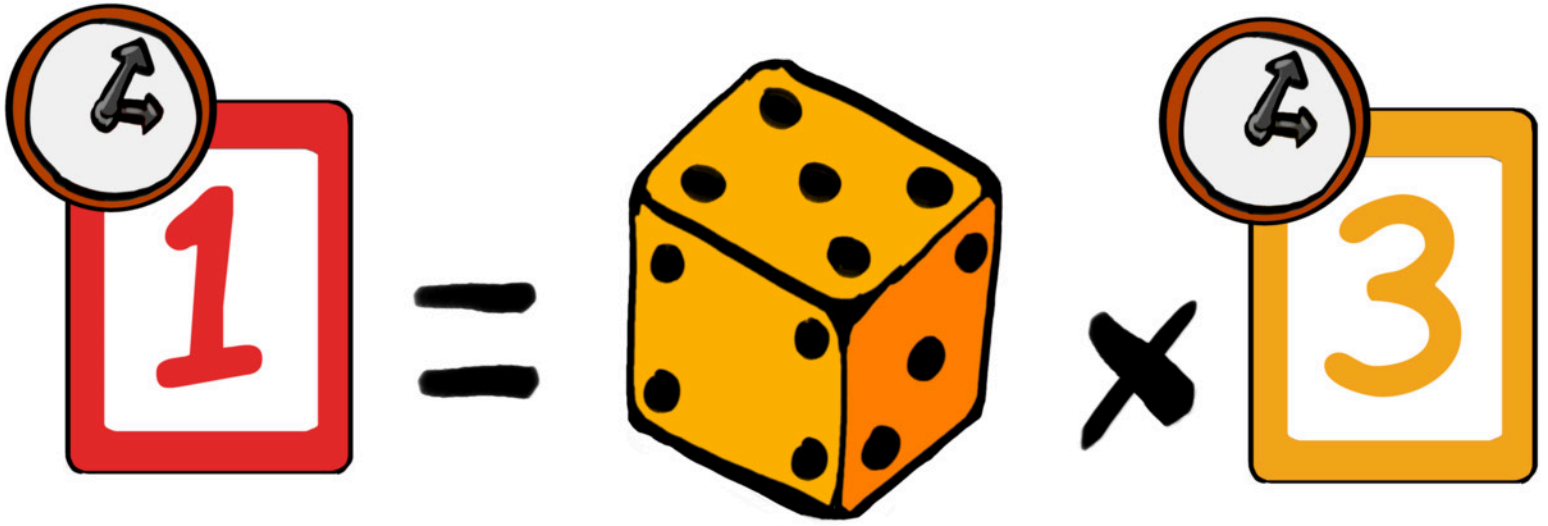
32 PROJECTS

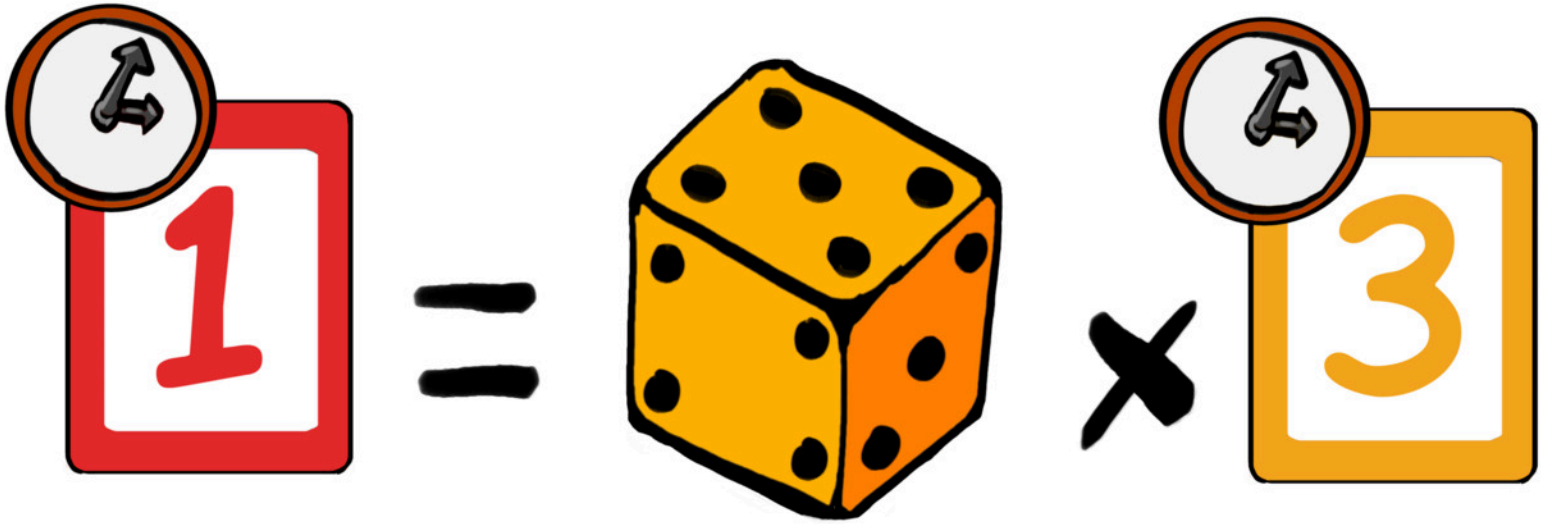
25

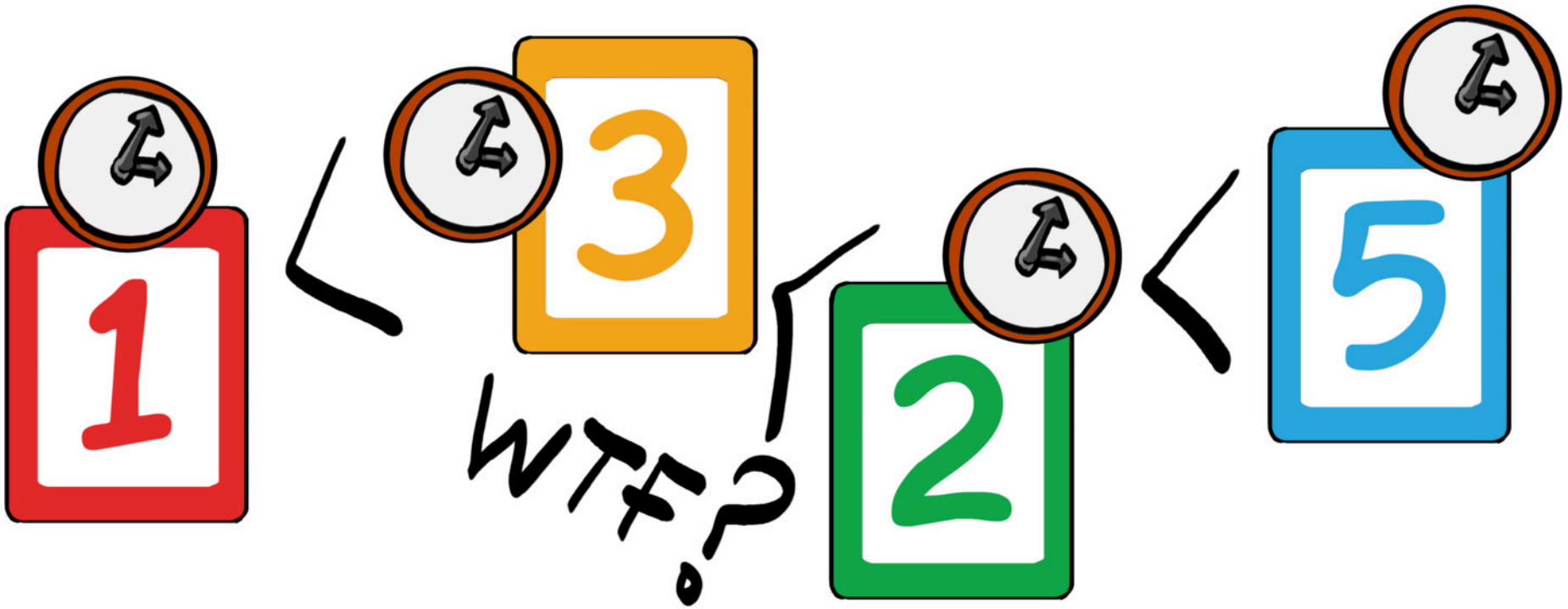


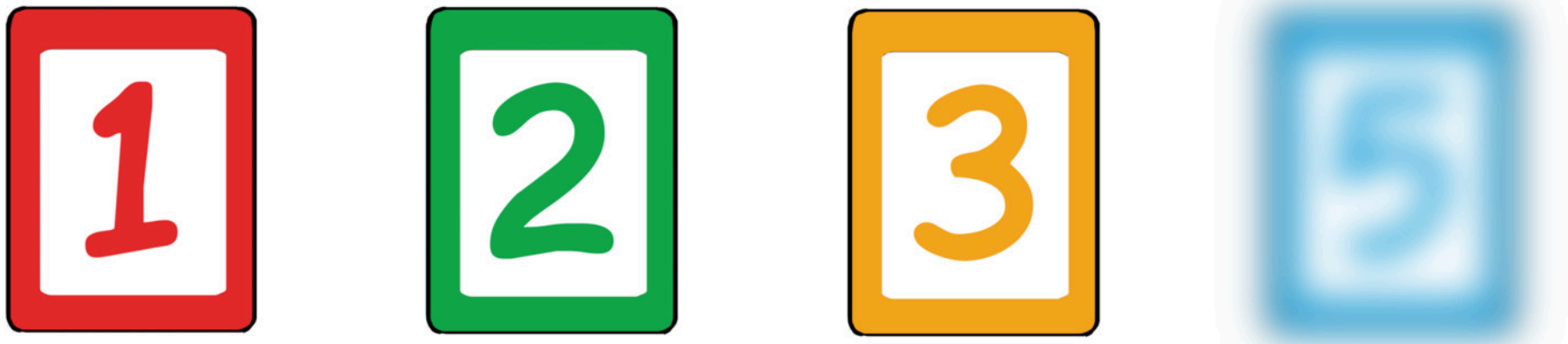
7

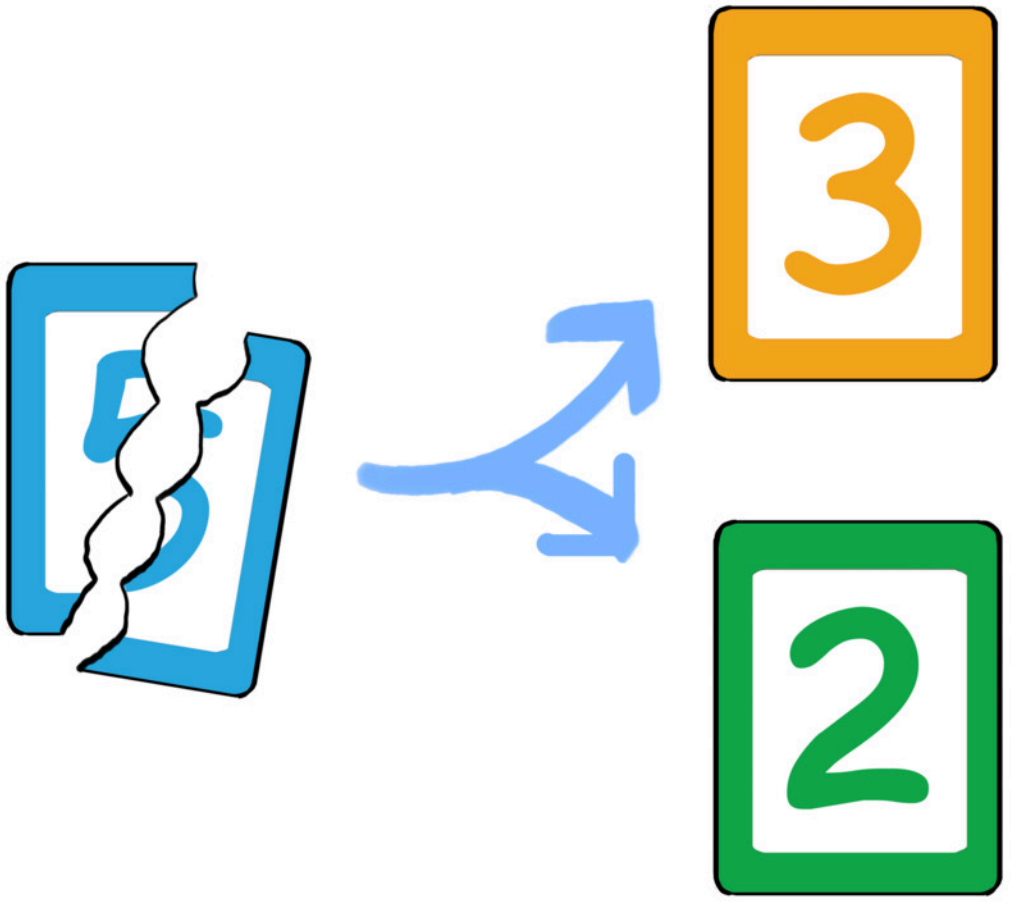

32 PROJECTS
25
7
4
ACTUAL
21
5
2
3
1
3
3
2


1
=
×
3


1
=
×
3


1
<
3
WTF?
2
<
5


1
2
3
5


5
3
2

SP showed a low (50%) or medium (25%) correlation with Timespent

SEXTAMT: A SYSTEMATIC MAP TO NAVIGATE THE WIDE SEAS OF FACTORS AFFECTING EXPERT JUDGMENT SOFTWARE ESTIMATES

Matsubara, Patrícia and Gadelha, Bruno and Steinmacher, Igor and Conte, Tayana

2021 - Journal of Systems and Software

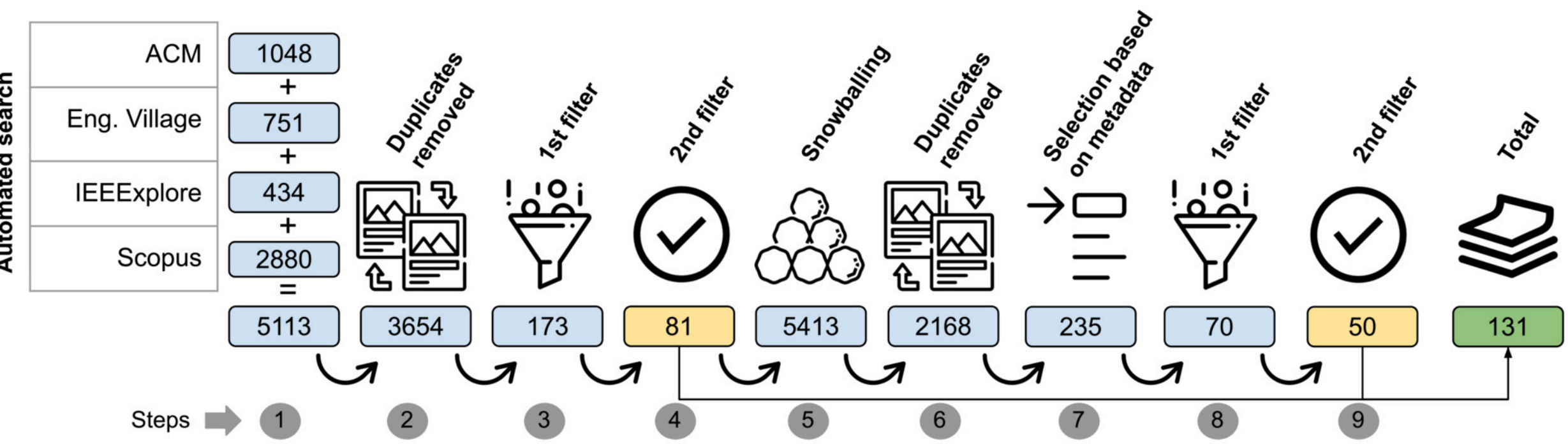


Figure 1 - Search and selection results

235 FACTORS

69 EXPLORED SEVERAL TIMES

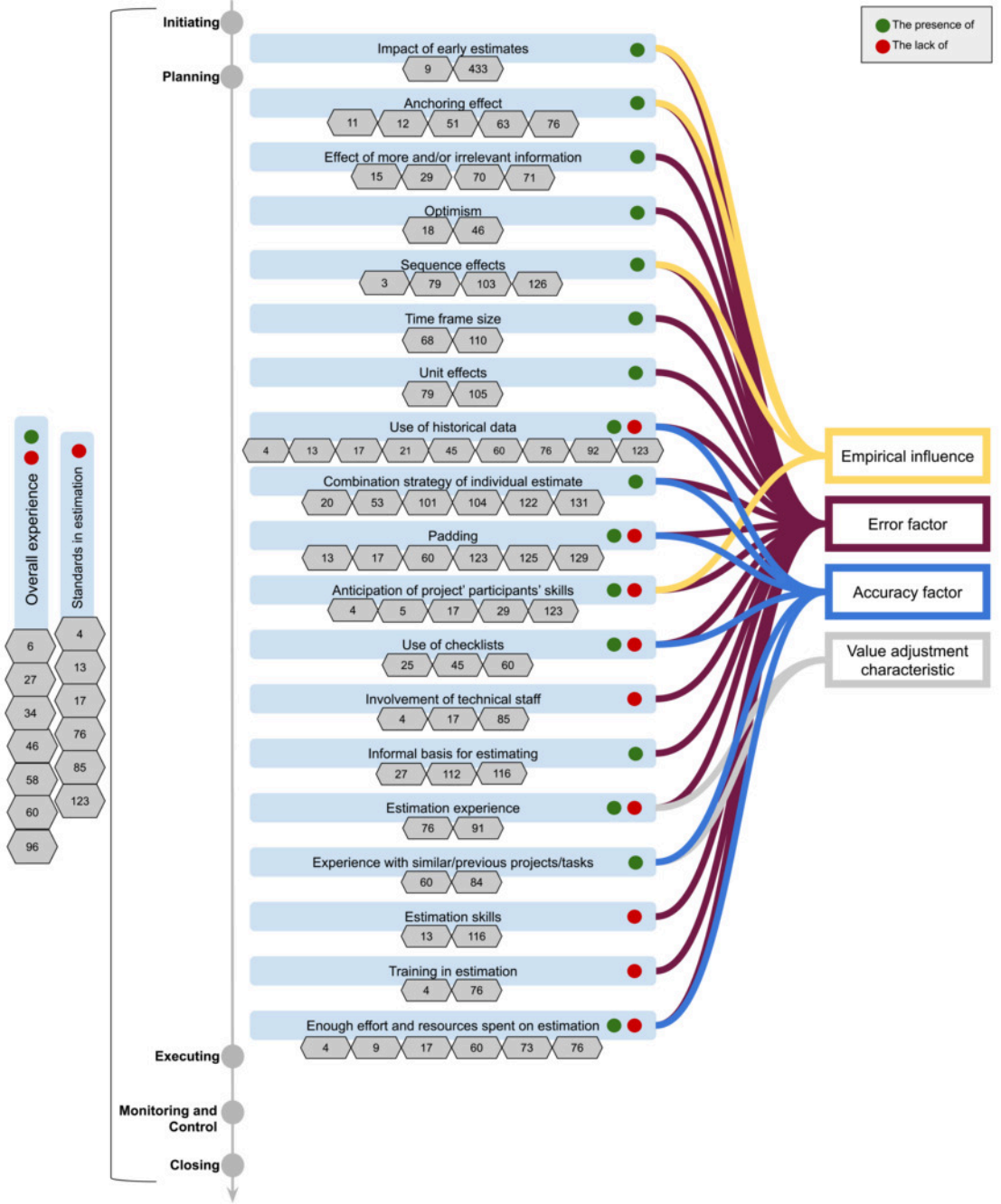
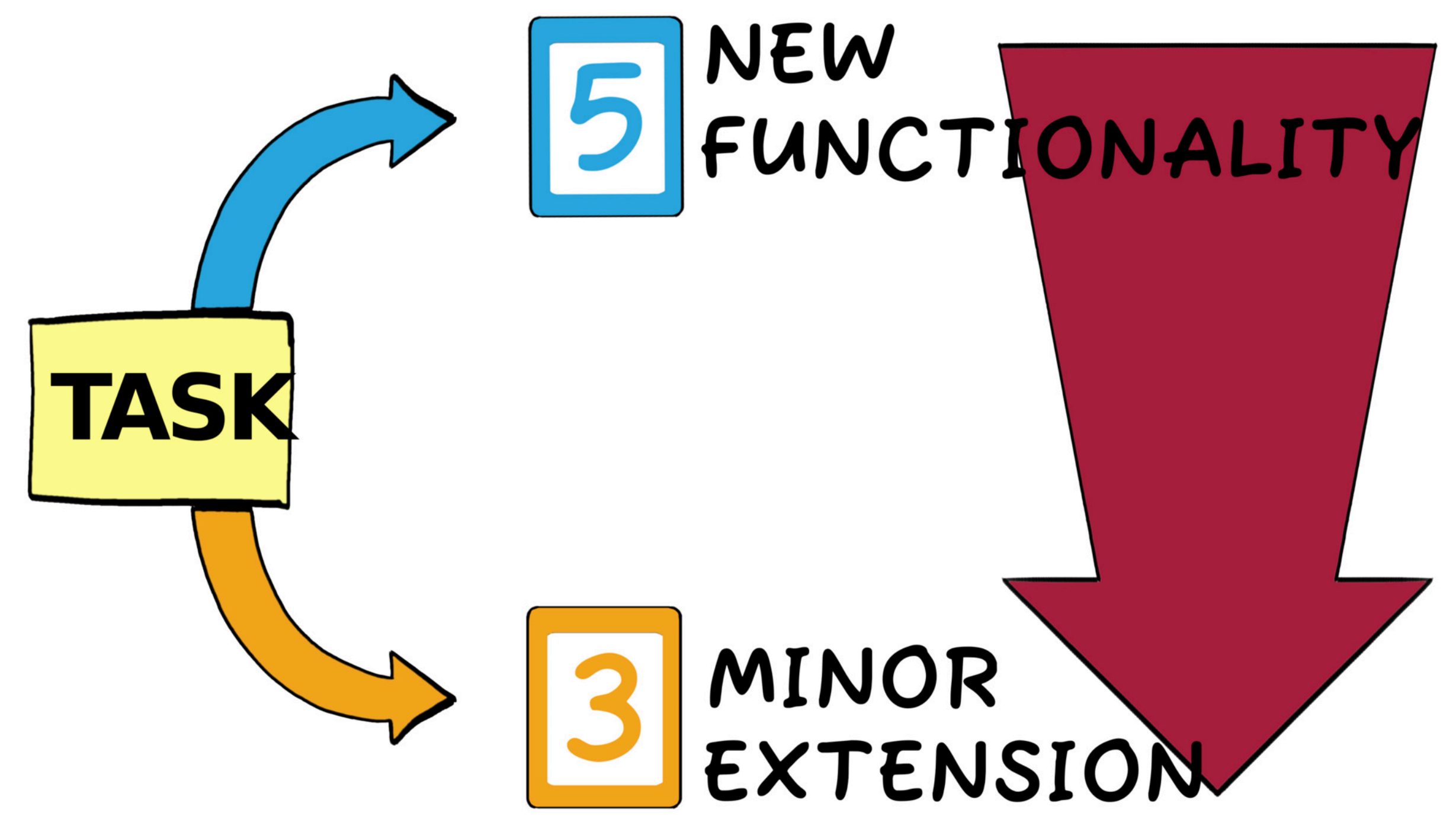
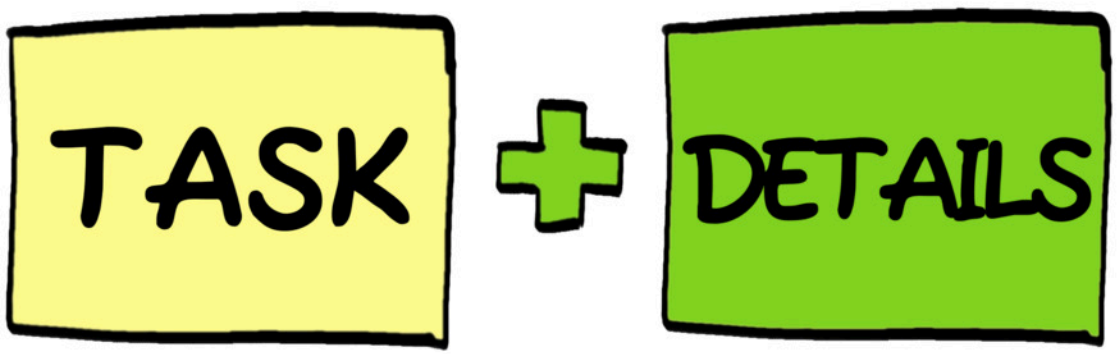


