



