Figure 9 - Factors related to Estimators


5
NEW
FUNCTIONALITY
TASK

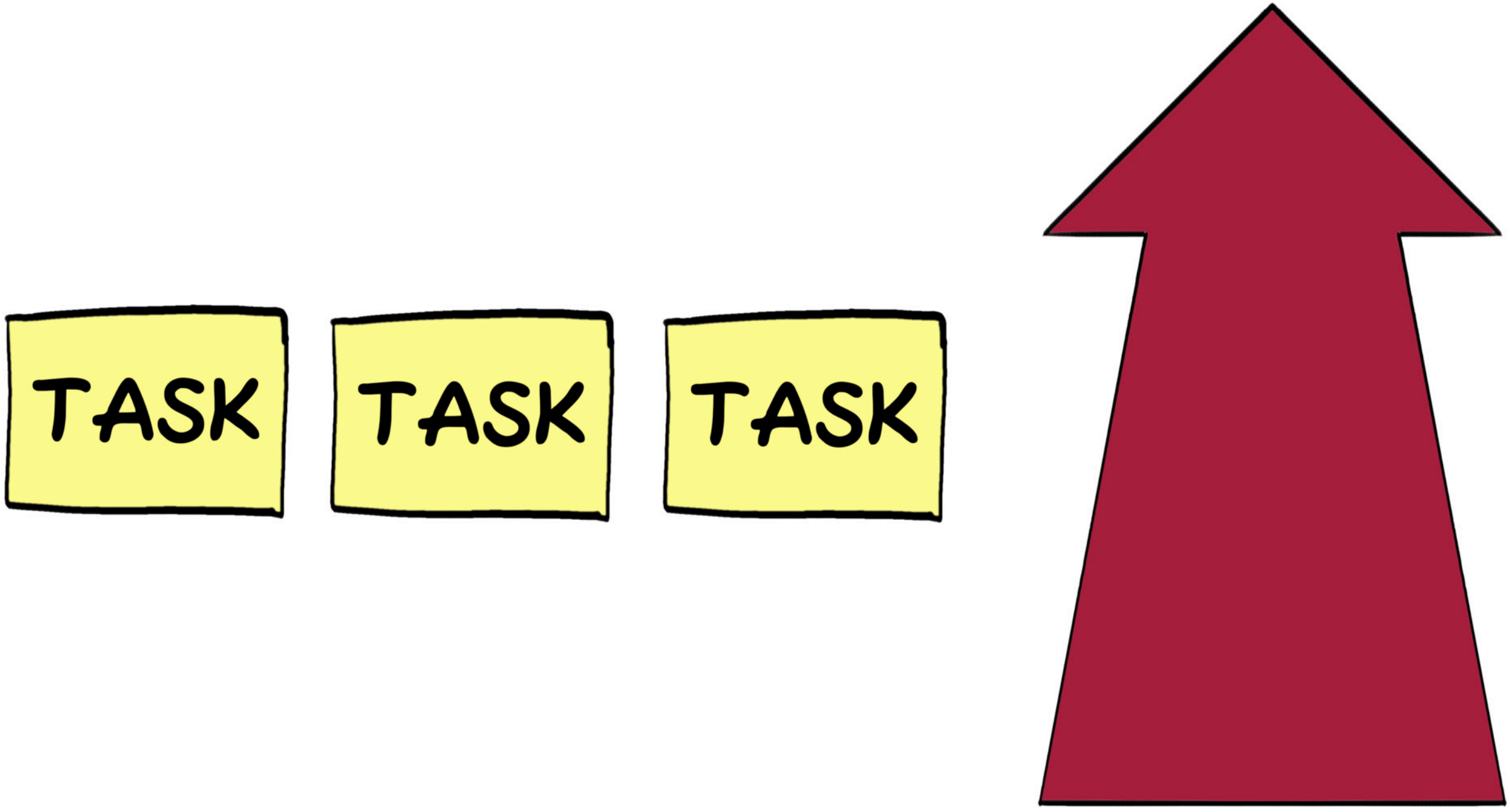

TASK
5
NEW FUNCTIONALITY
3
MINOR EXTENSION


TASK

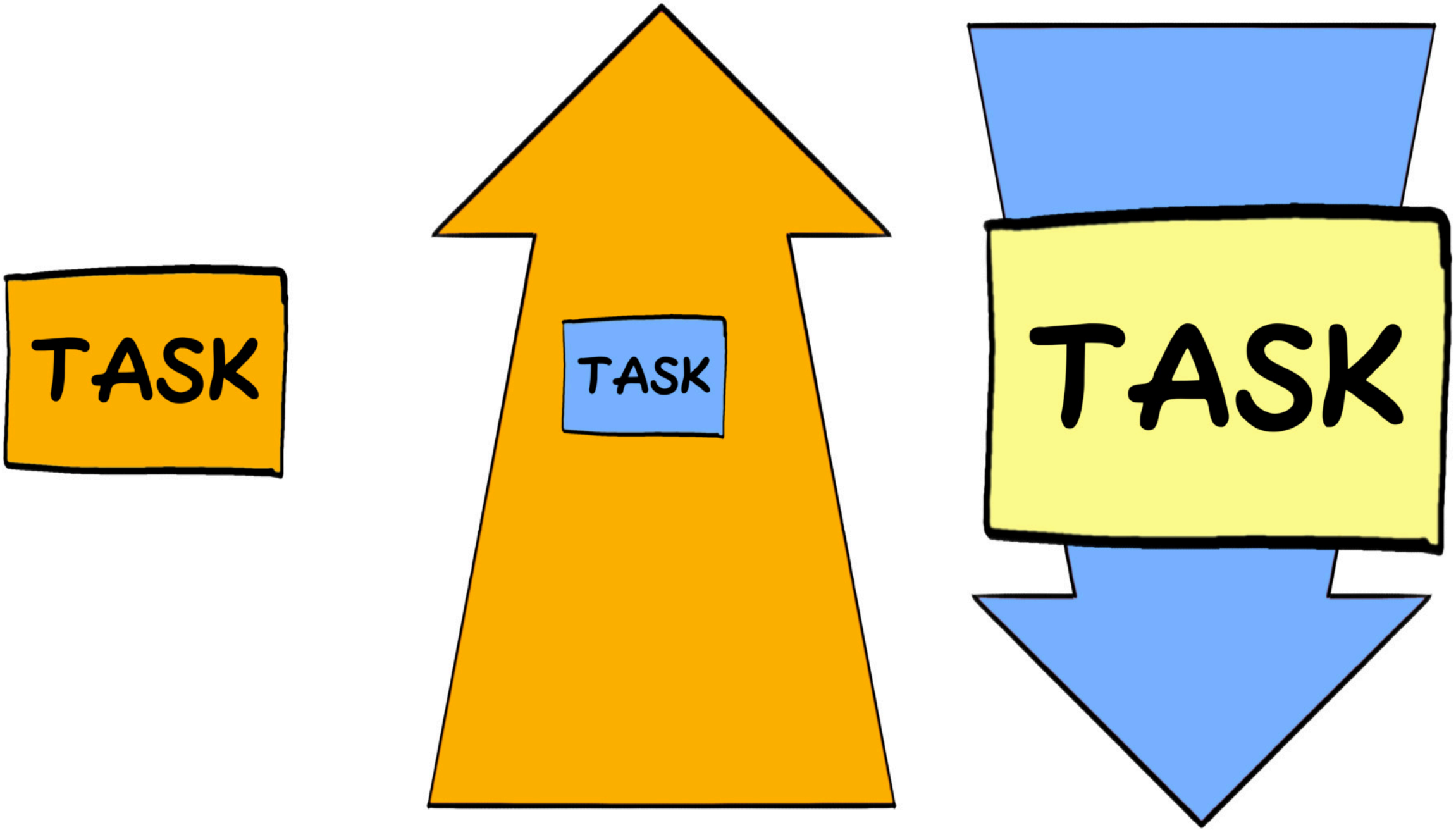

TASK
DETAILS


EASY!
TASK
DETAILS


TASK
TASK
TASK

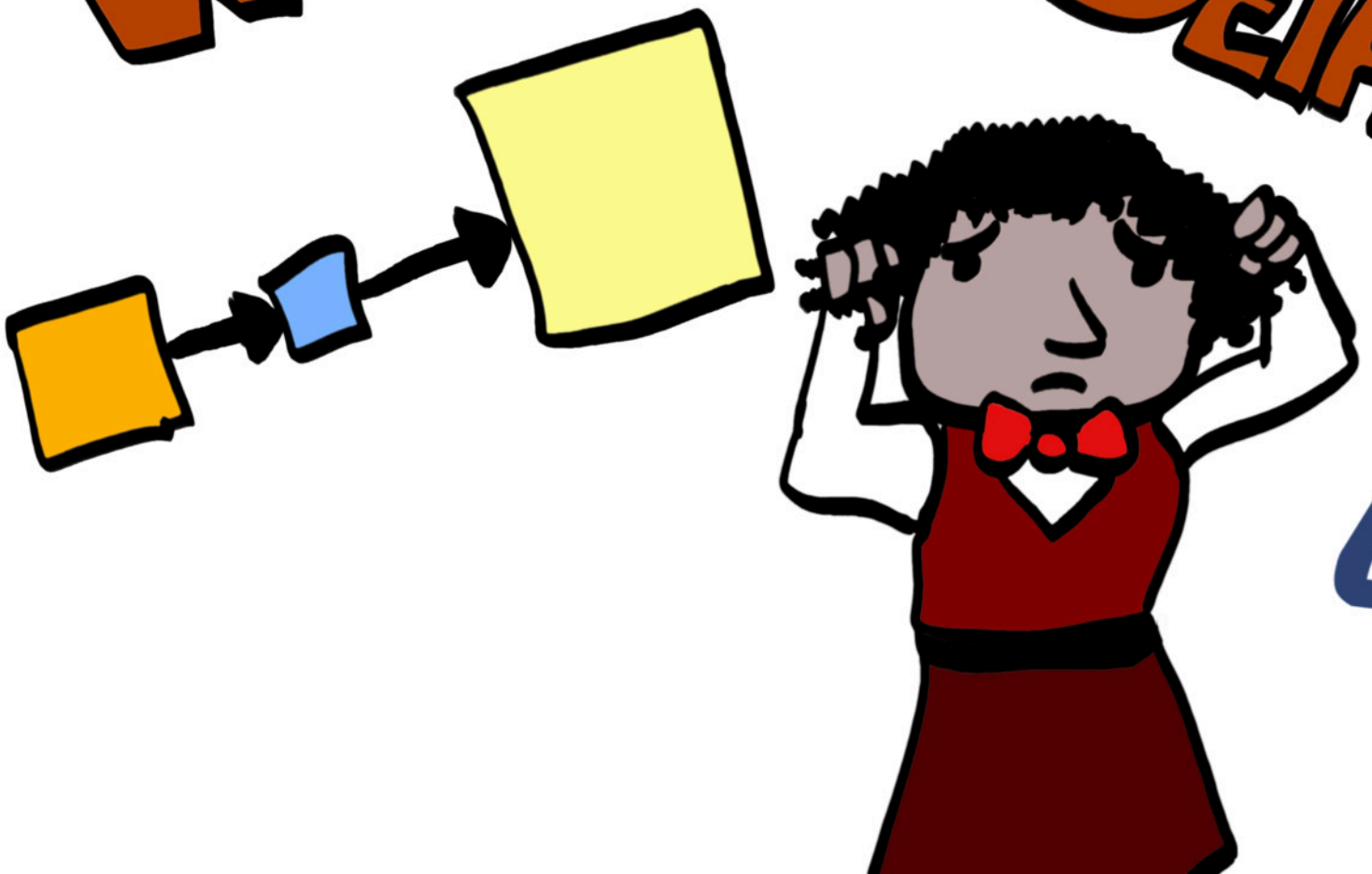

TASK
TASK
TASK


TASK
TASK
TASK


TASK
TASK
TASK

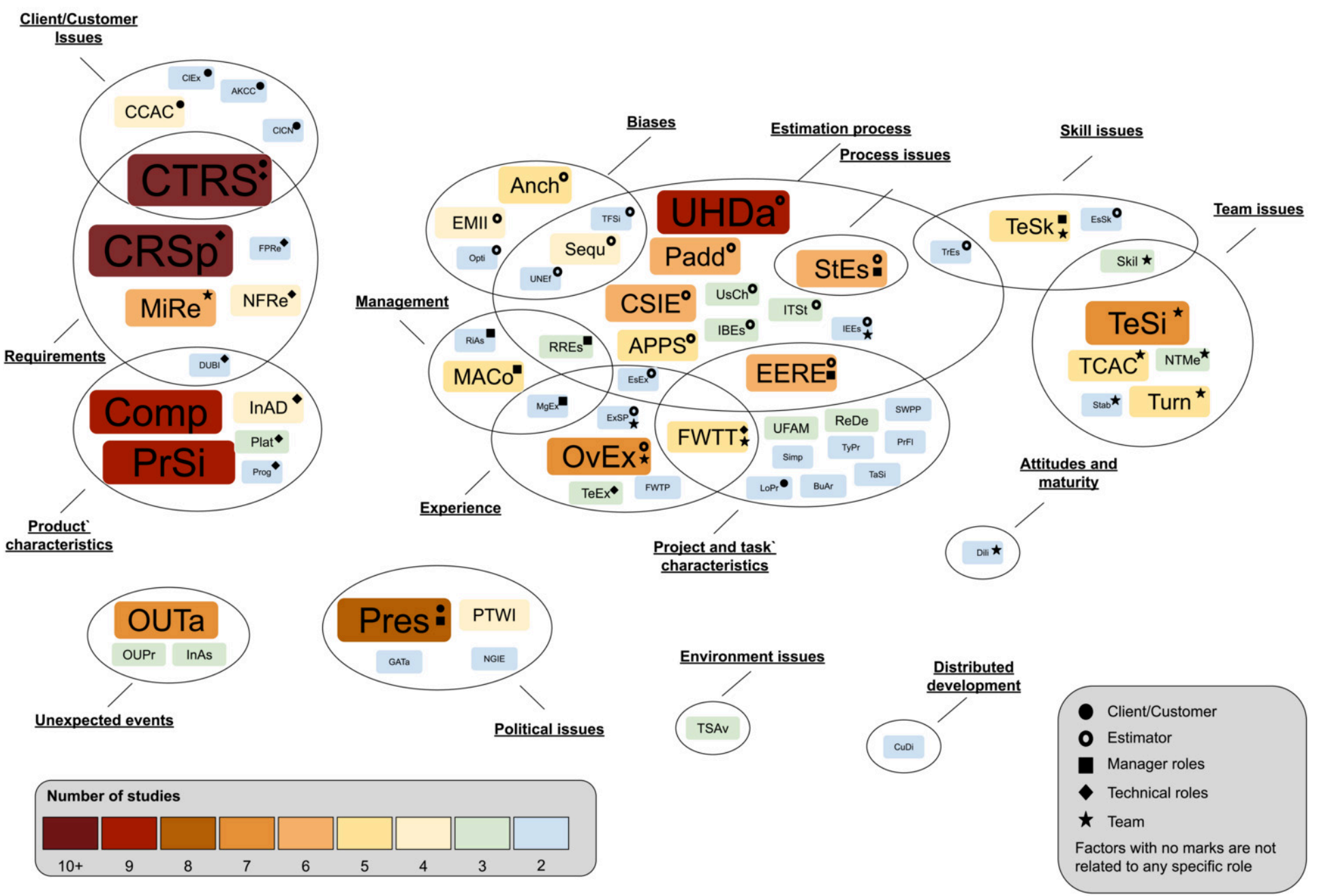
WORDING

DETAILS



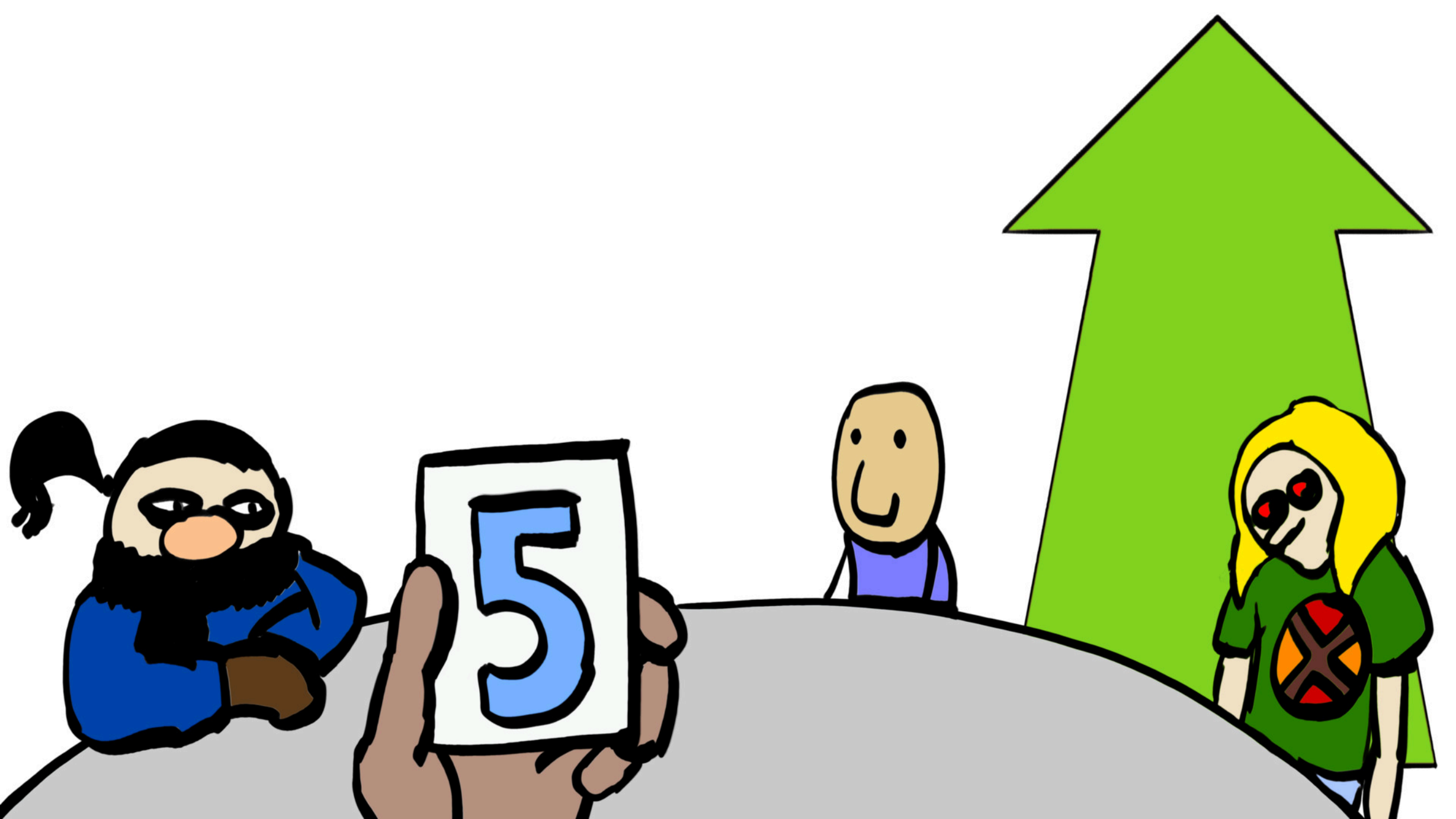
4/69

AKCC	Availability of knowledgeable/competent clients
Anch	Anchoring effect
APPS	Antecipation of project' participants' skills
BuAr	Business area
CCAC	Collaboration and communication
CICN	Clarity of client's needs
CIEx	Client's expectations
Comp	Complexity
CRSp	Clear requirements specification
CSIE	Combination strategy of individual estimates
CTRS	Changes to requirements or scope
CuDi	Cultural differences
Dili	Diligence
DUBI	Dependencies between user stories/backlog items
EERE	Enough effort and resources spent on estimation
EMII	The effect of more and/or irrelevant information
EsEx	Estimation experience
EsSk	Estimation skills
ExSP	Experience with similar/previous projects/tasks
FPRe	Familiar problem or requirements
FWTP	Familiarity with the product
FWTT	Familiarity with the technology
GATa	Goals and targets
IBEs	Informal basis for estimating
IEEs	Impact of early estimates
InAD	Integration and dependencies
InAs	Incorrect assumptions
ITSt	Involvement of technical staff
LoPr	Longer projects
MACo	Monitoring and control
MgEx	Manager experience



MiRe	Misunderstanding of requirements
NFRe	Non-functional requirements
NGIE	Negotiations games in estimates
NTMe	New team members
Opti	Optimism
OUPr	Occurrence of unforeseen problems
OUTa	Overlooked and unplanned tasks
OvEx	Overall experience
Padd	Padding
Plat	Platform
Pres	Pressure
PrFl	Project flexibility
Prog	Programming language
PrSi	Size
PTWI	Price-to-win issues
ReDe	Resources dependencies
RiAs	Risk assessment
RREs	Reestimation and revision of estimates
Sequ	Sequence effects
Simp	Simplicity
Skil	Skill
Stab	Stability
StEs	Standards in estimation
SWPP	Similarity with previous tasks/projects
TaSi	Task size
TCAC	Collaboration and communication
TeEx	Technical experience
TeSi	Size
TeSk	Technical skill
TFSi	Time frame size
TrEs	Training in estimation
TSAv	Tool support and avaliability
Turn	Turnover
TyPr	Type of project
UFAM	Use of flexible/agile development model
UHDa	Use of historical data
UnEf	Unit effects (Lower granularity time unit)
UsCh	Use of checklists

Figure 7 - The SEXTAMT


5

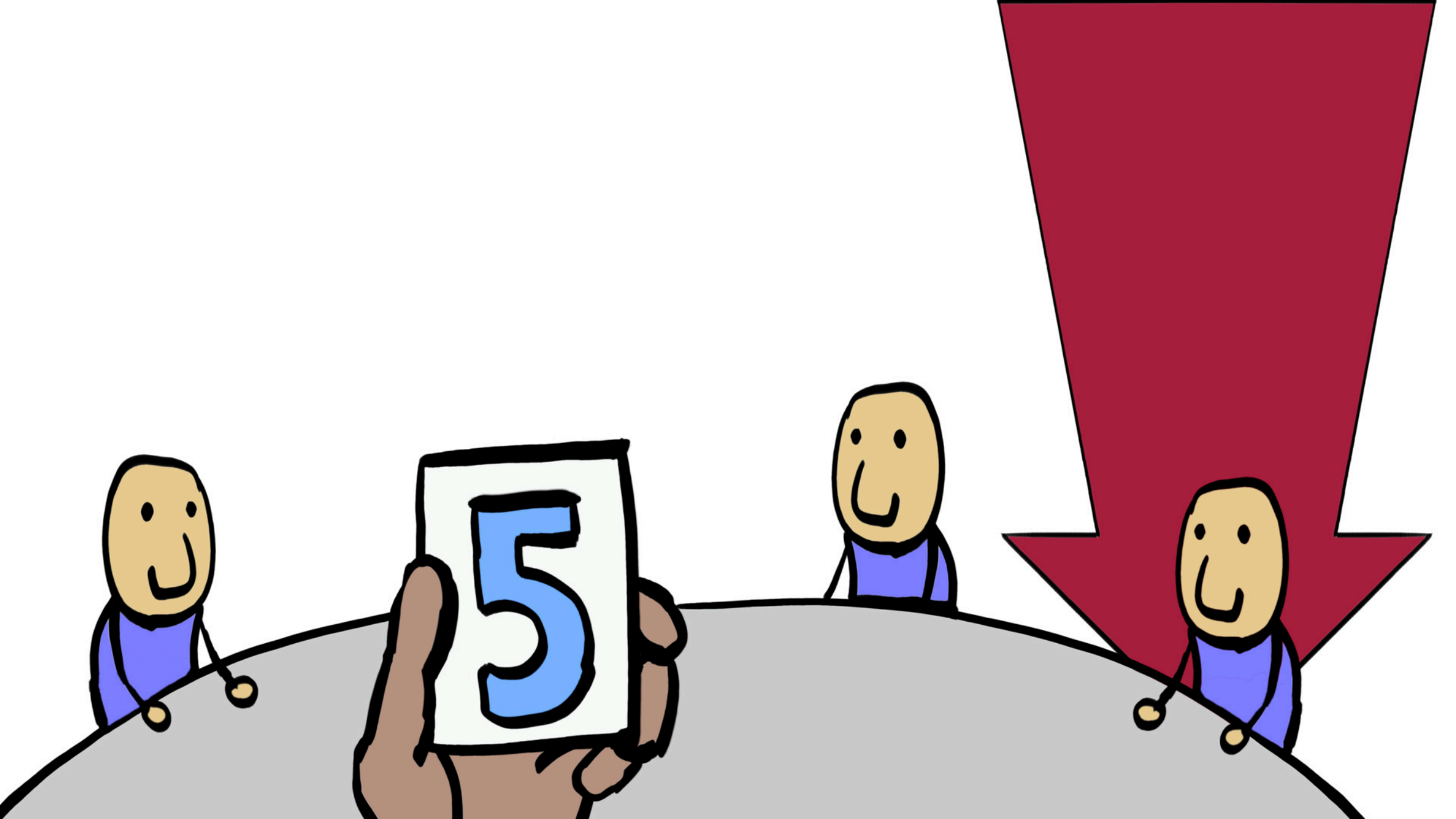
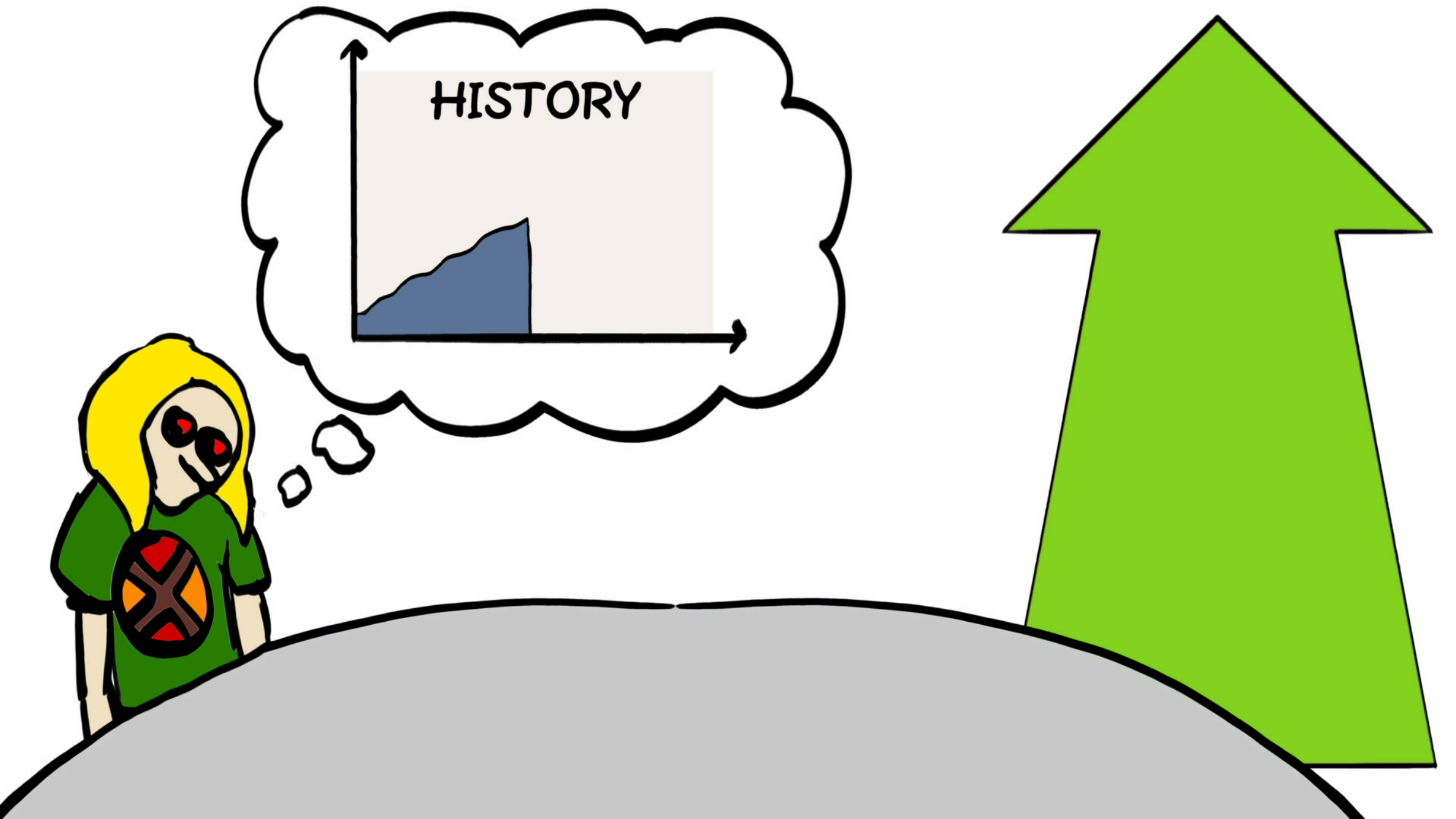

5

Table A.2: Infoway checklist V3 (prioritized and classified after dynamic validation)

No.	Checklist factor	Scale
	Mandatory Part	
1	What is the type of the task?	Implementation, testing, research
2	The task is implementing a Functional Requirement (FR) or Non-Functional Requirement (NFR) or both?	Only an NFR, FR with major NFR constraints, FR with minor NFR constraints, only FR
3	Rank team's technical competence/skills required to implement and test this task	Very good, good, average, below average, lacking
4	How similar is the task to the previously developed tasks?	To a great extent, somewhat, very little, not at all
5	Rank team's domain knowledge required to implement this task	Very good, good, average, below average, lacking
6	How clear is your understanding of the task?	Very good, good, average, below average, need to talk to customer before moving on
7	The implementation of the task requires understanding and/or changing legacy code	Yes, not sure, no
8	Does the task involve communication between multiple (sub) systems?	Yes, not sure, no
	Optional Part	
9	Does the task require an architectural change?	Yes, not sure, no
10	Does the team has new member(s)?	Yes, no
11	Will the team or some member(s) be working on other products/projects/tasks in parallel?	Yes, not sure, no
12	Does the task involve accessing and/or modifying several different elements/tables in the persistence/DB layer?	Yes, not sure, no
13	Is the task size suitable to fit in the planned sprint?	Yes, not sure, no
14	Rank team's recent productivity	Very good, good, average, below average, lacking
Based on the above characterizations, most likely effort estimate for this task is: ——— hours		


HISTORY


TASK
TASK
TASK
TASK


TASK
TASK
TASK

I CONTEXT

II STORY POINTS

III NOESTIMATES TECHNIQUES

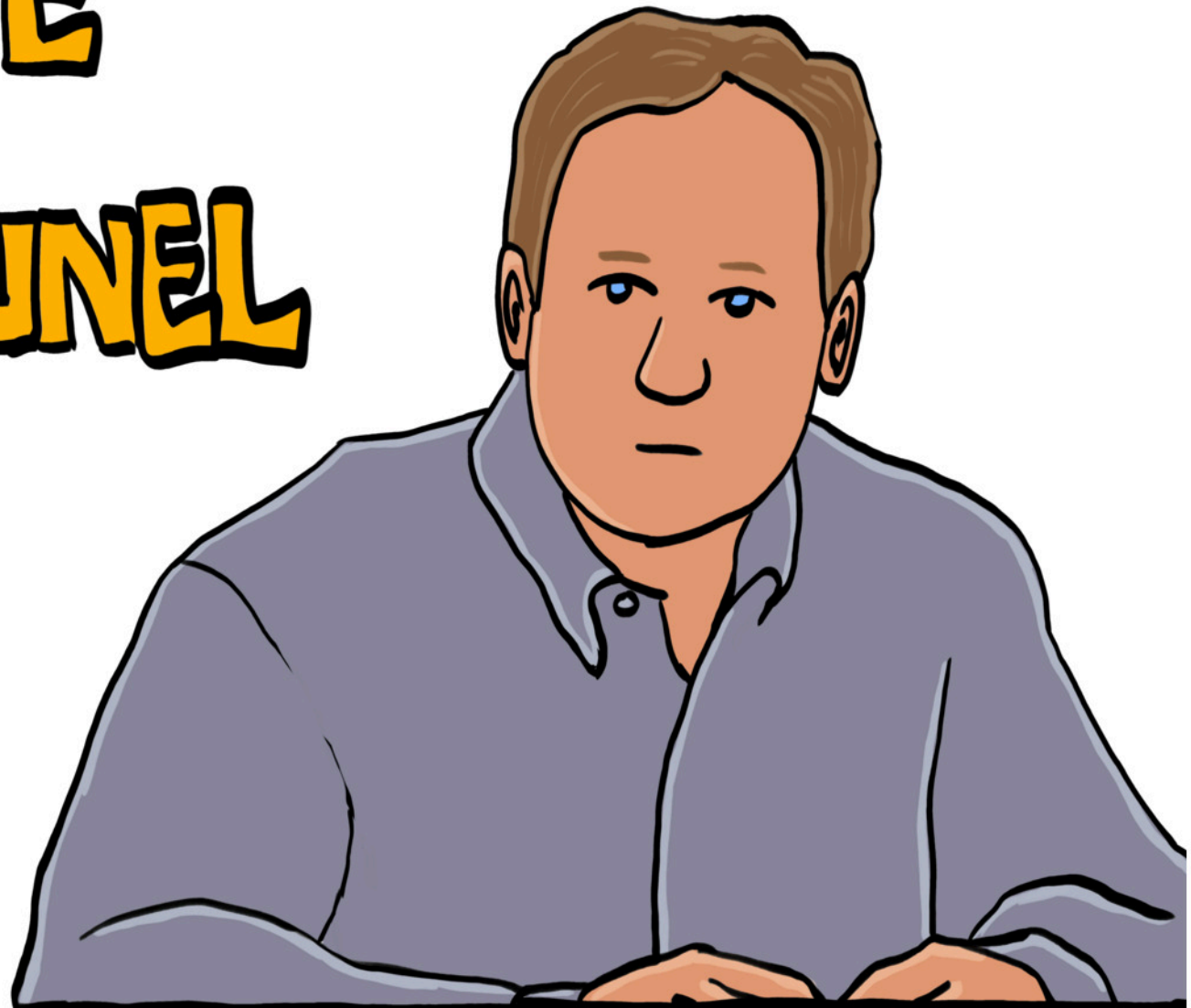
IV

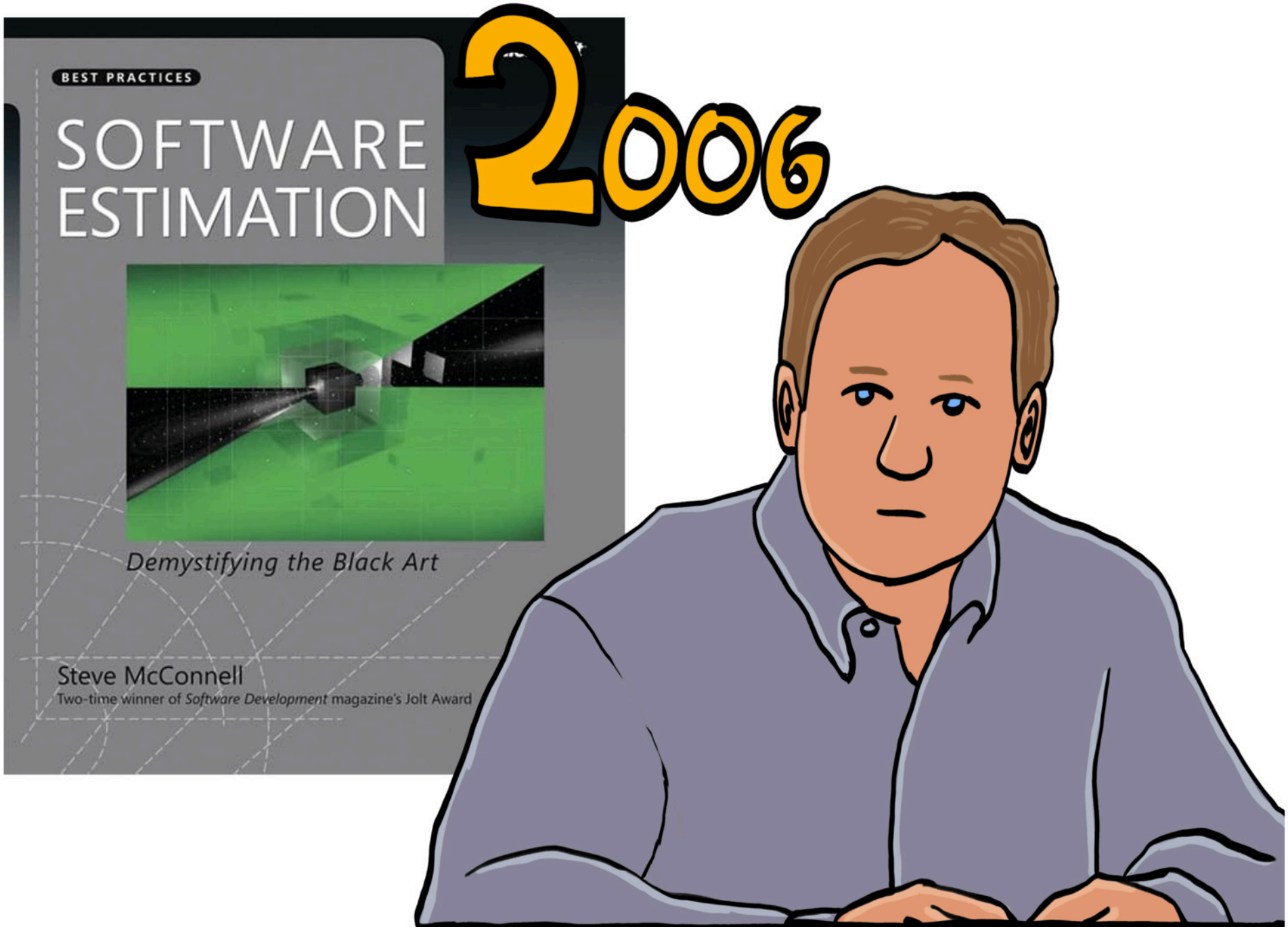
V

VI

VII

Steve McConnel

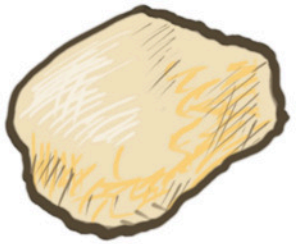


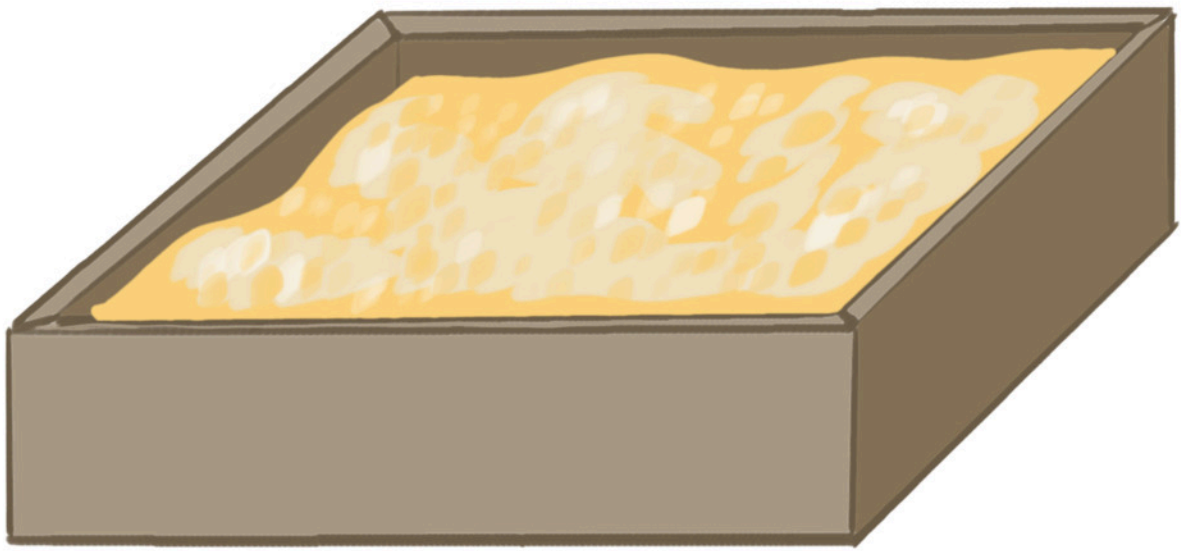

BEST PRACTICES
SOFTWARE ESTIMATION
Demystifying the Black Art
Steve McConnell
Two-time winner of Software Development magazine's Jolt Award
2006

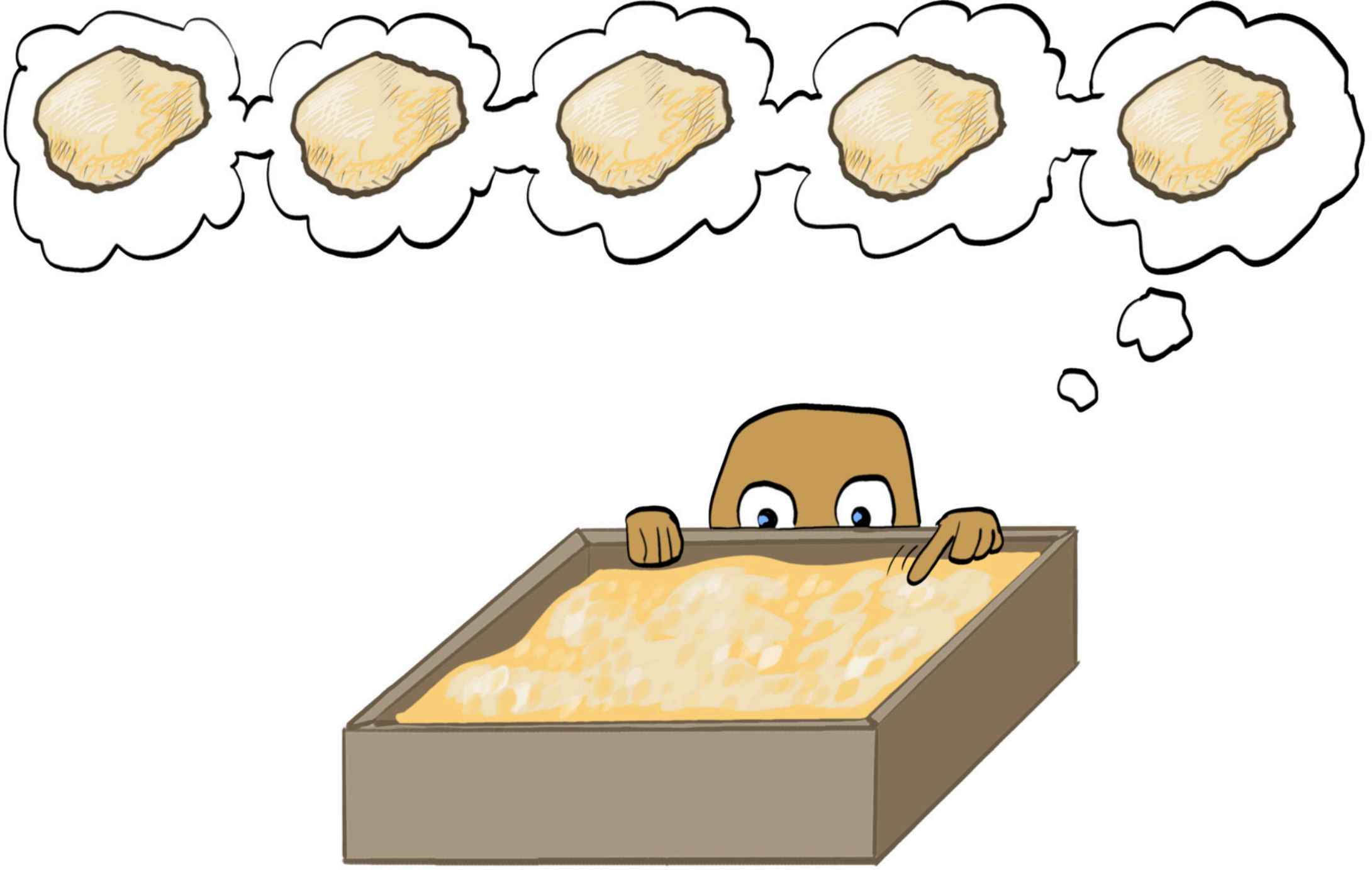
SOFTWARE ESTIMATION : DEMYSTIFYING THE BLACK ART / S. MCCONNELL.

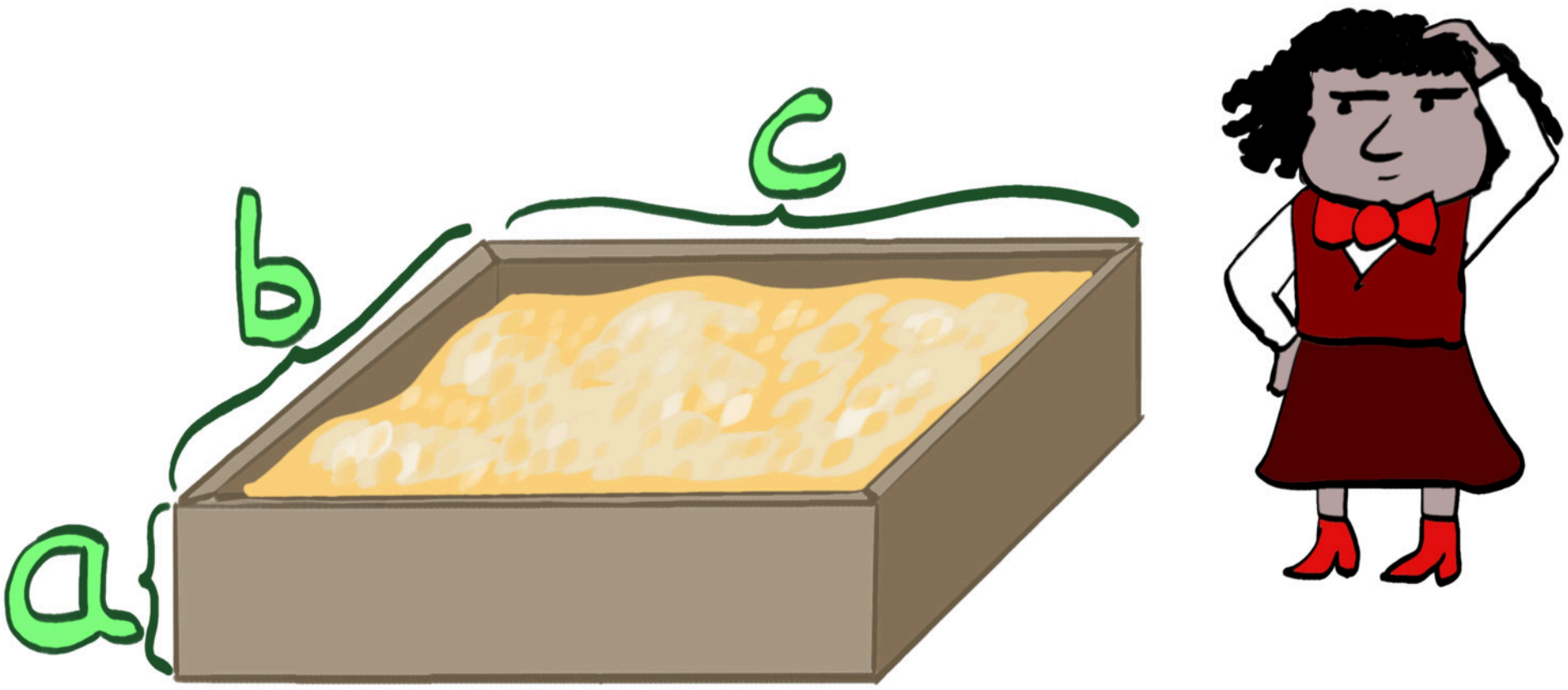
Mcconnell, Steve

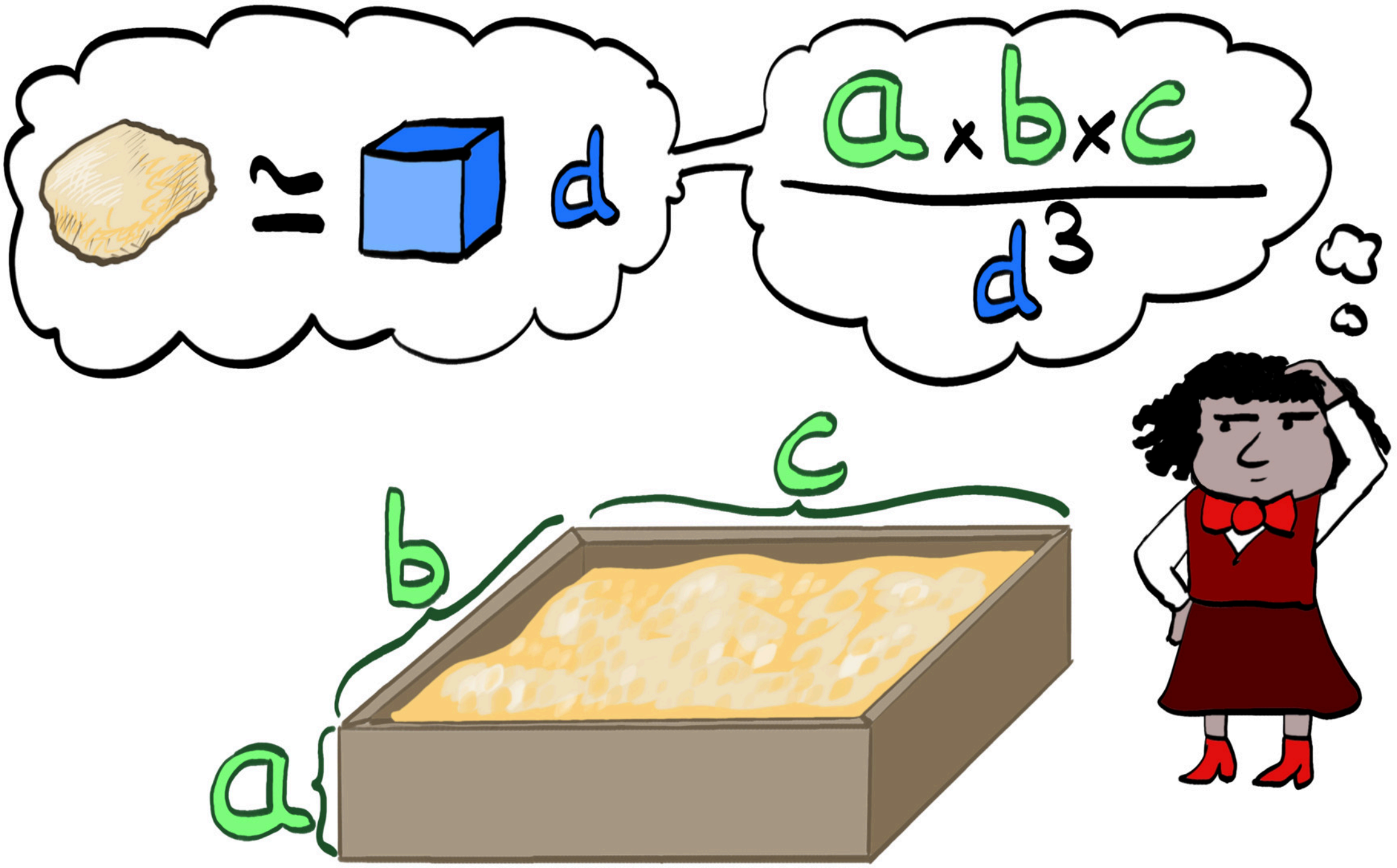
2006

n x  ?






c
b
a

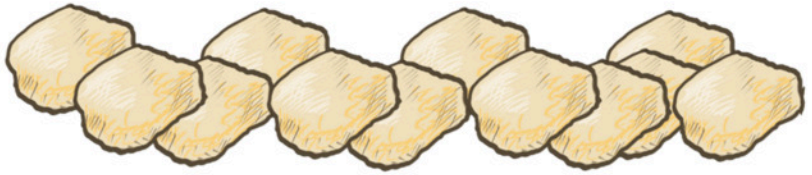

≃
d
$\frac{a \times b \times c}{d^3}$
c
b
a


Hum...

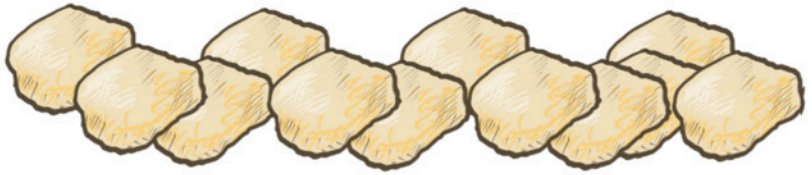

two million
I guess...


two million
I guess...
Client signed
for one


two million I guess...
Client signed for one
Ok, one million then



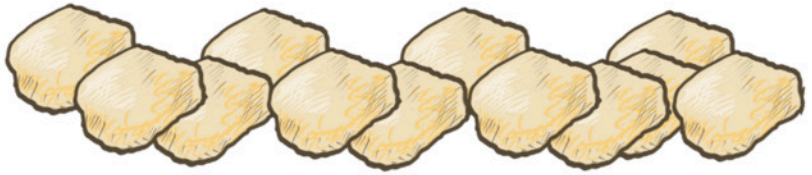
1



1

2

$$\frac{a \times b \times c}{\square^3}$$



1

2

$$\frac{a \times b \times c}{\square^3}$$

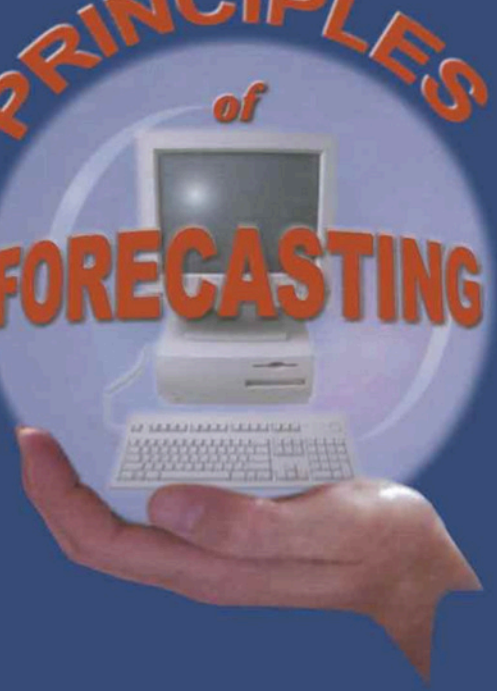
3

J. Scott Armstrong

About

570	472,844	40,117
Publications	Reads ⓘ	Citations

PRINCIPLES of FORECASTING



A Handbook for Researchers and Practitioners

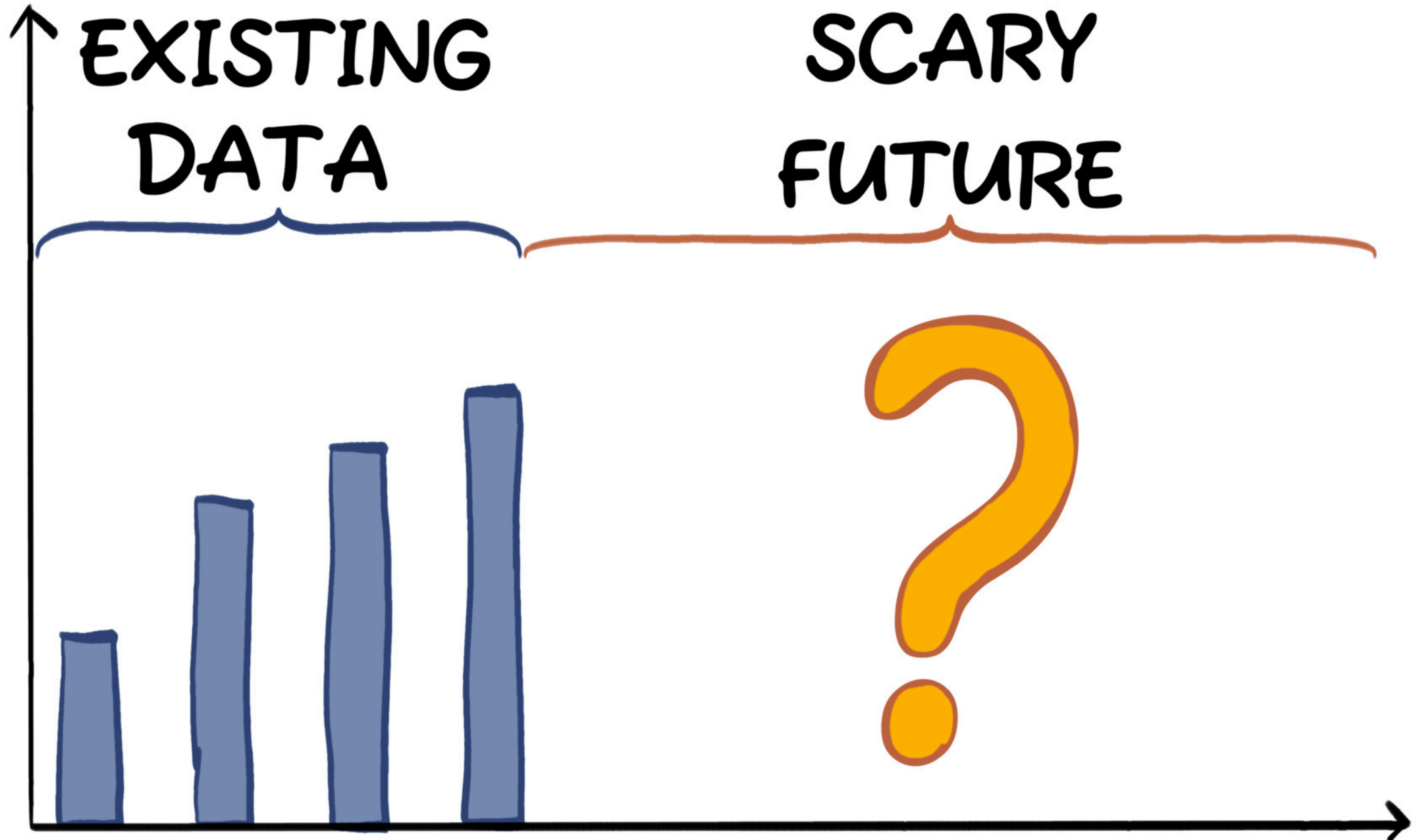
edited by

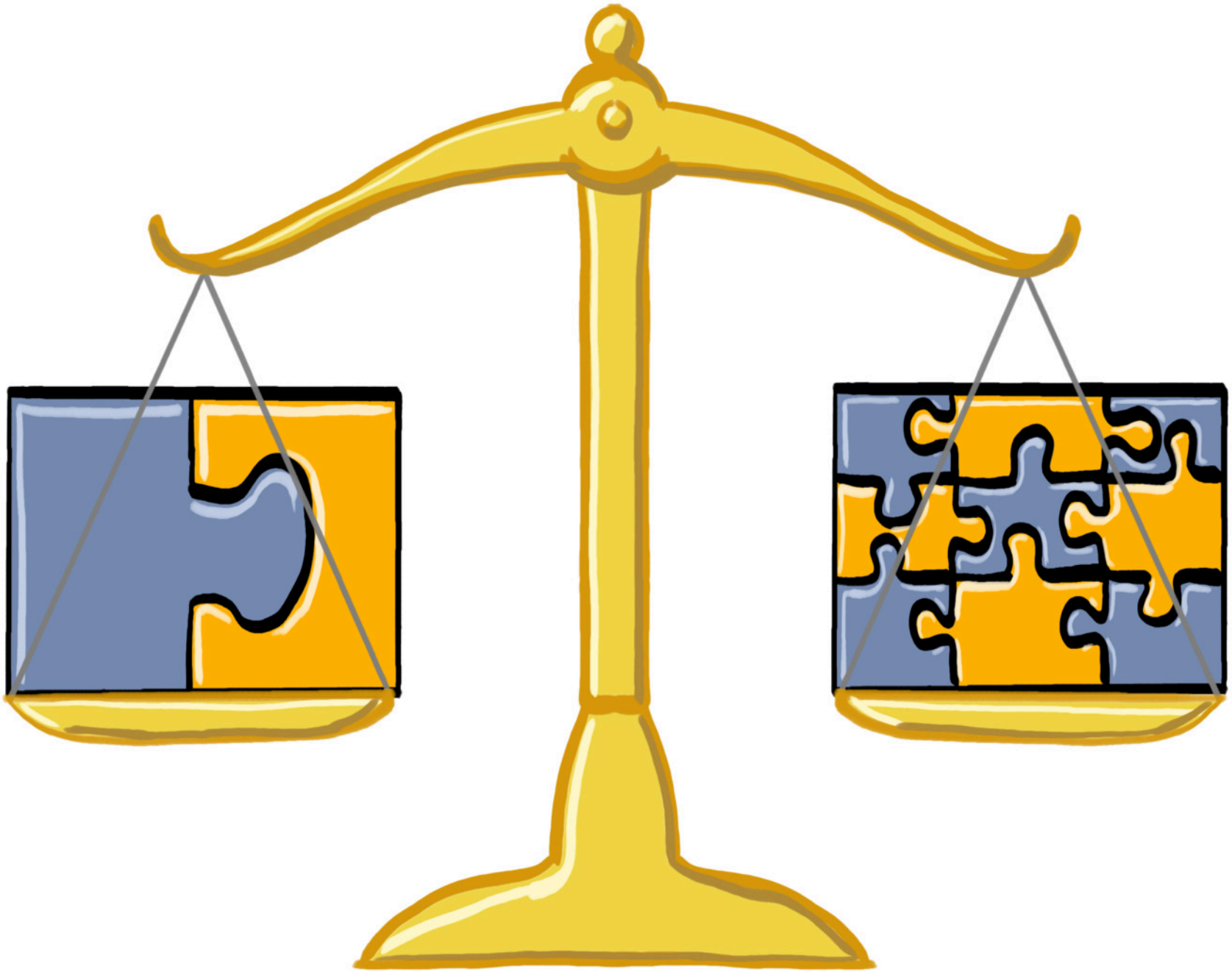
J. Scott Armstrong

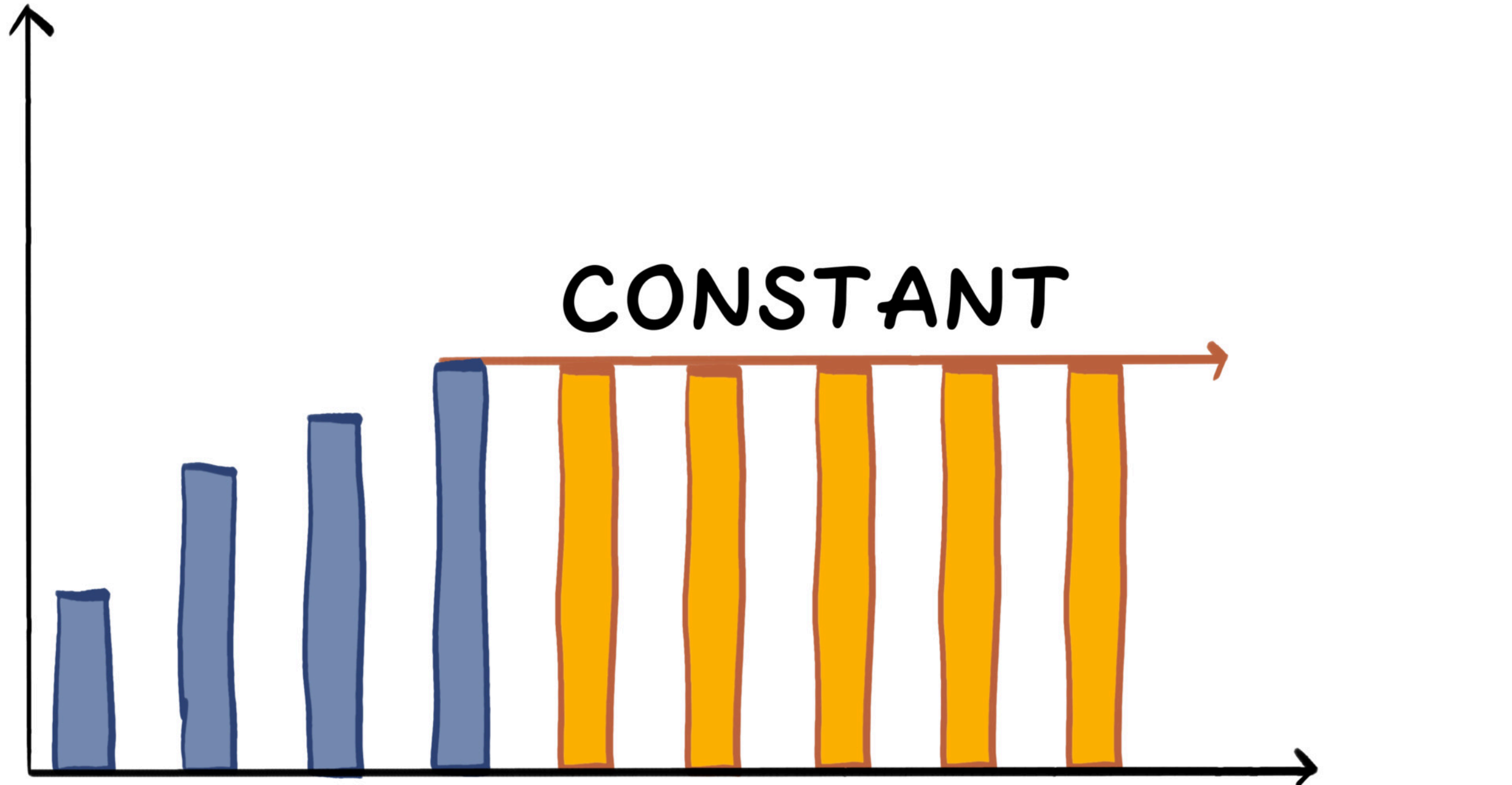
SIMPLE VERSUS COMPLEX FORECASTING: THE EVIDENCE (SIMPLE-FORECASTING.COM)

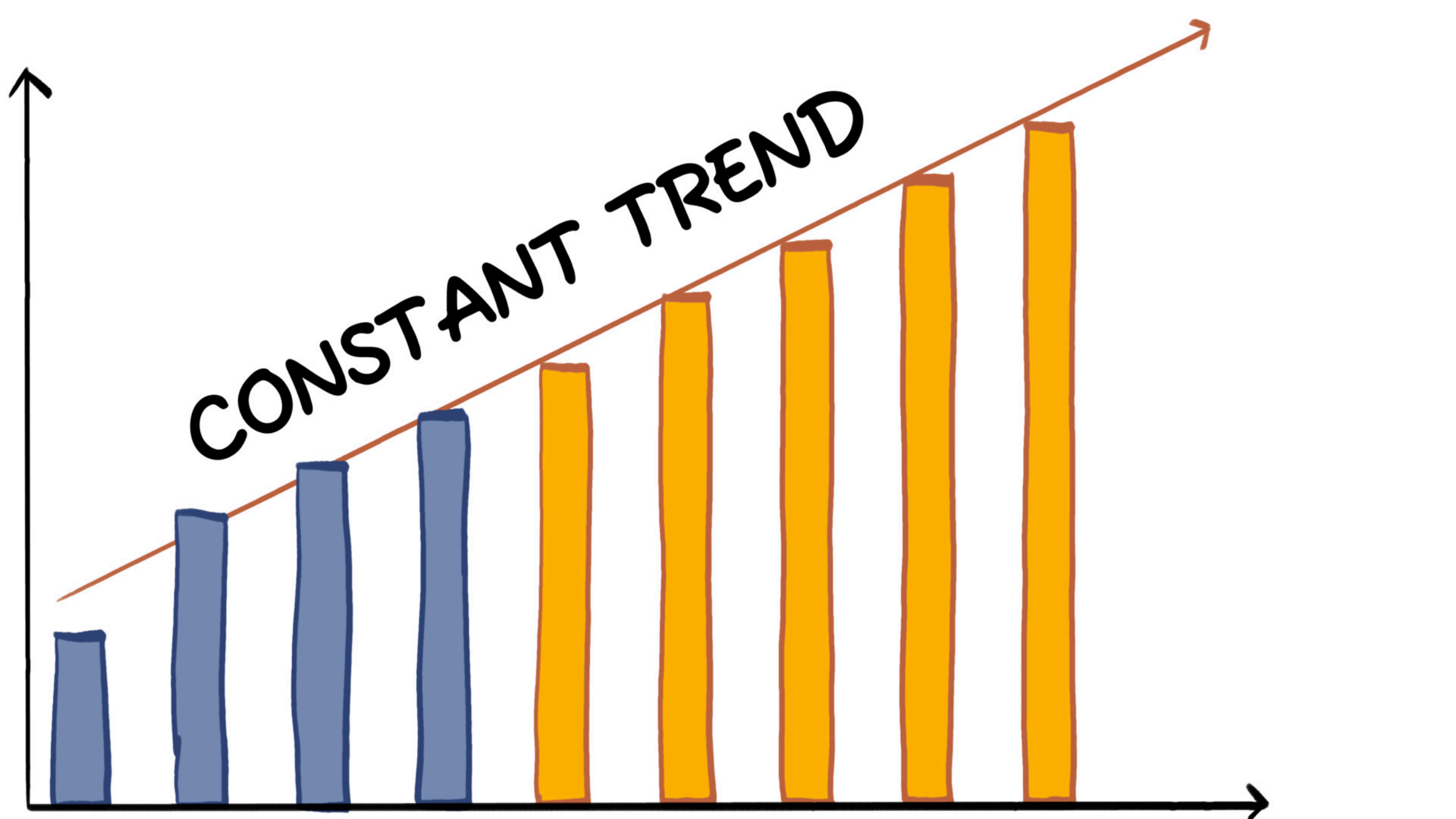
Green, Kesten and Armstrong, J.

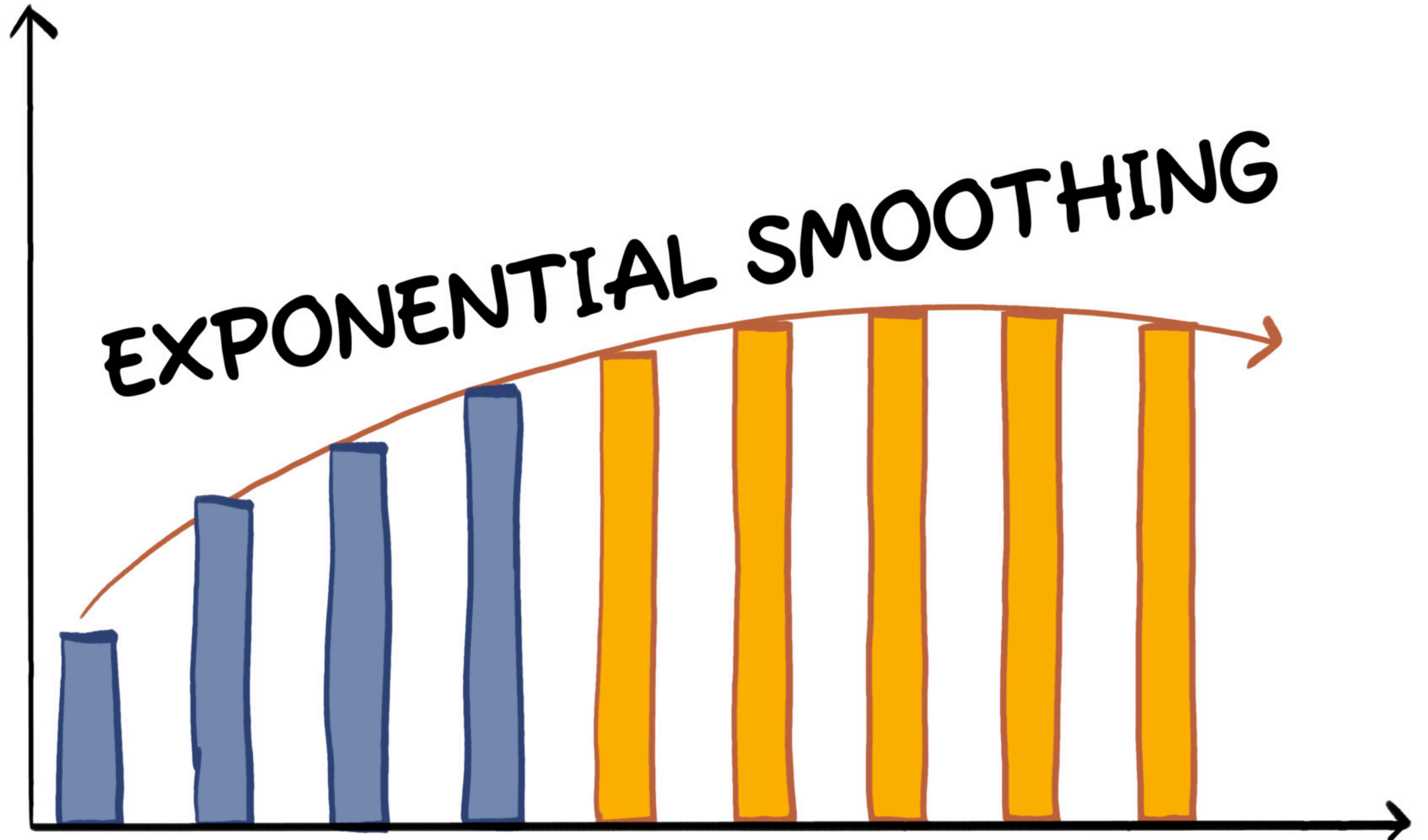
2015

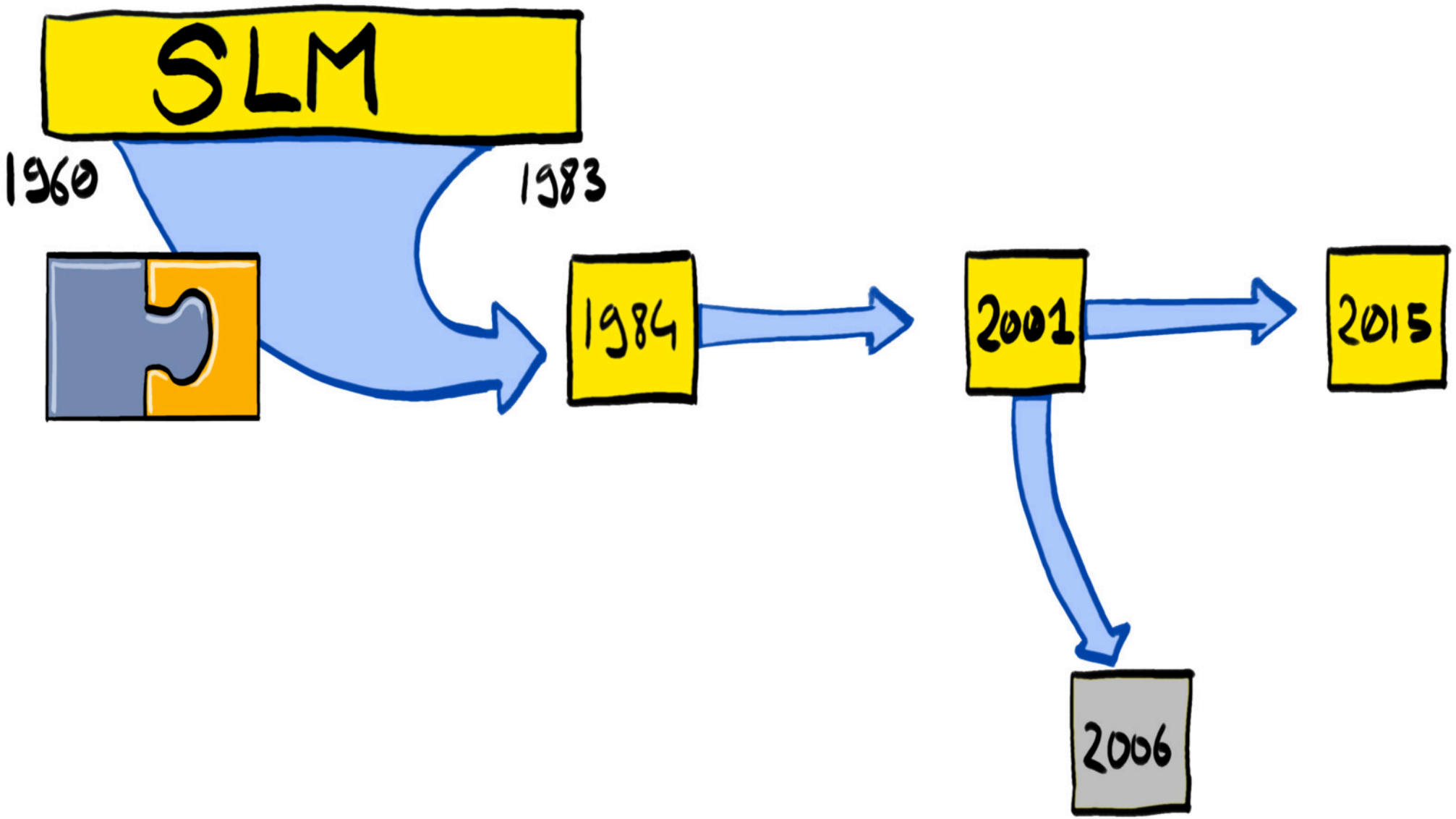

EXISTING DATA
SCARY FUTURE
?




CONSTANT


CONSTANT TREND


EXPONENTIAL SMOOTHING

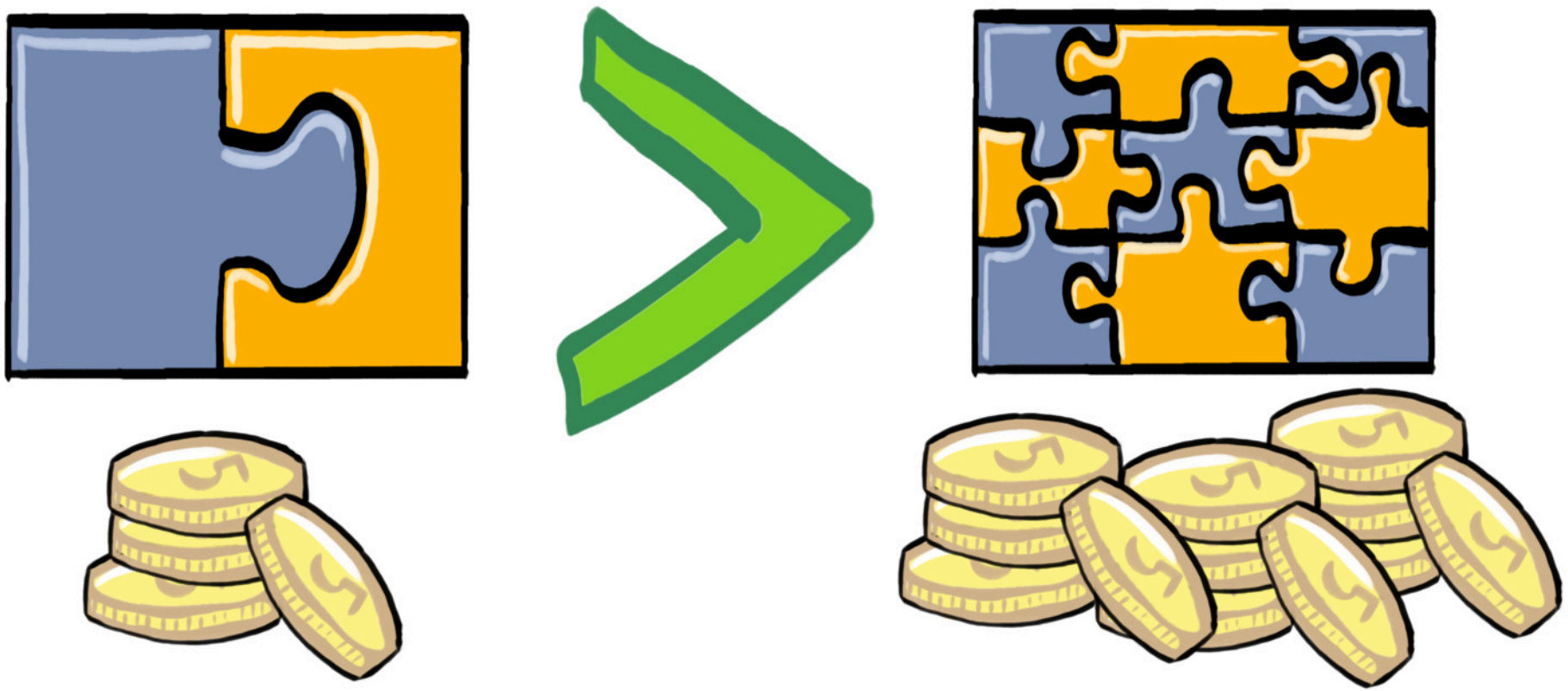

SLM
1960
1983
1984
2001
2015
2006

Table

Summary of evidence on accuracy of forecasts from complex vs. simple methods

	-------- Number of Comparisons --------				
Method type	**Total papers**	**Total compar-isons**	**Simple better or similar**	**Effect size**	**Error increase vs simple (%)**
Judgmental	4	4	4	4	28.2
Extrapolative	17	62	51	12	27.5
Causal	8	23	19	5	25.3
Combined	3	8	7	4	23.9
All method types	**32**	**97**	**81**	**25**	
Weighted average*					**26.7**

*Weighted by total papers



PRINCIPLES OF FORECASTING: A HANDBOOK FOR RESEARCHERS AND

Armstrong, J.

2001 - Technological Forecasting and Social Change - TECHNOL FORECAST SOC CHANGE

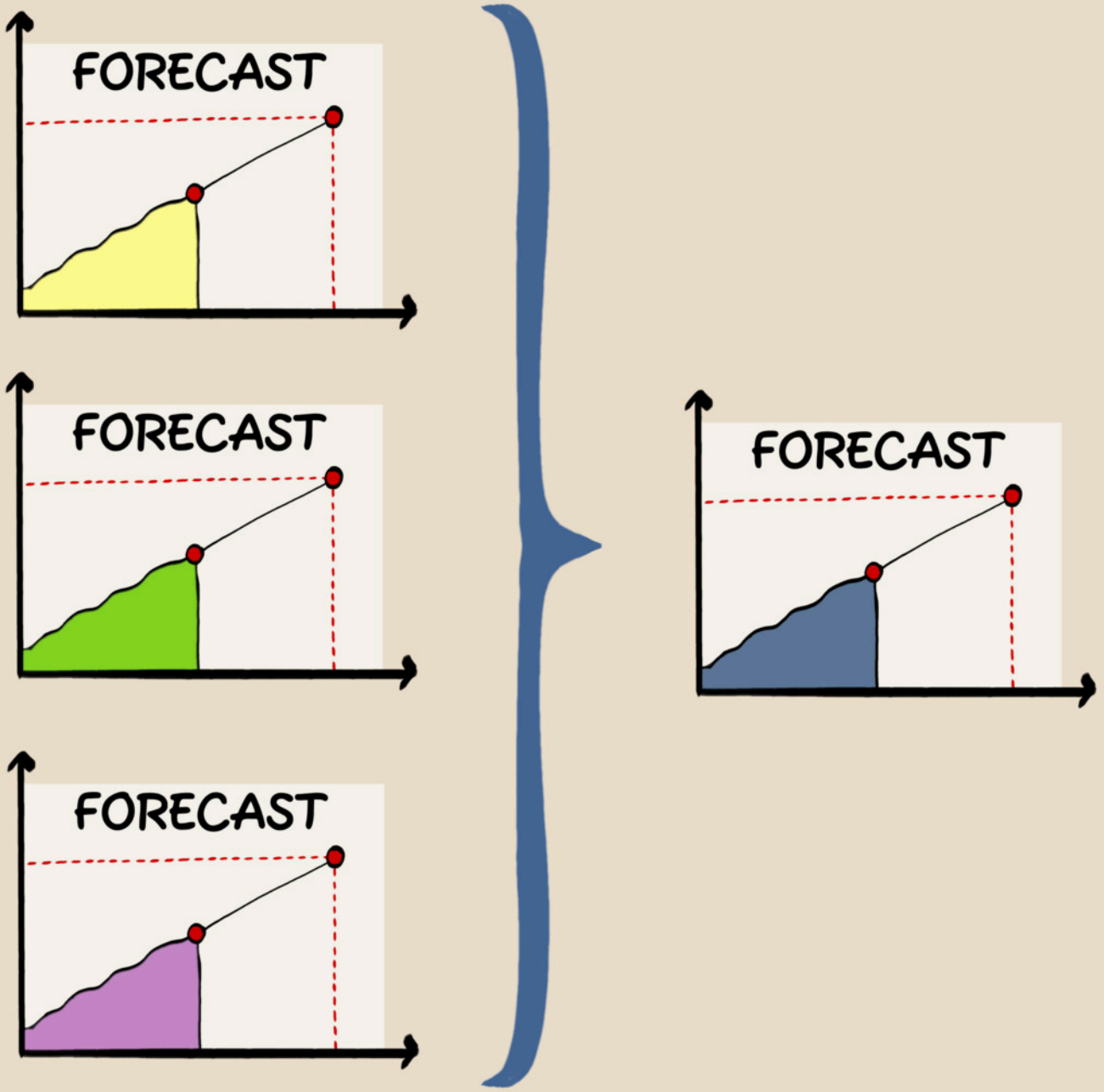
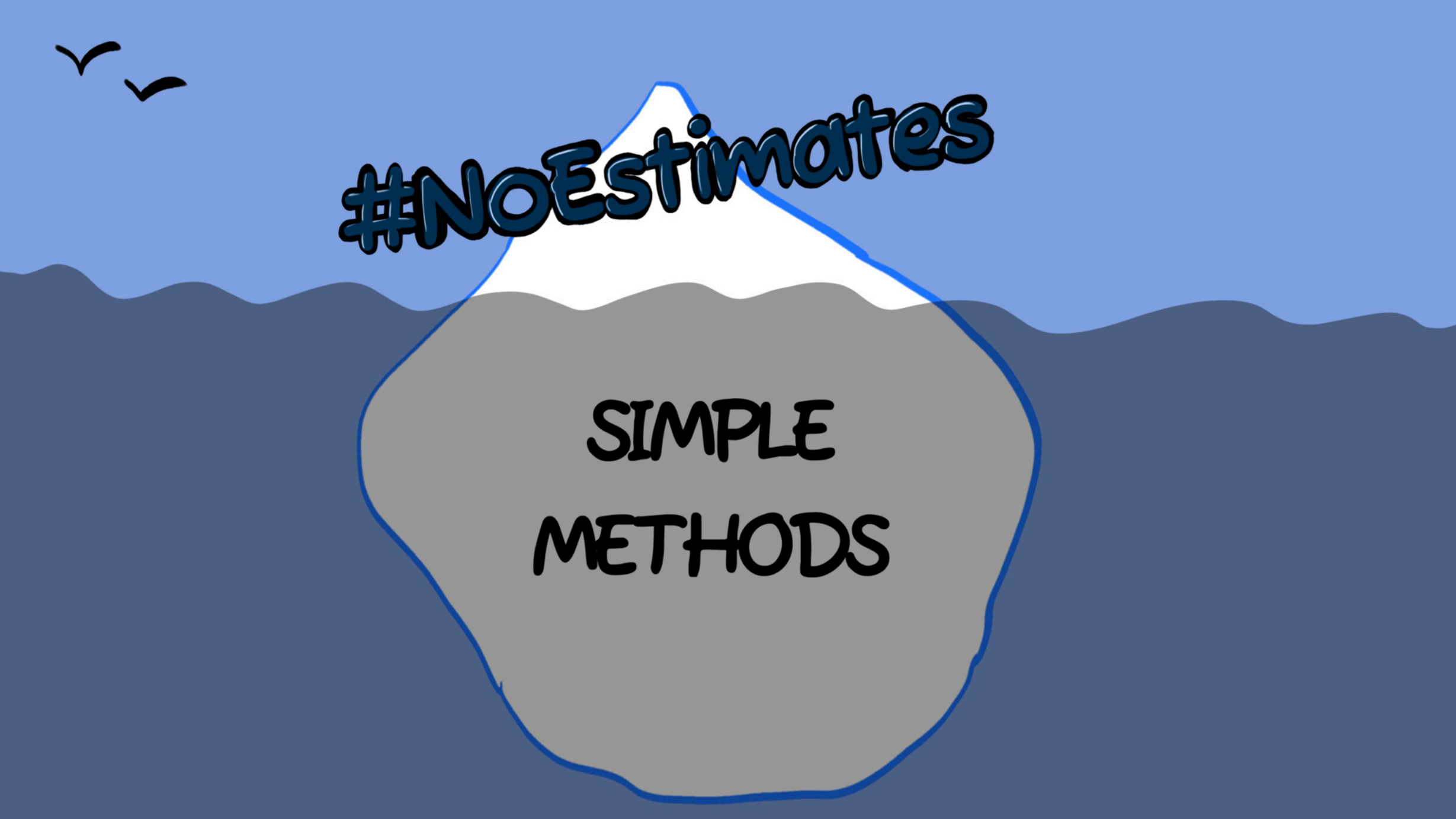

FORECAST
FORECAST
FORECAST
FORECAST

Exhibit 1. Error reductions from combining ex ante forecasts

Study	Methods	Compo-nents	Criterion	Data	Situation	Validation forecasts	Forecast horizon	Percent error reduction
Levine (1960)	intentions	2	MAPE	annual	capital expenditures	6	1	18.0
Okun (1960)	"	2	"	"	housing starts	6	1	7.0
Landefeld & Seskin (1986)	"	2	MAE	"	plant & equipment	11	1	20.0
Armstrong et al. (2000)	"	4	RAE	"	consumer products	65	varied	5.5
Winkler & Poses (1993)	expert	4	Brier	cross-section	survival of patients	231	varied	12.2
Thorndike (1938)	"	4 to 6	% wrong	"	knowledge questions	30	varied	6.6
Makridakis et al. (1993)	"	5	MAPE	monthly	economic time series	322	1 – 14	19.0
Richards & Fraser (1977)	"	5	"	annual	company earnings	213	1	8.1
Batchelor & Dua (1995)	"	10	MSE	"	macroeconomic	40	1	16.4
Kaplan et al. (1950)	"	26	% wrong	cross-section	technology events	16	varied	13.0
Zarnowitz (1984)	"	79	RMSE	quarterly	macroeconomic	288	1	10.0
Sanders & Ritzman (1989)	extrapolation	3	MAPE	daily	public warehouse	260	1	15.1
Makridakis & Winkler (1983)	"	5	"	monthly	economic time series	617	18	24.2
Makridakis et al. (1993)	"	5	"	"	"	322	1 – 14	4.3
Lobo (1992)	"	5	"	quarterly	company earnings	6,560	1 – 4	13.6
Schnaars (1986)	"	7	"	annual	consumer products	1,412	1 – 5	20.0
Landefeld & Seskin (1986)	econometric	2	MAE	annual	plant & equipment	7	1	21.0
Clemen & Winkler (1986)	"	4	MAD	quarterly	GNP (real & nominal)	45	1 – 4	3.4
Shamseldin et al. (1997)	"	5	MAPE	annual	rainfall runoff	22	1	9.4
Lobo (1992)	expert/extrap	2	MAPE	annual	company earnings	6,560	1 – 4	11.0
Lawrence et al. (1986)	"	3	"	monthly	economic time series	1,224	1 – 18	10.7
Sanders & Ritzman (1989)	"	3	"	daily	public warehouse	260	1	15.5
Lobo & Nair (1990)	"	4	"	annual	company earnings	768	1	6.4
Landefeld & Seskin (1986)	inten-tions/econ	2	MAE	annual	plant & equipment	11	1	11.5
Vandome (1963)	extrap/econ	2	MAPE	quarterly	macroeconomic	20	1	10.1
Armstrong (1985)	"	2	"	annual	photo sales by country	17	6	4.2
Weinberg (1986)	expert/econ	2	"	cross-sectio n	performing arts	15	varied	12.5
Bessler & Brandt (1981)	experrt/ extrap/ econ	3	"	quarterly	cattle & chicken prices	48	1	13.6
Fildes (1991)	"	3	MAE	annual	construction	72	1 & 2	8.0
Brandt & Bessler (1983)	"	6	MAPE	quarterly	hog prices	24	1	23.5
							Unweighted average	12.5


#NoEstimates
SIMPLE
METHODS

I CONTEXT

II STORY POINTS

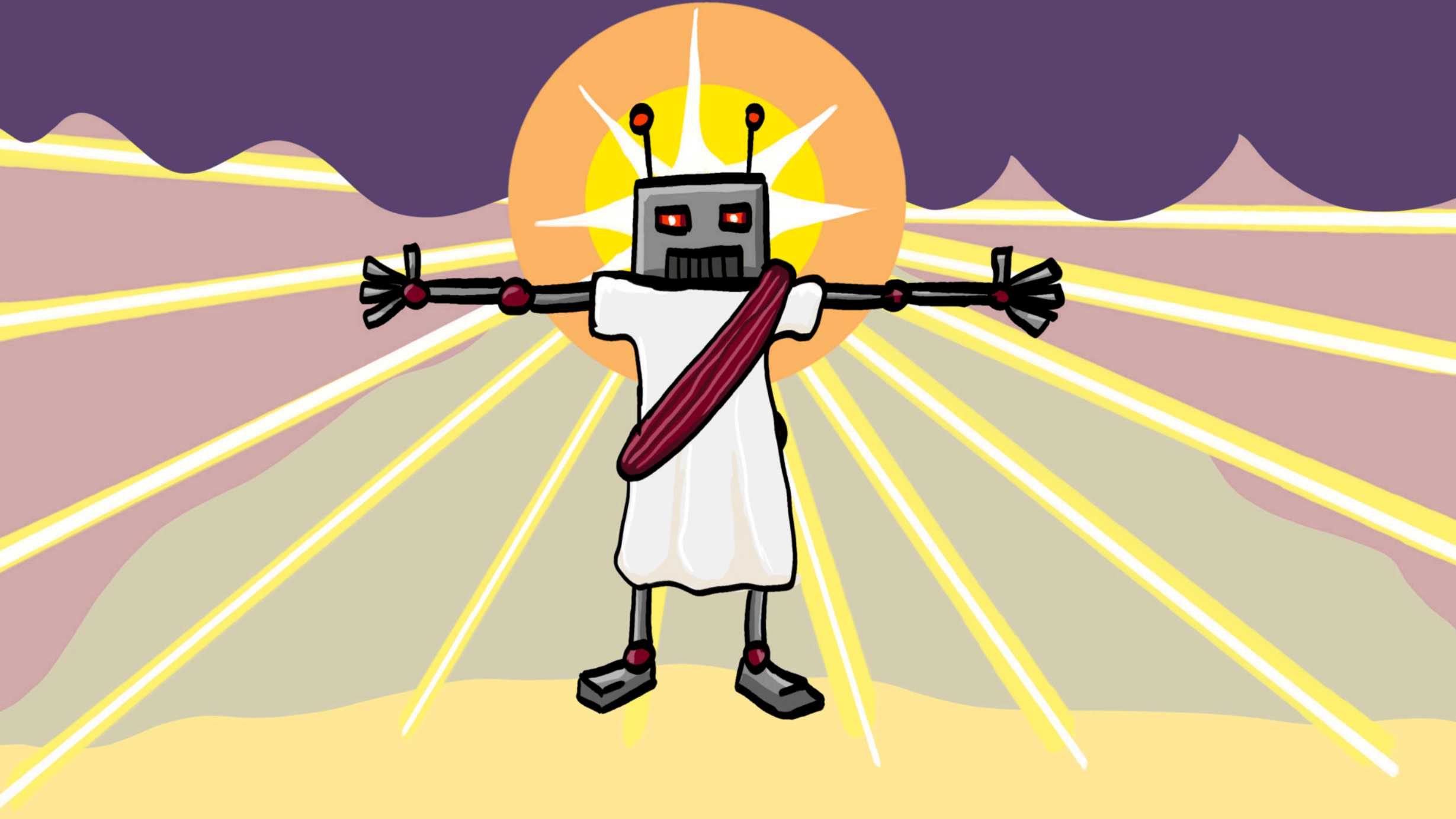
III NOESTIMATES TECHNIQUES

IV WILL AI SAVE US ?

V

VI

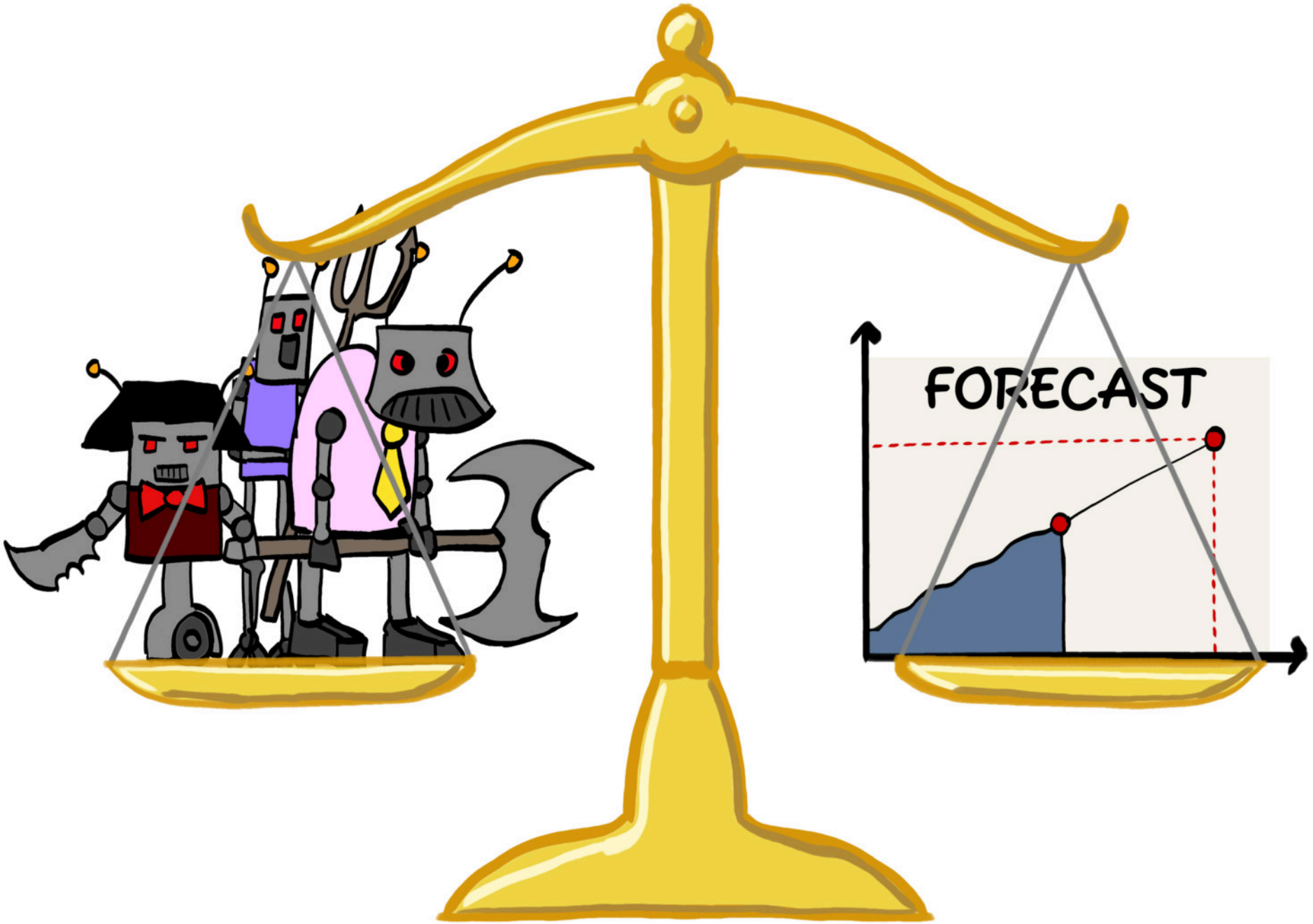
VII

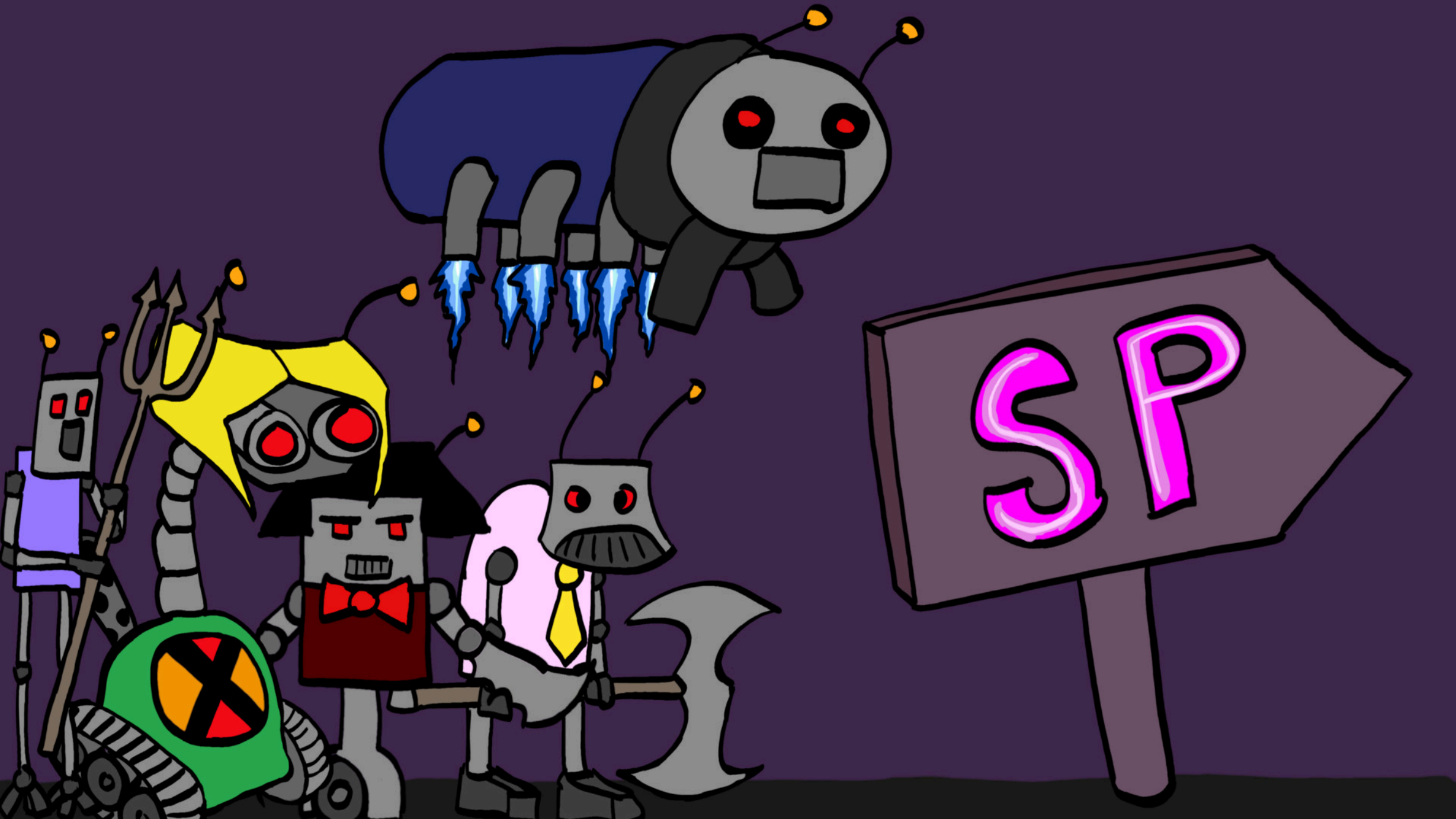


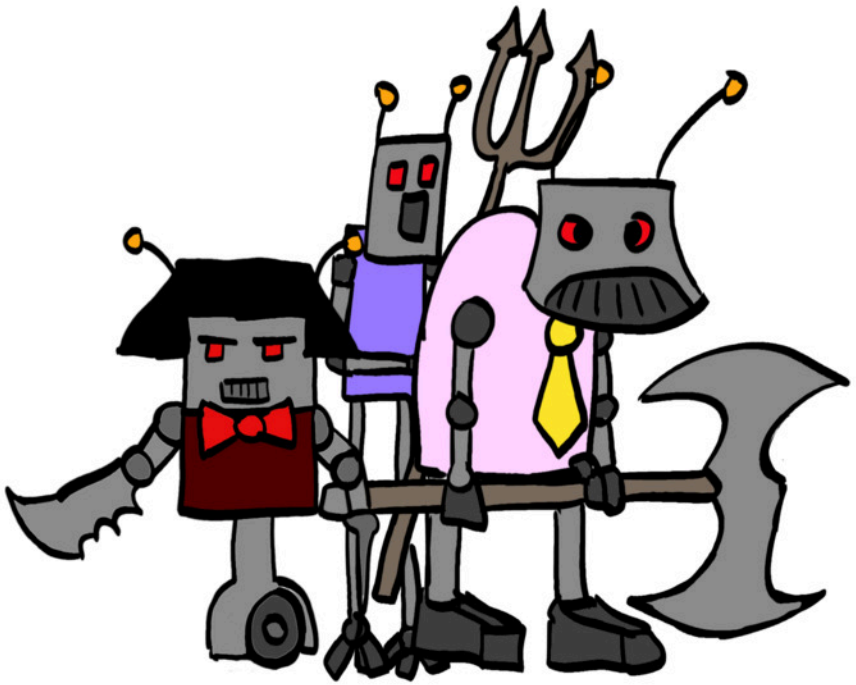
AGILE EFFORT ESTIMATION: HAVE WE SOLVED THE PROBLEM YET? INSIGHTS FROM A REPLICATION STUDY

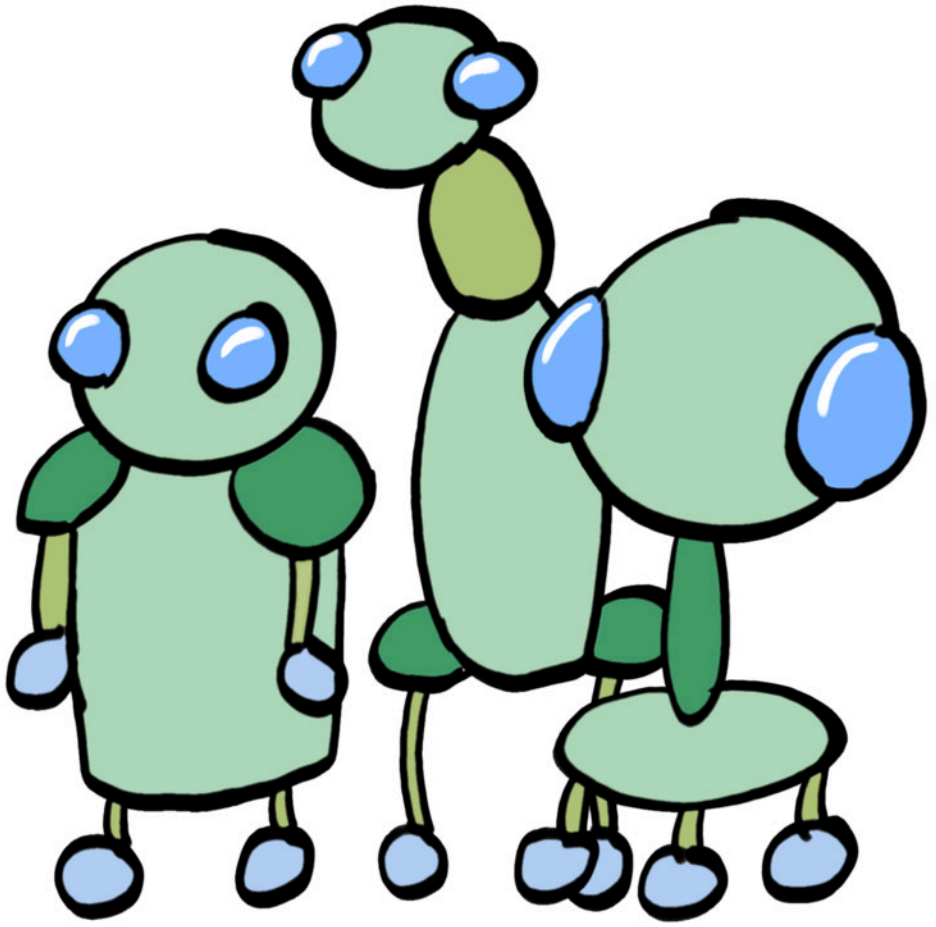
Tawosi, Vali and Moussa, Rebecca and Sarro, Federica

2022 - IEEE Transactions on Software Engineering


FORECAST


SP





I CONTEXT

II STORY POINTS

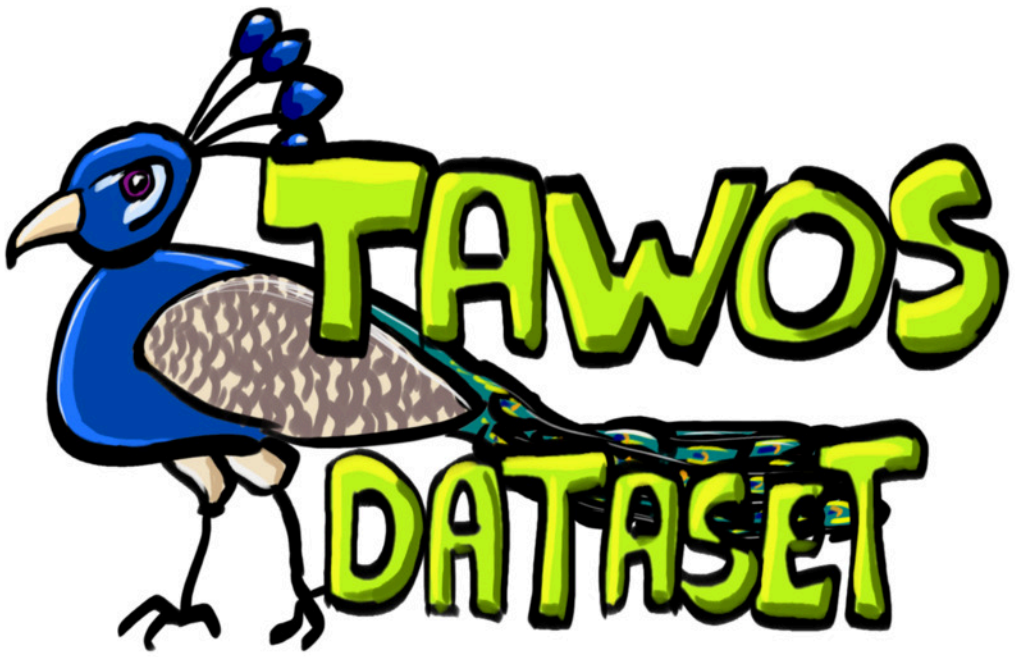
III NOESTIMATES TECHNIQUES

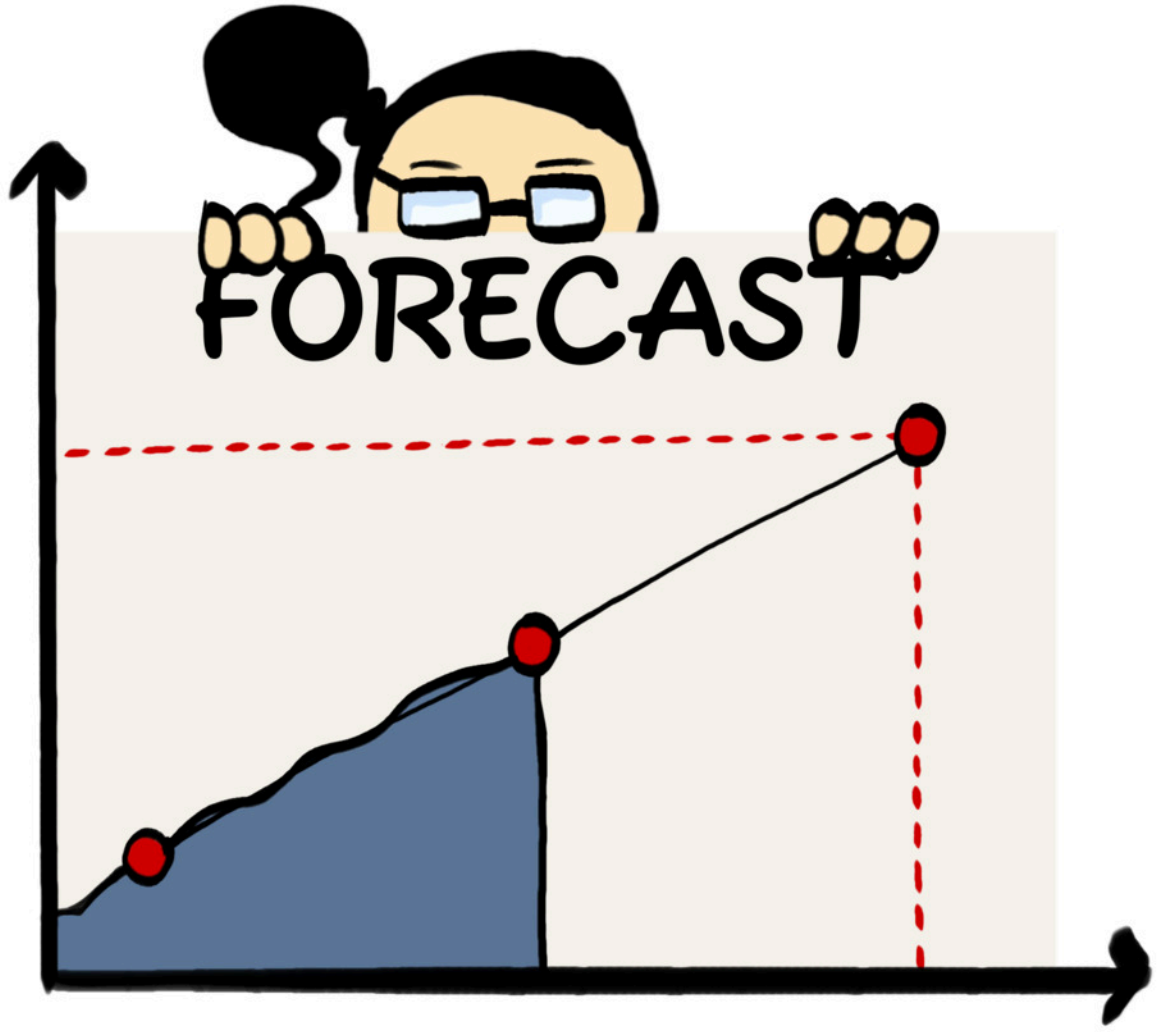
IV WILL AI SAVE US ?

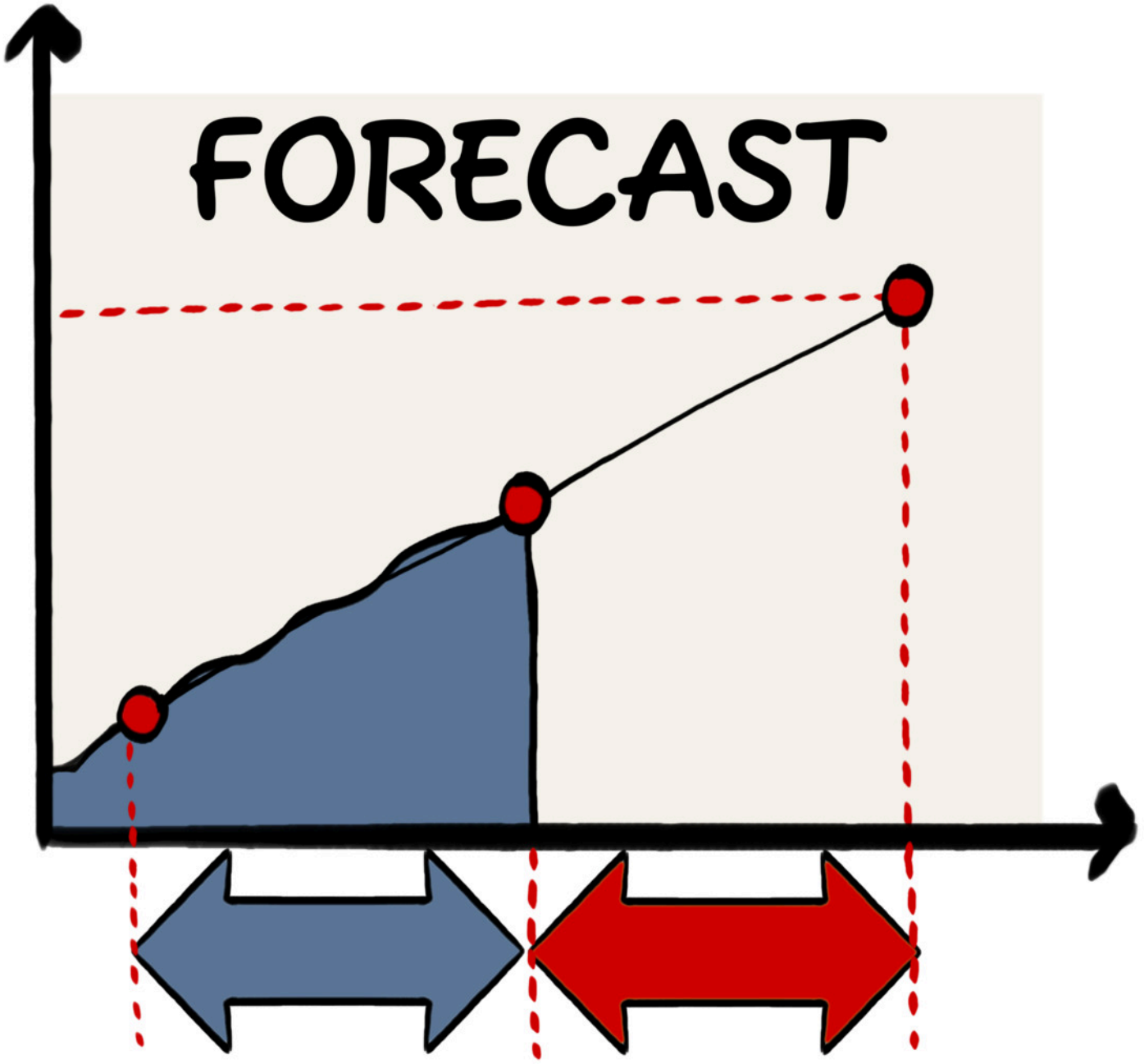
V IN PRACTICE

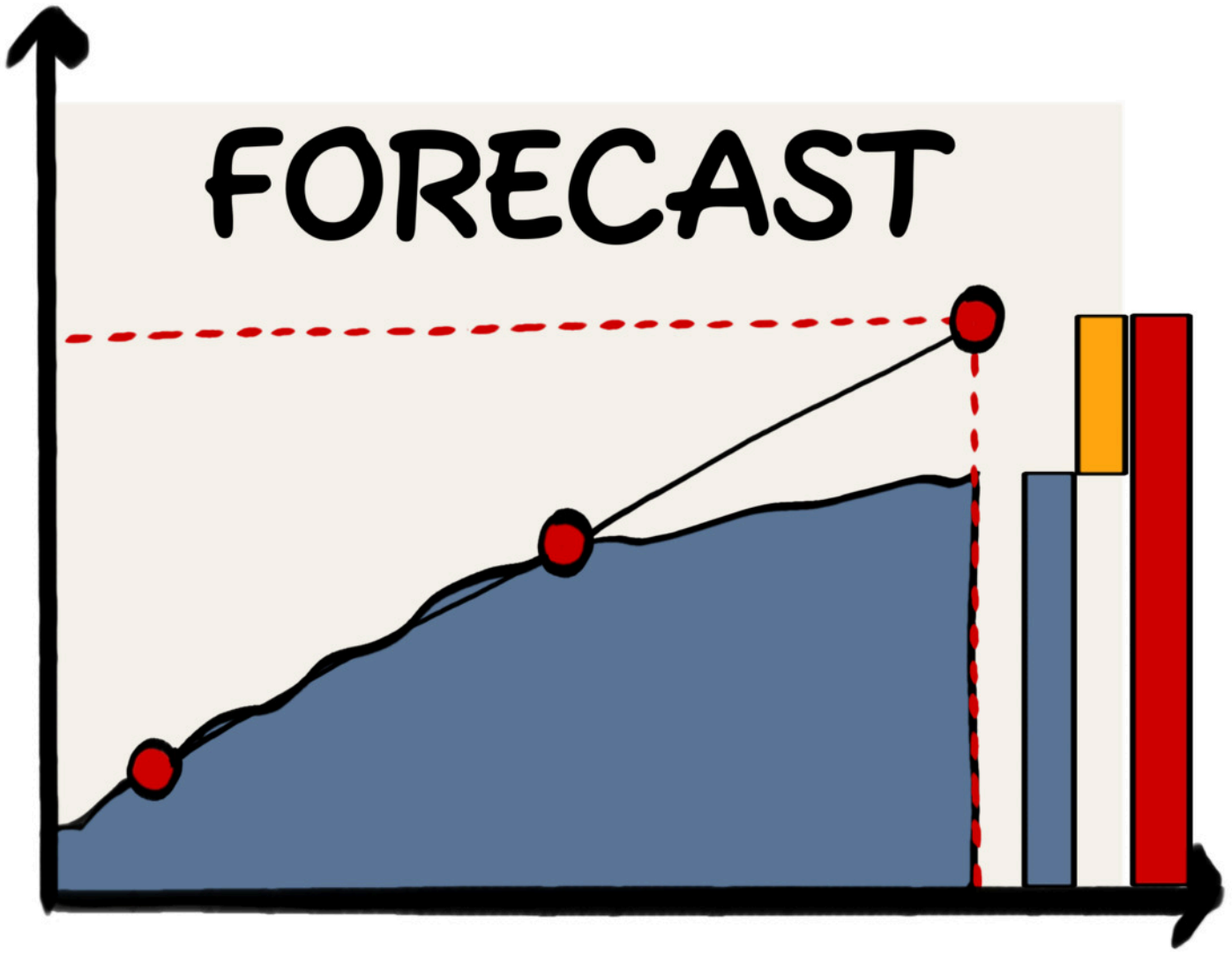
VI

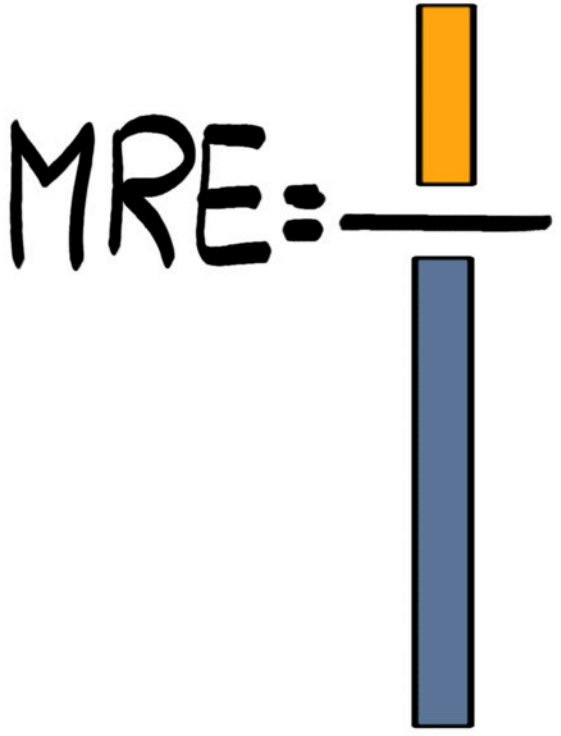
VII


TAWOS
DATASET

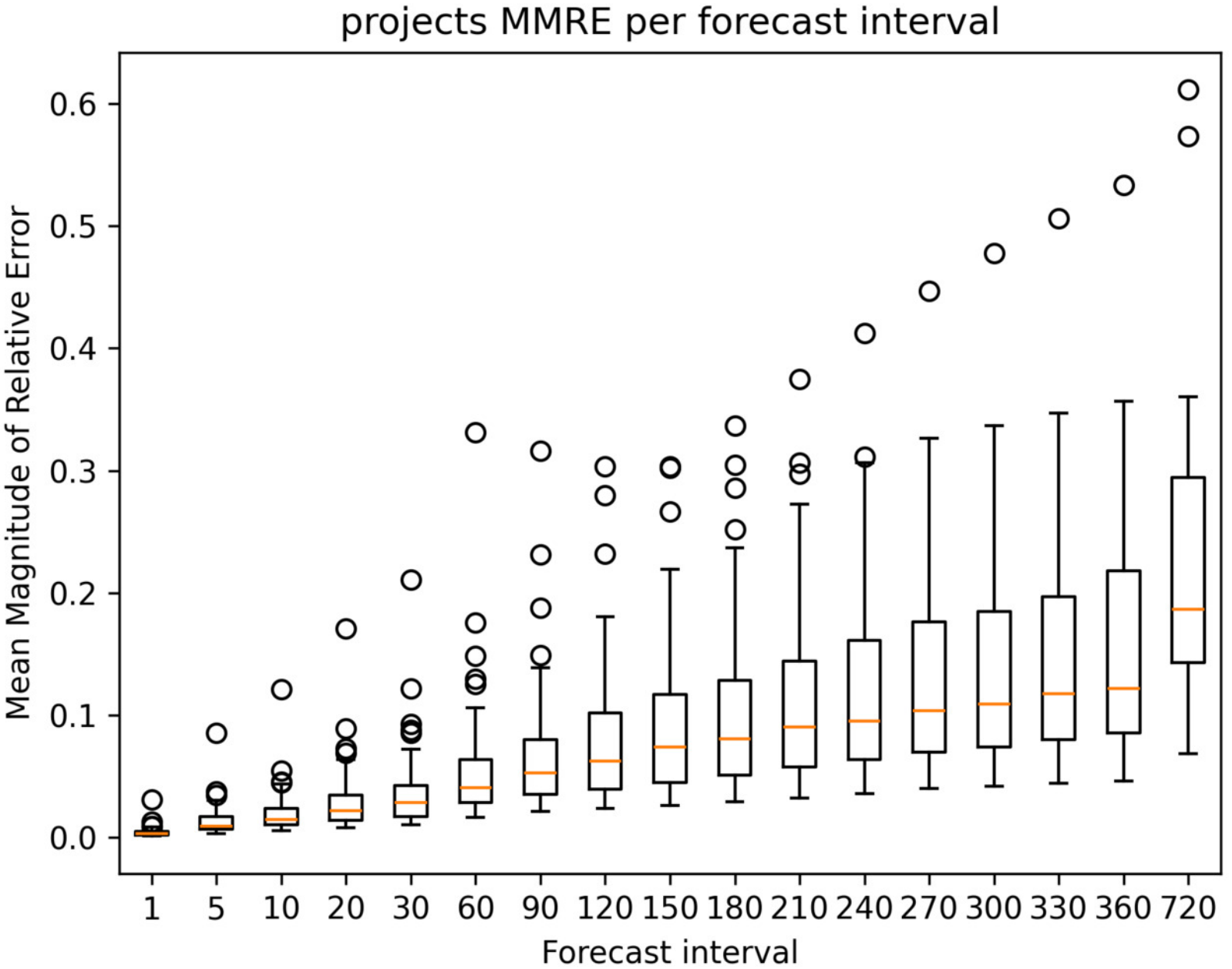

FORECAST

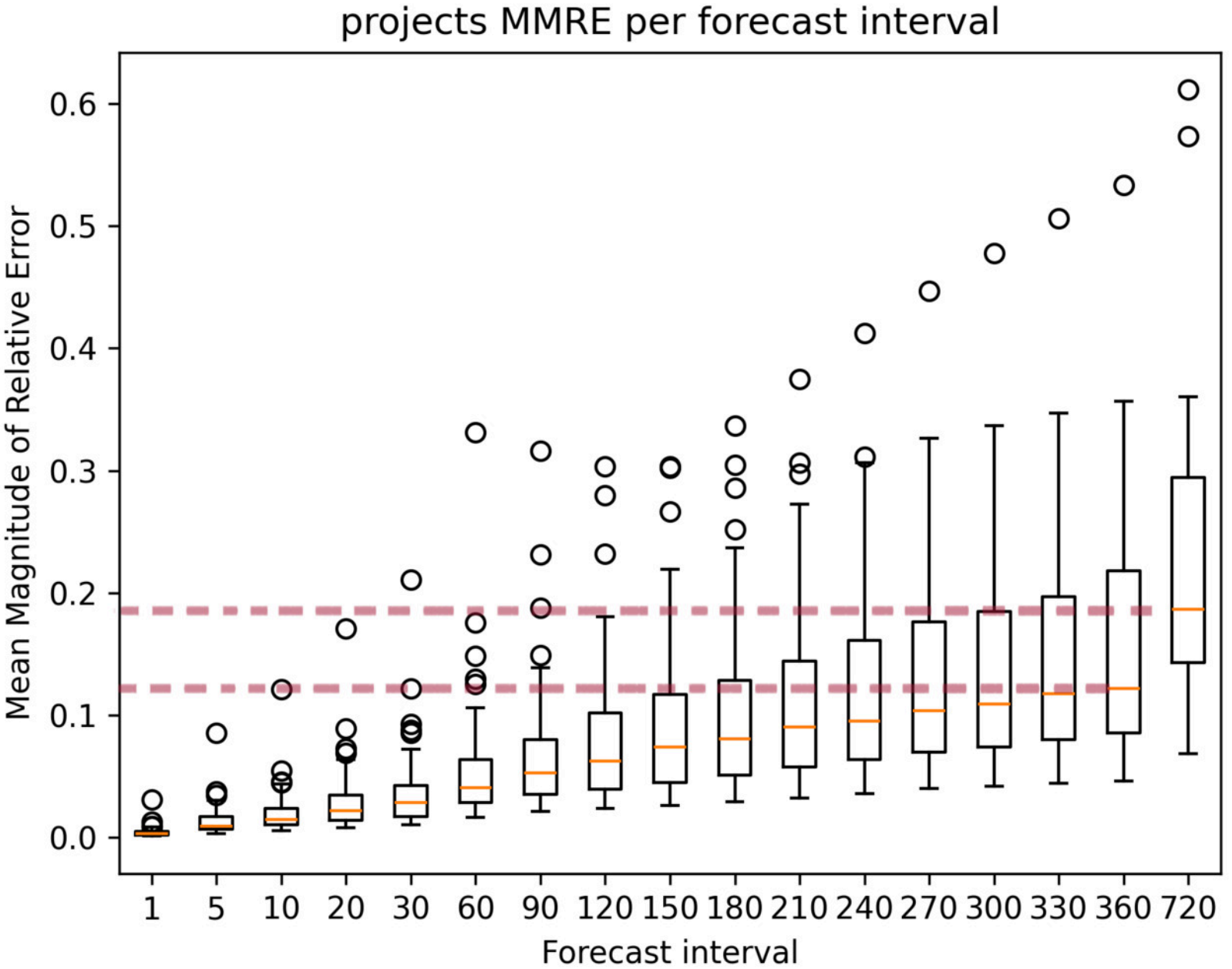

FORECAST

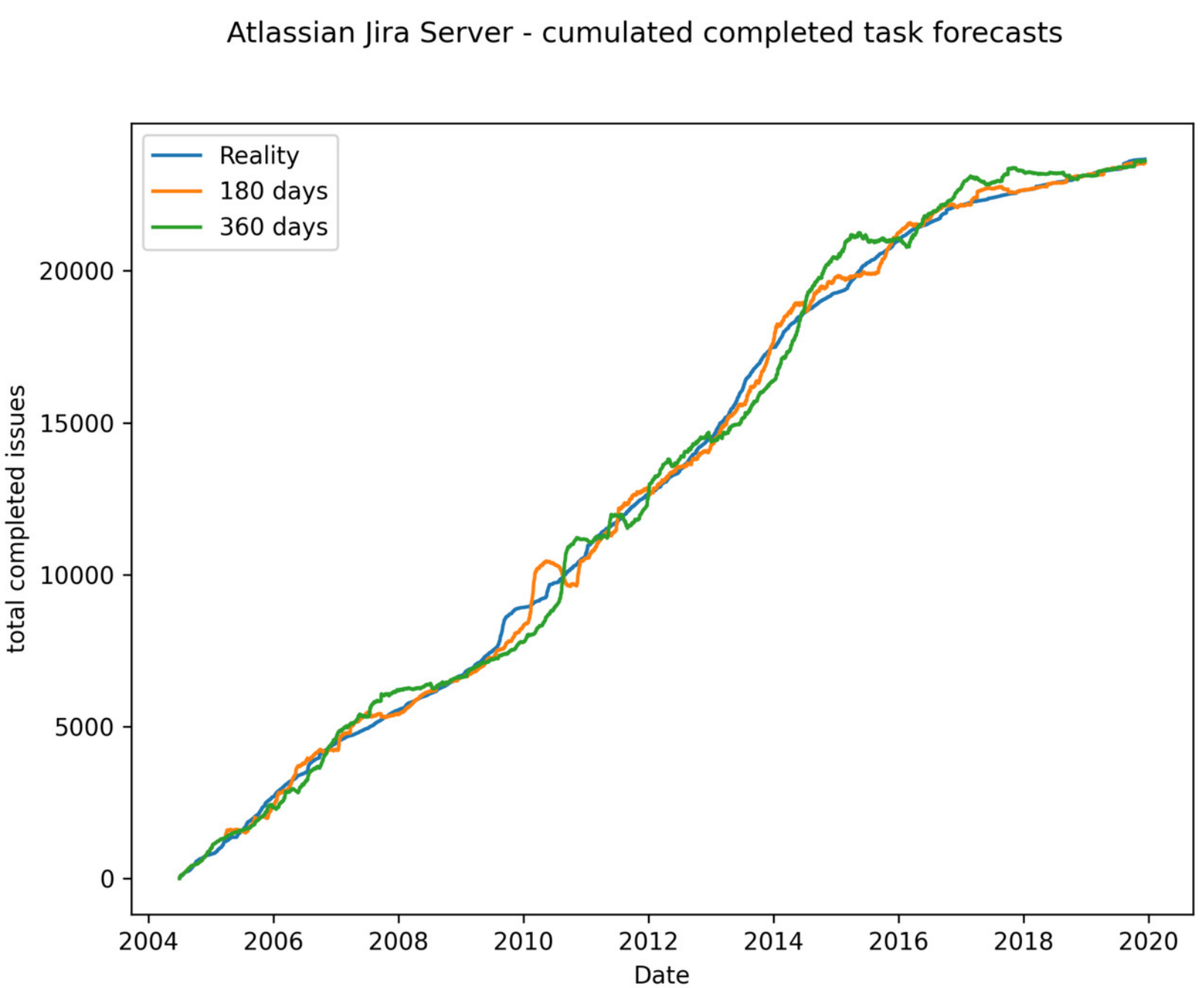


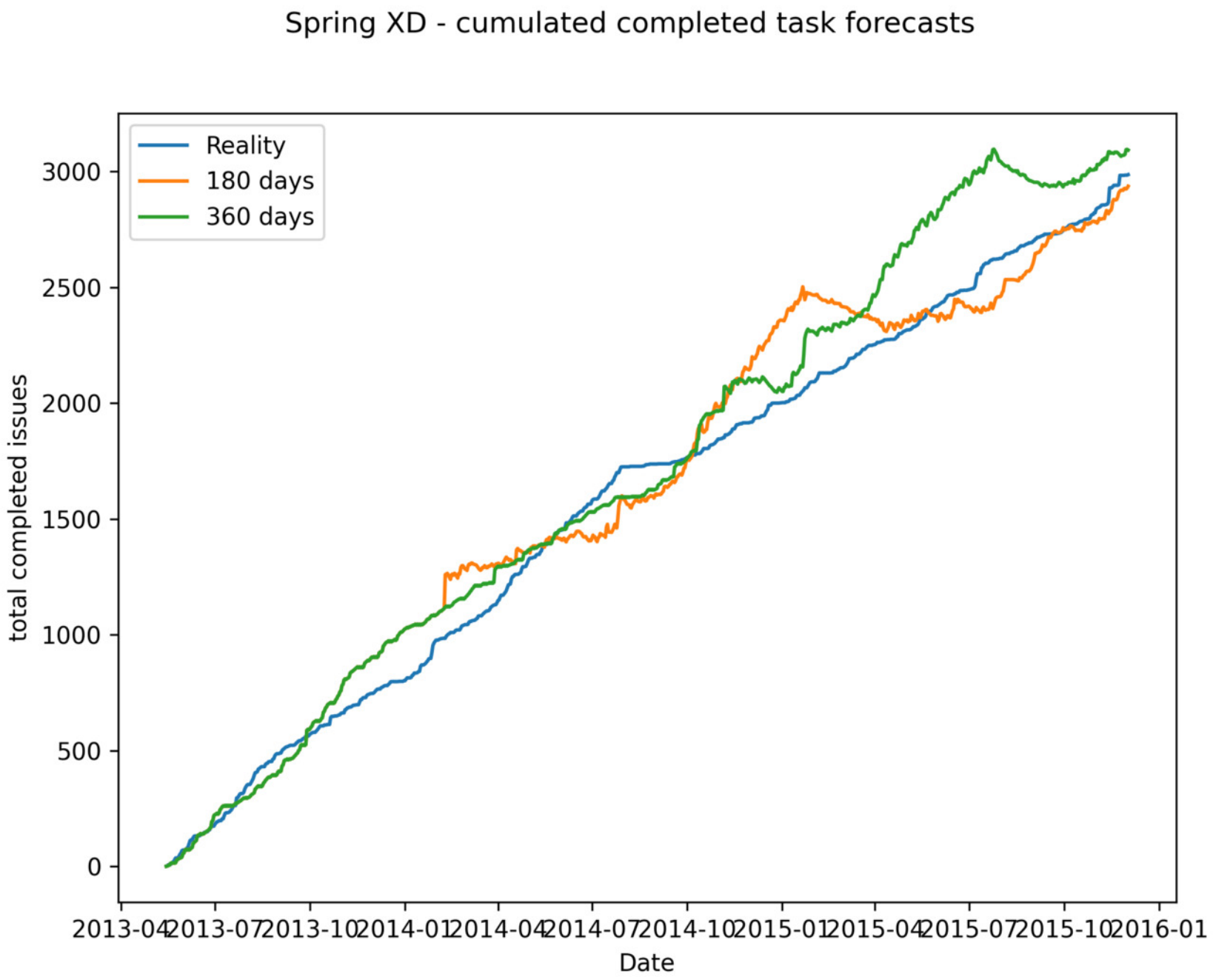


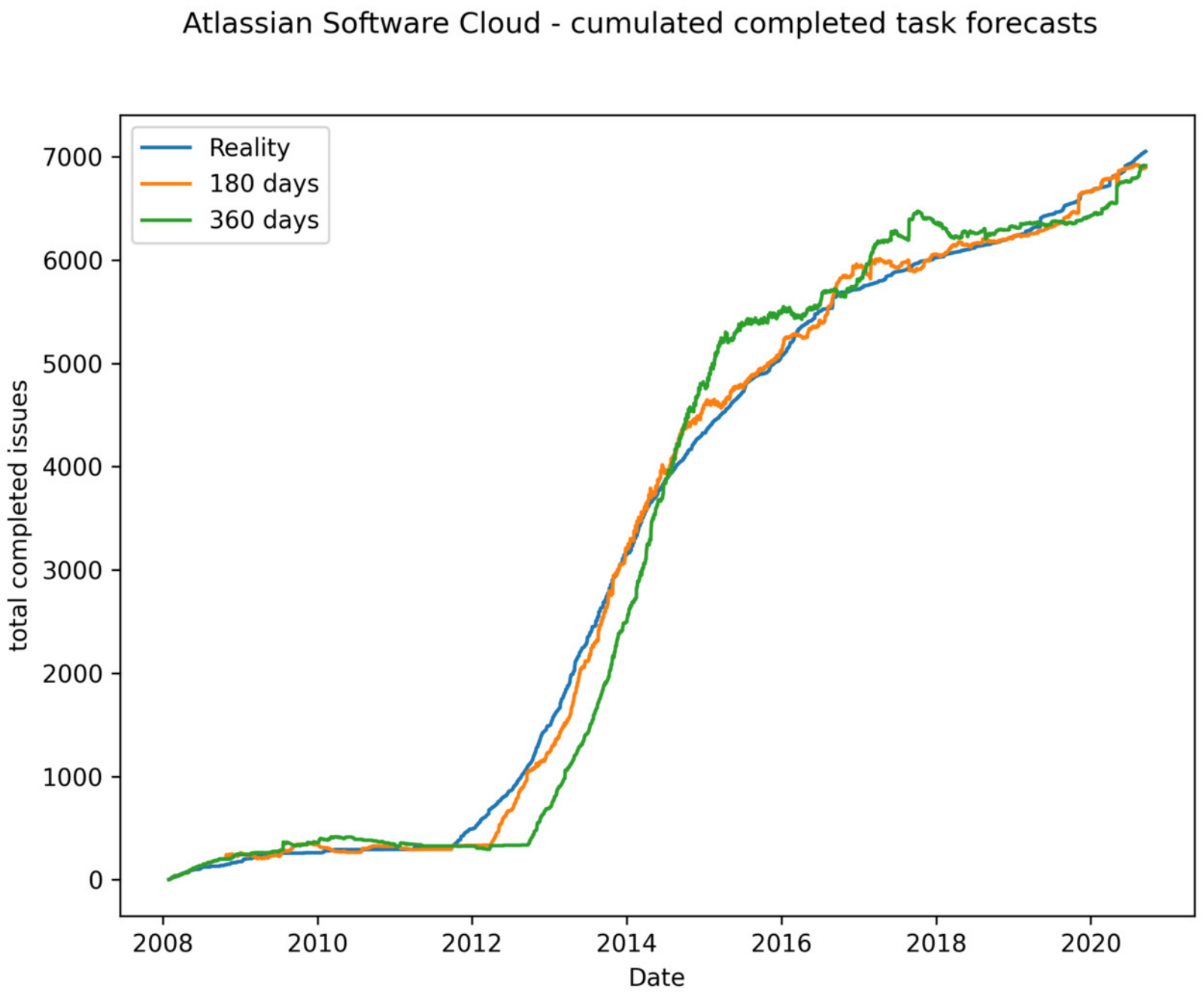
MRE: MAGNITUDE OF RELATIVE ERROR

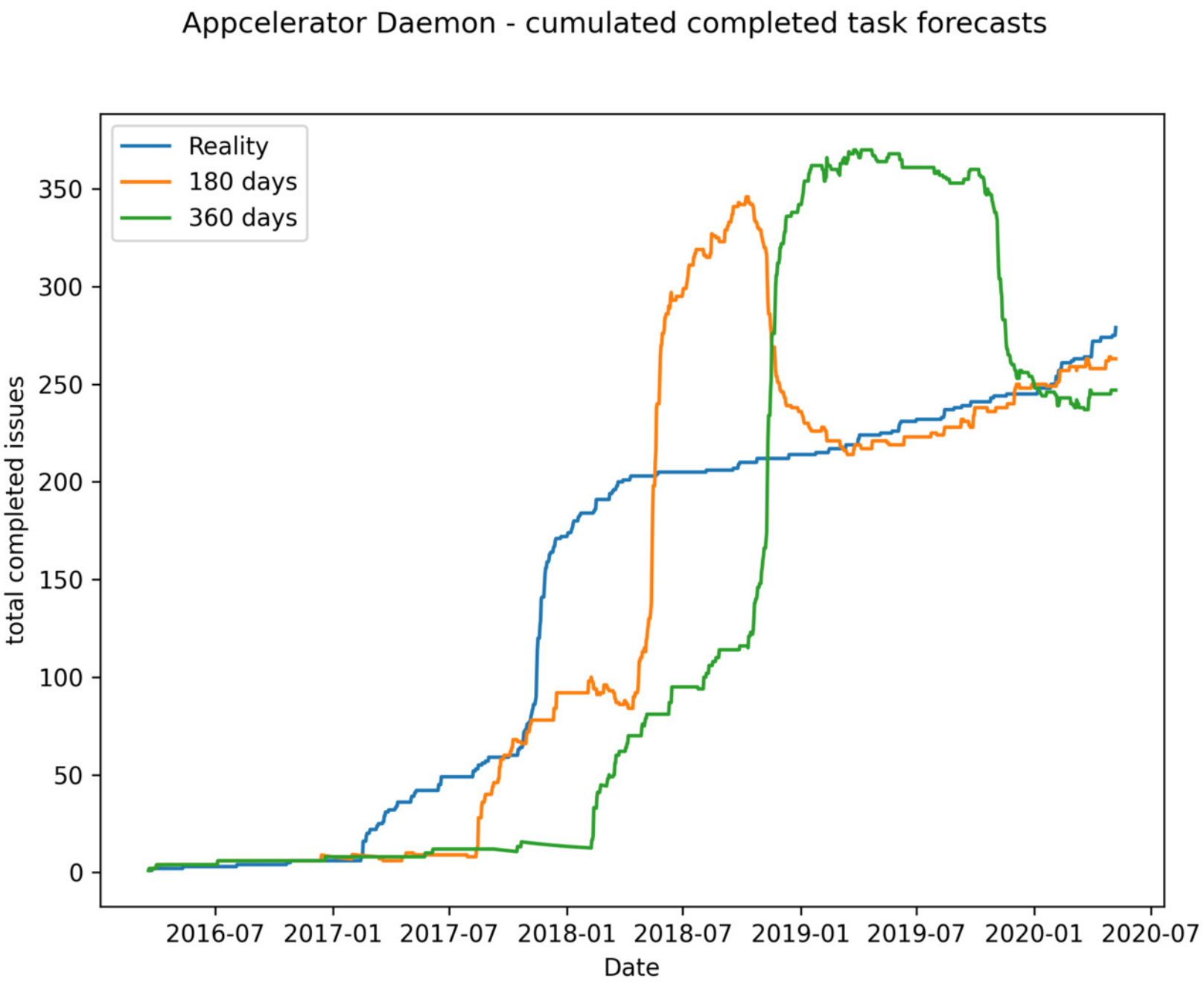

projects MMRE per forecast interval
Mean Magnitude of Relative Error
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1
5
10
20
30
60
90
120
150
180
210
240
270
300
330
360
720
Forecast interval


projects MMRE per forecast interval
Mean Magnitude of Relative Error
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1
5
10
20
30
60
90
120
150
180
210
240
270
300
330
360
720
Forecast interval


Atlassian Jira Server - cumulated completed task forecasts
Reality
180 days
360 days
total completed issues
0
5000
10000
15000
20000
2004
2006
2008
2010
2012
2014
2016
2018
2020
Date


Spring XD - cumulated completed task forecasts
Reality
180 days
360 days
total completed issues
0
500
1000
1500
2000
2500
3000
2013-04
2013-07
2013-10
2014-01
2014-04
2014-07
2014-10
2015-01
2015-04
2015-07
2015-10
2016-01
Date


Atlassian Software Cloud - cumulated completed task forecasts
Reality
180 days
360 days
total completed issues
0
1000
2000
3000
4000
5000
6000
7000
2008
2010
2012
2014
2016
2018
2020
Date


Appcelerator Daemon - cumulated completed task forecasts
Reality
180 days
360 days
total completed issues
0
50
100
150
200
250
300
350
2016-07
2017-01
2017-07
2018-01
2018-07
2019-01
2019-07
2020-01
2020-07
Date

I CONTEXT

II STORY POINTS

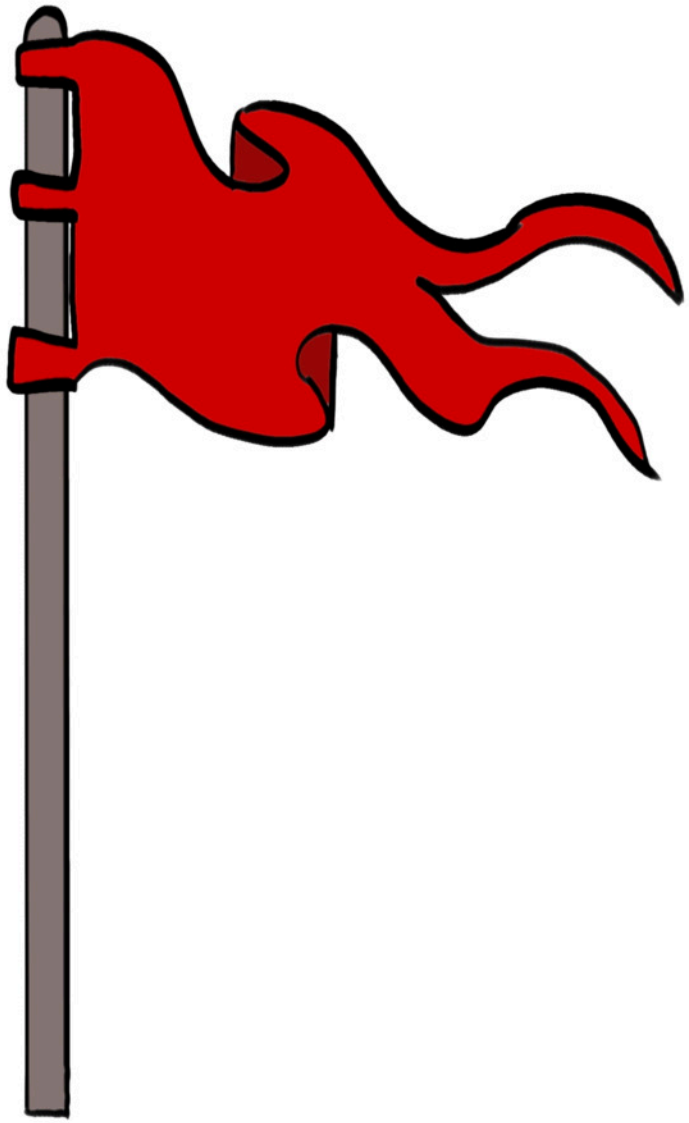
III NOESTIMATES TECHNIQUES

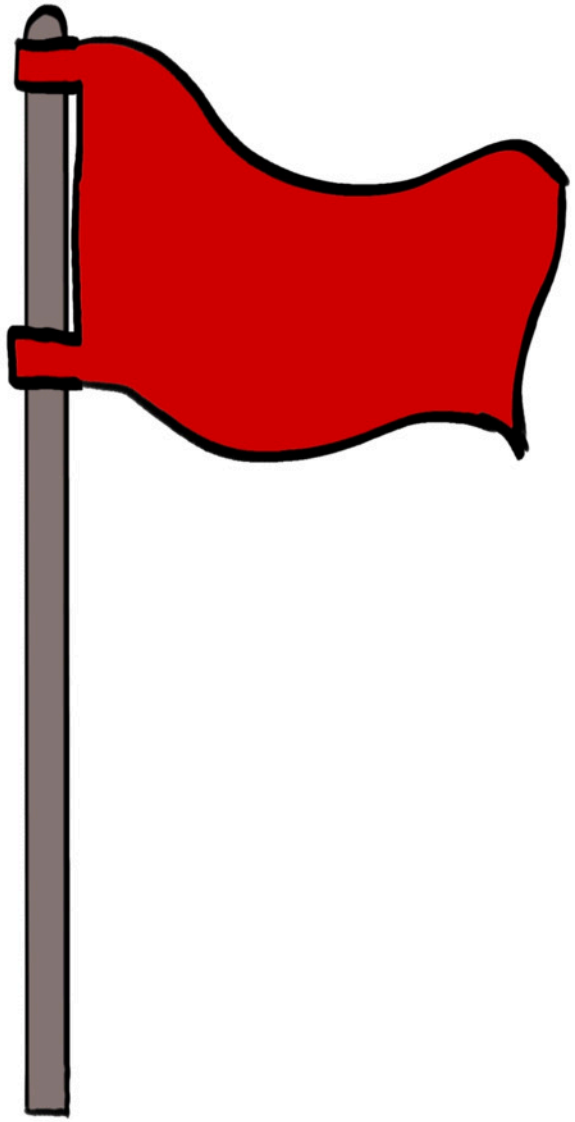
IV WILL AI SAVE US ?

V IN PRACTICE

VI CONTROL

VII






DEADLINE


DEADLINE


DEADLINE


DEADLINE

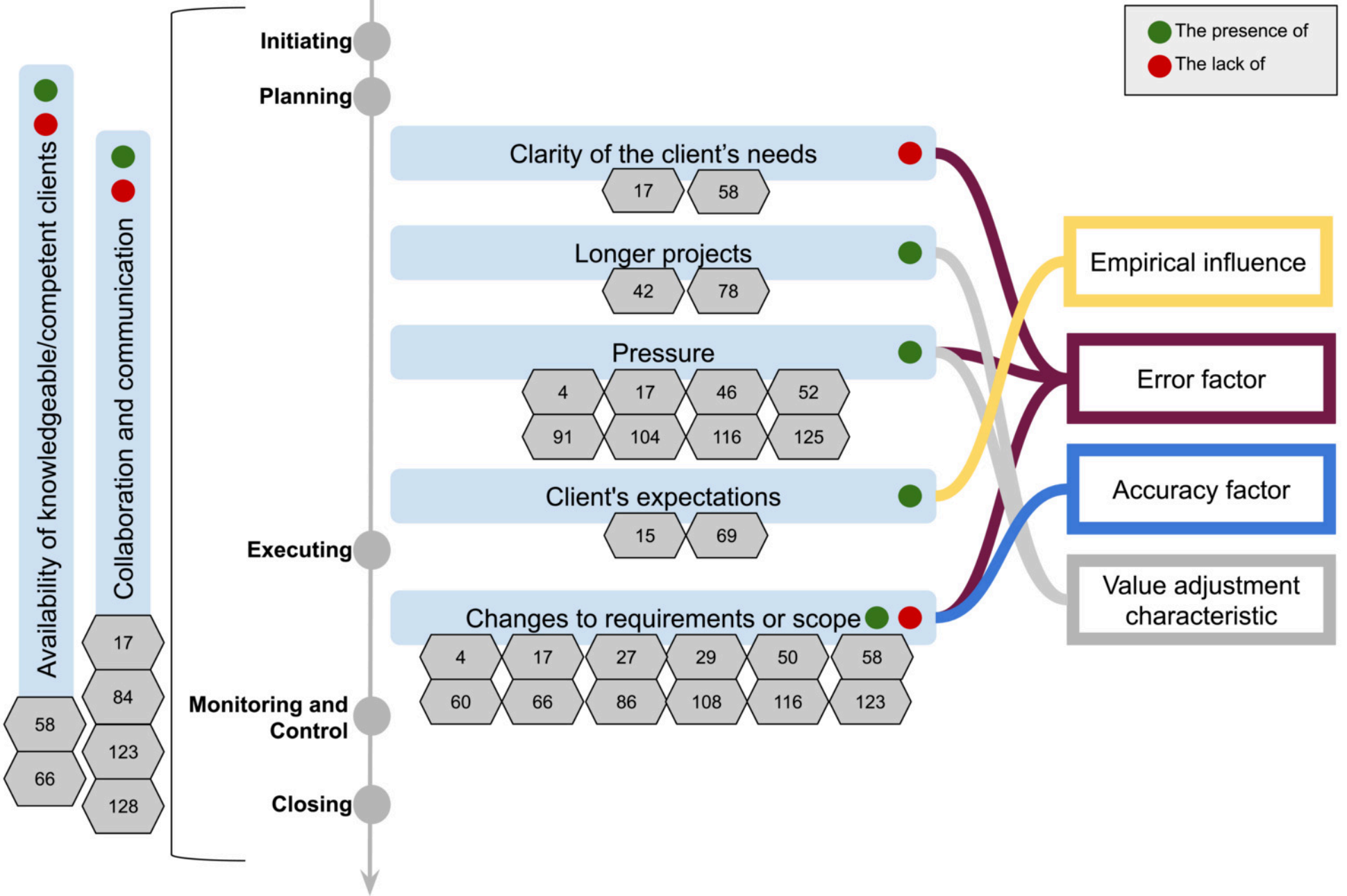


Figure 8 - Factors related to Customer/Client

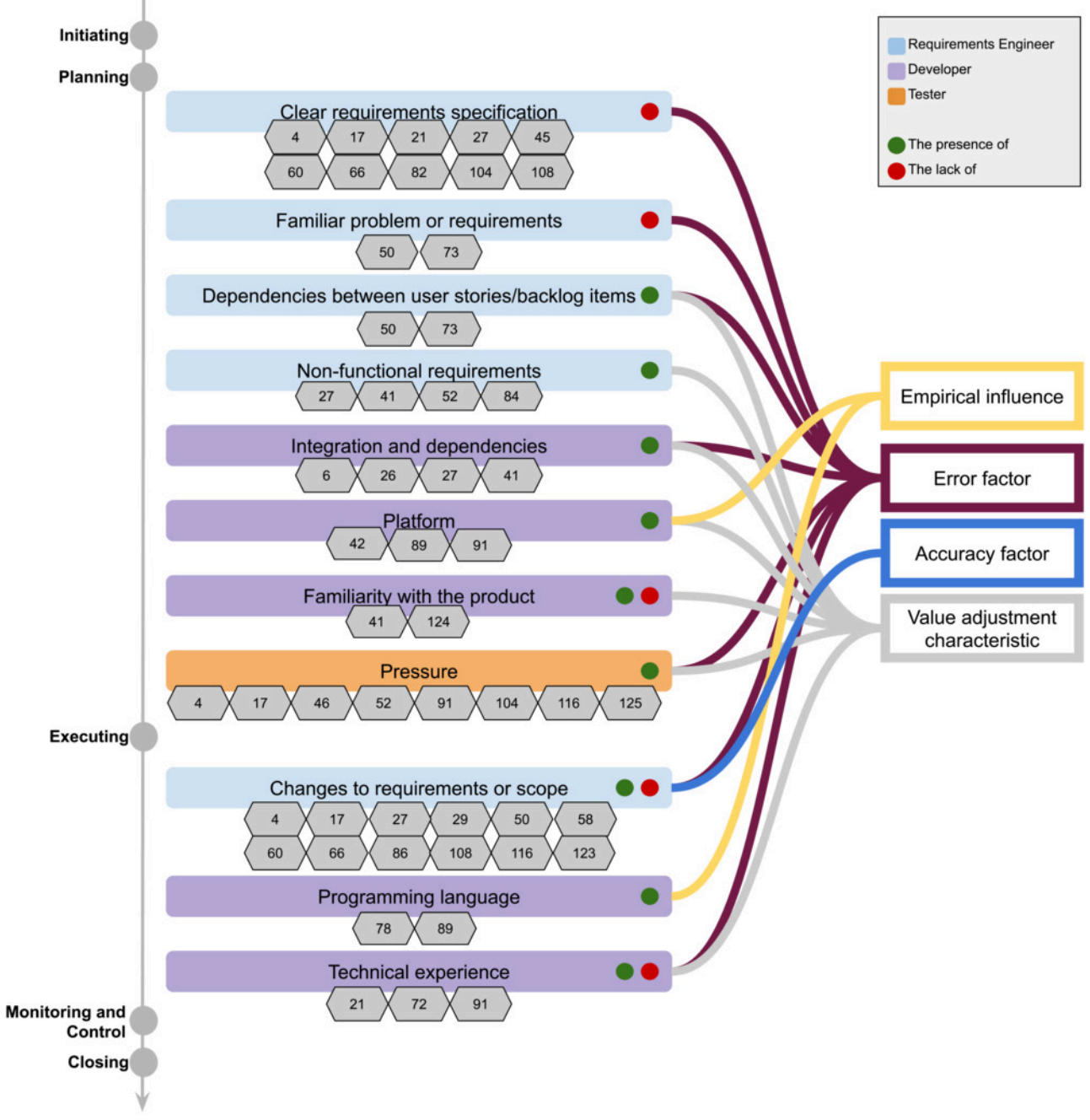


Figure 11 - Factors related to people in technical roles


DEADLINE

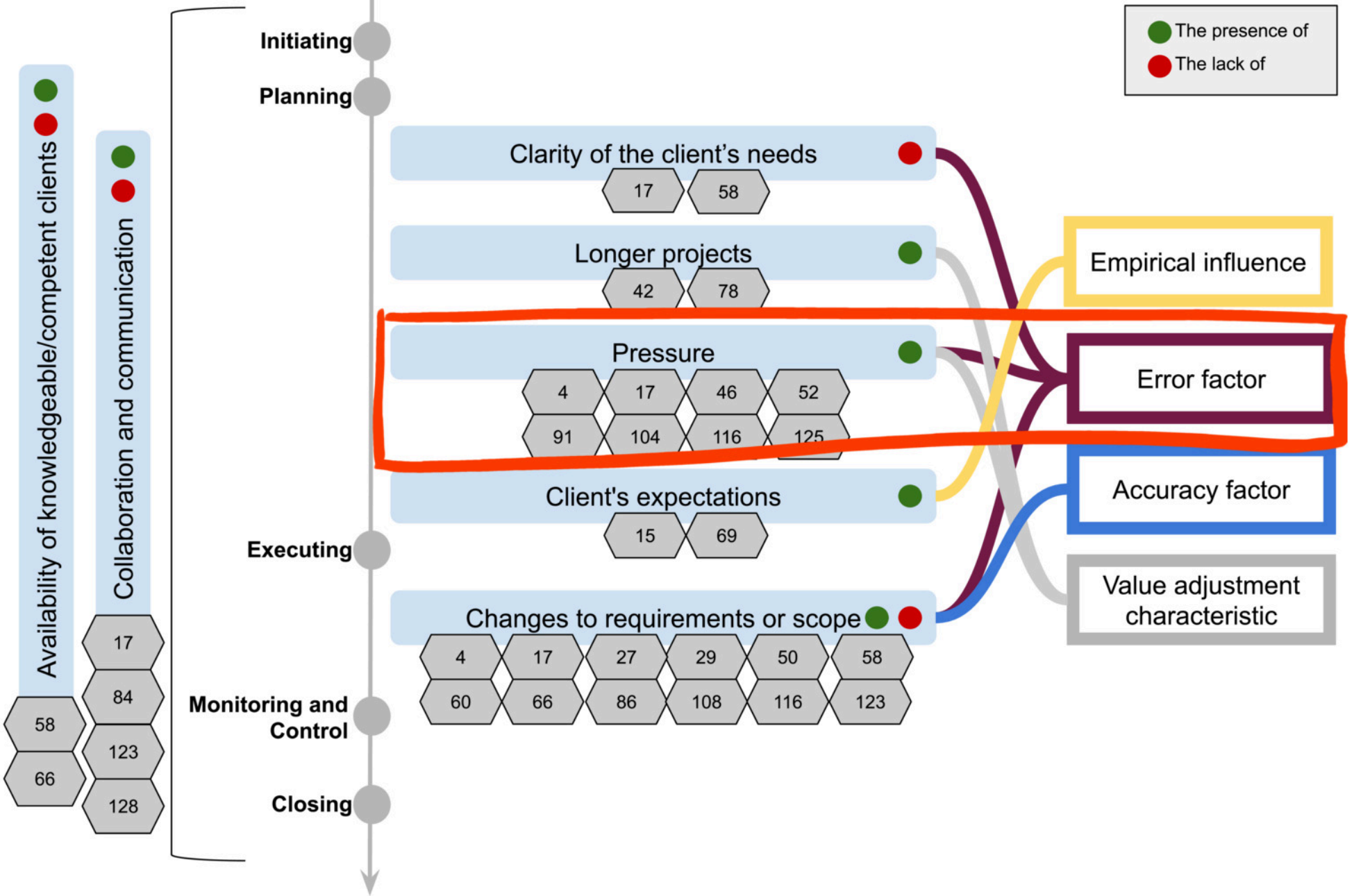


Figure 8 - Factors related to Customer/Client

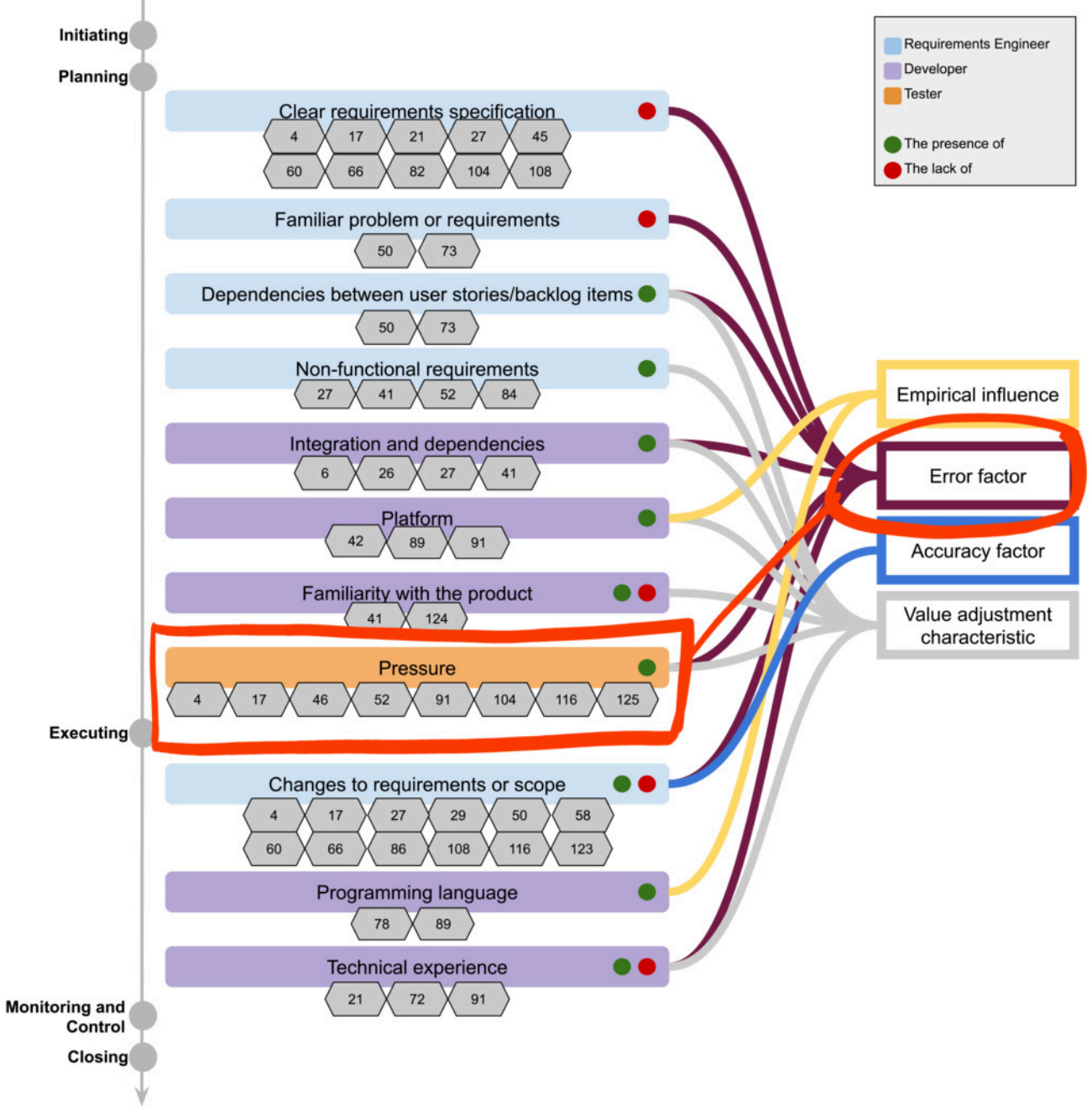


Figure 11 - Factors related to people in technical roles

I CONTEXT

II STORY POINTS

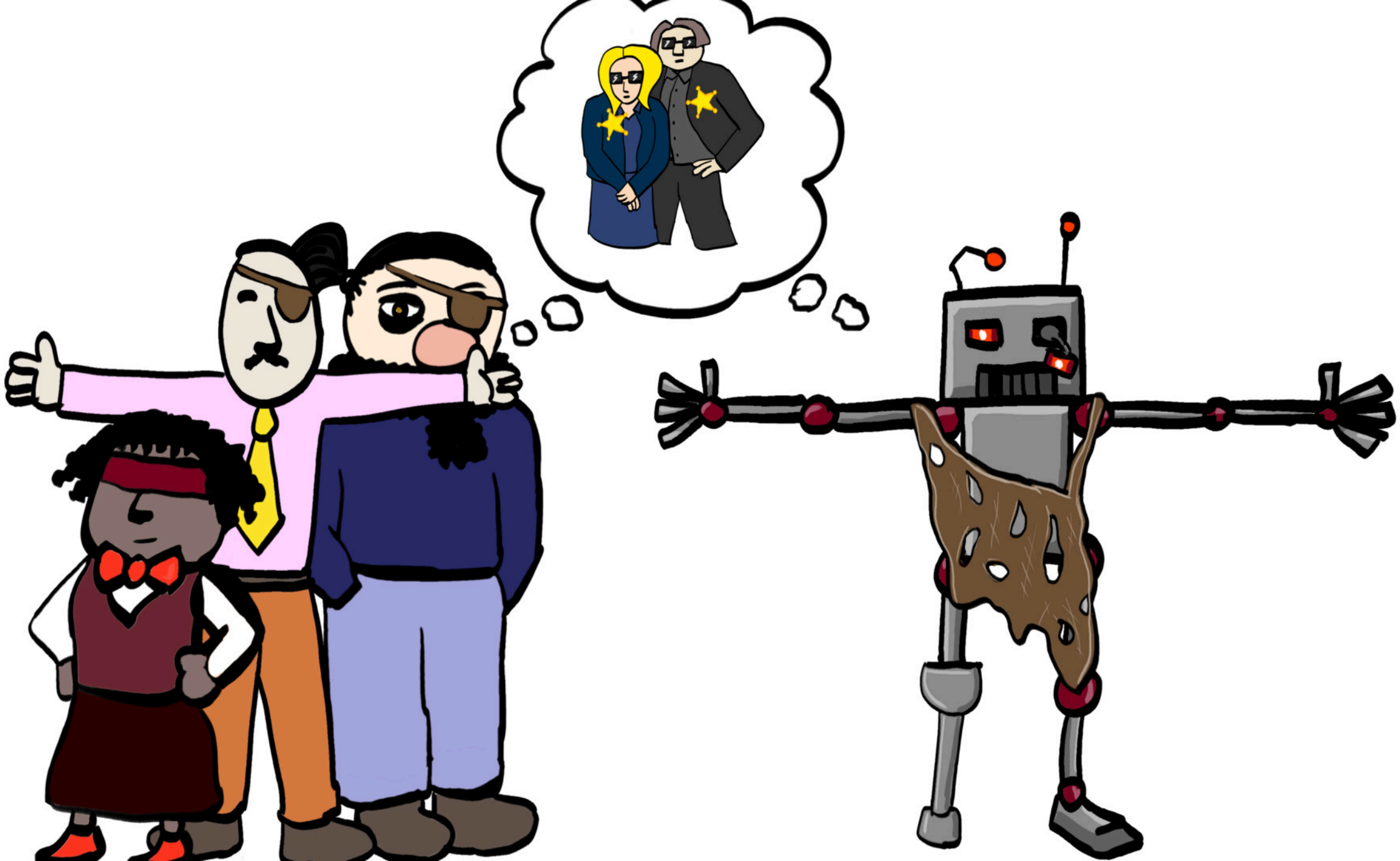
III NOESTIMATES TECHNIQUES

IV WILL AI SAVE US ?

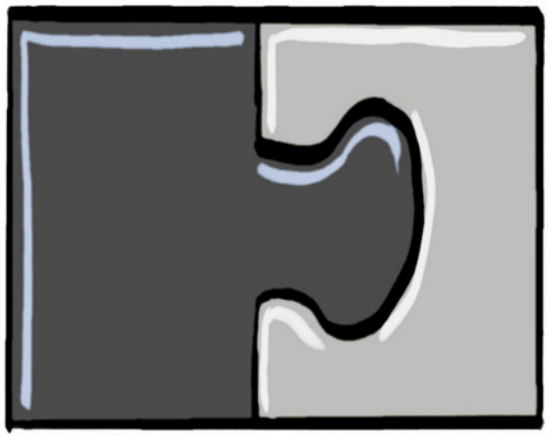
V IN PRACTICE

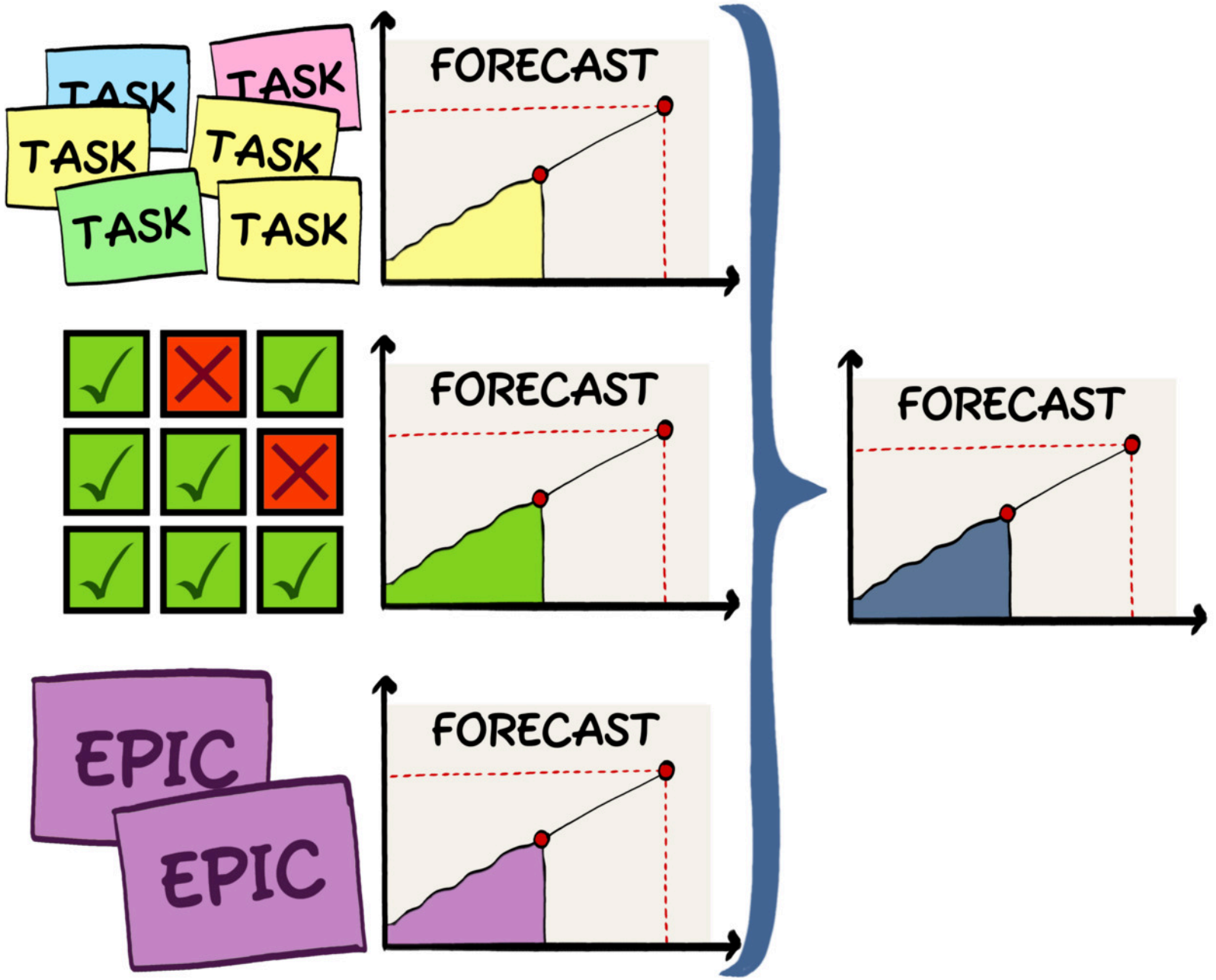
VI CONTROL

VII CONCLUSION

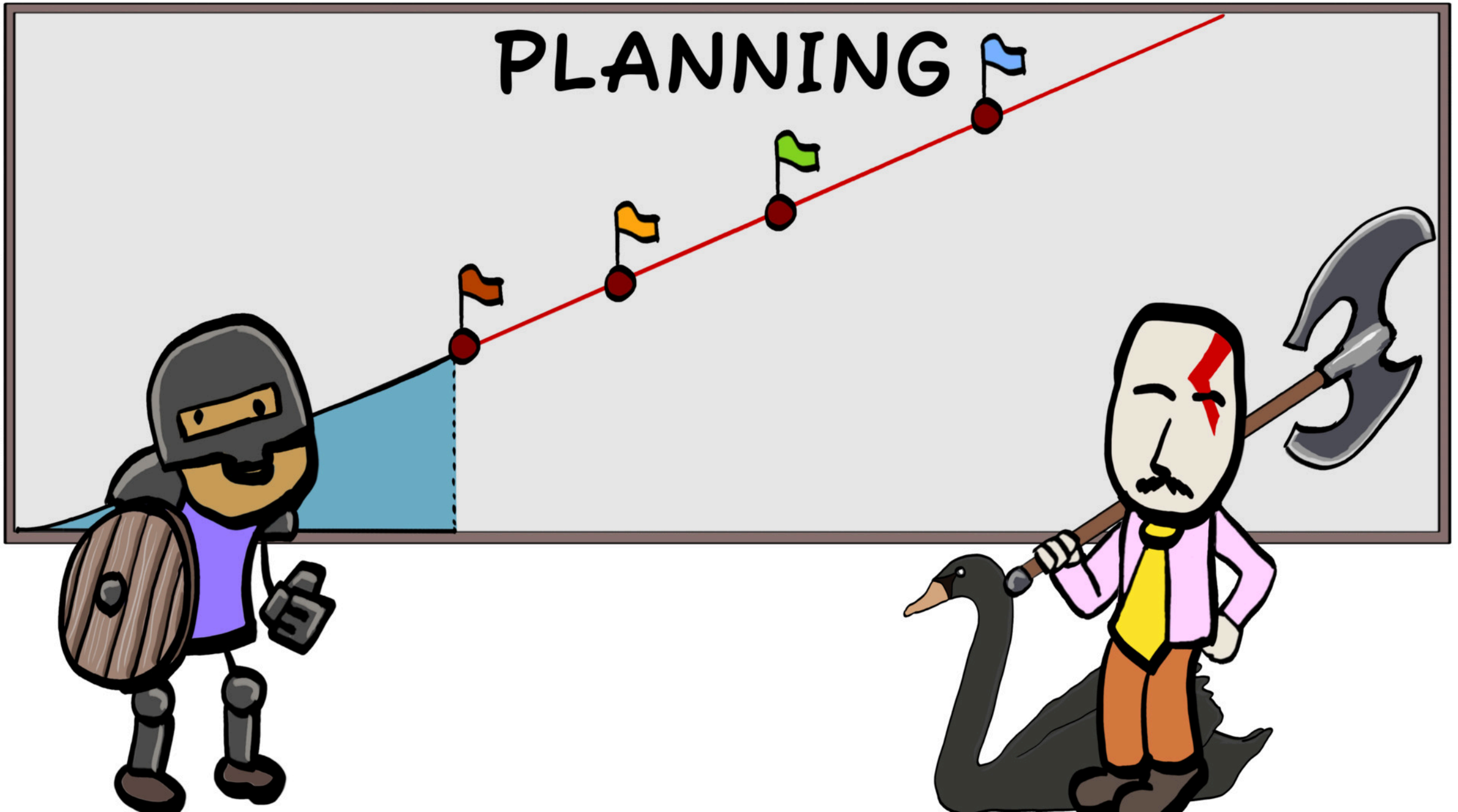


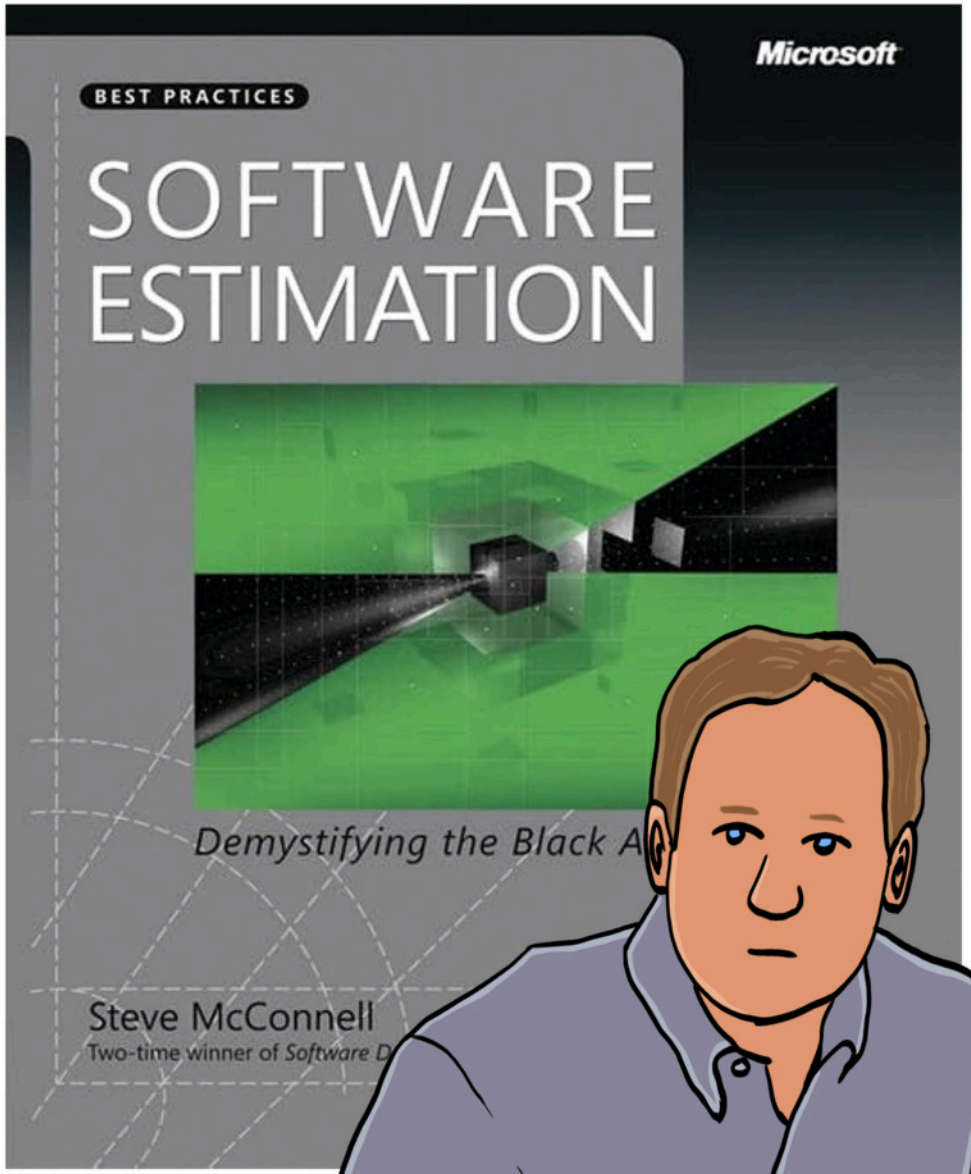

3
2
1
8
13
5
8
1
25
2
3
1

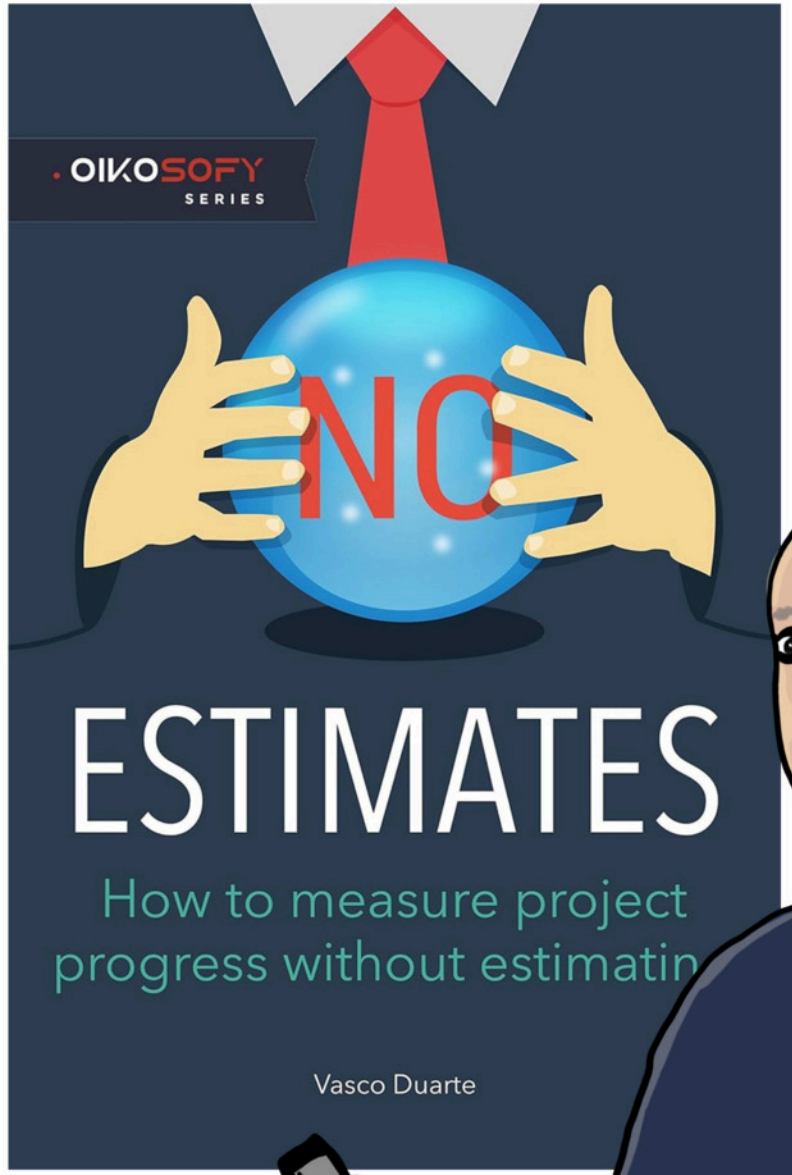



TASK
TASK
TASK
TASK
TASK
TASK
FORECAST
FORECAST
EPIC
EPIC
FORECAST
FORECAST

PLANNING


PLANNING


Microsoft
BEST PRACTICES
SOFTWARE ESTIMATION
Demystifying the Black A
Steve McConnell
Two-time winner of Software D


OIKOSOFY
SERIES
NO
ESTIMATES
How to measure project progress without estimatin
Vasco Duarte

MERCI


MERCI


MERCI
AGICAP

BORINGDEV.EU

@VICTORLAMBRET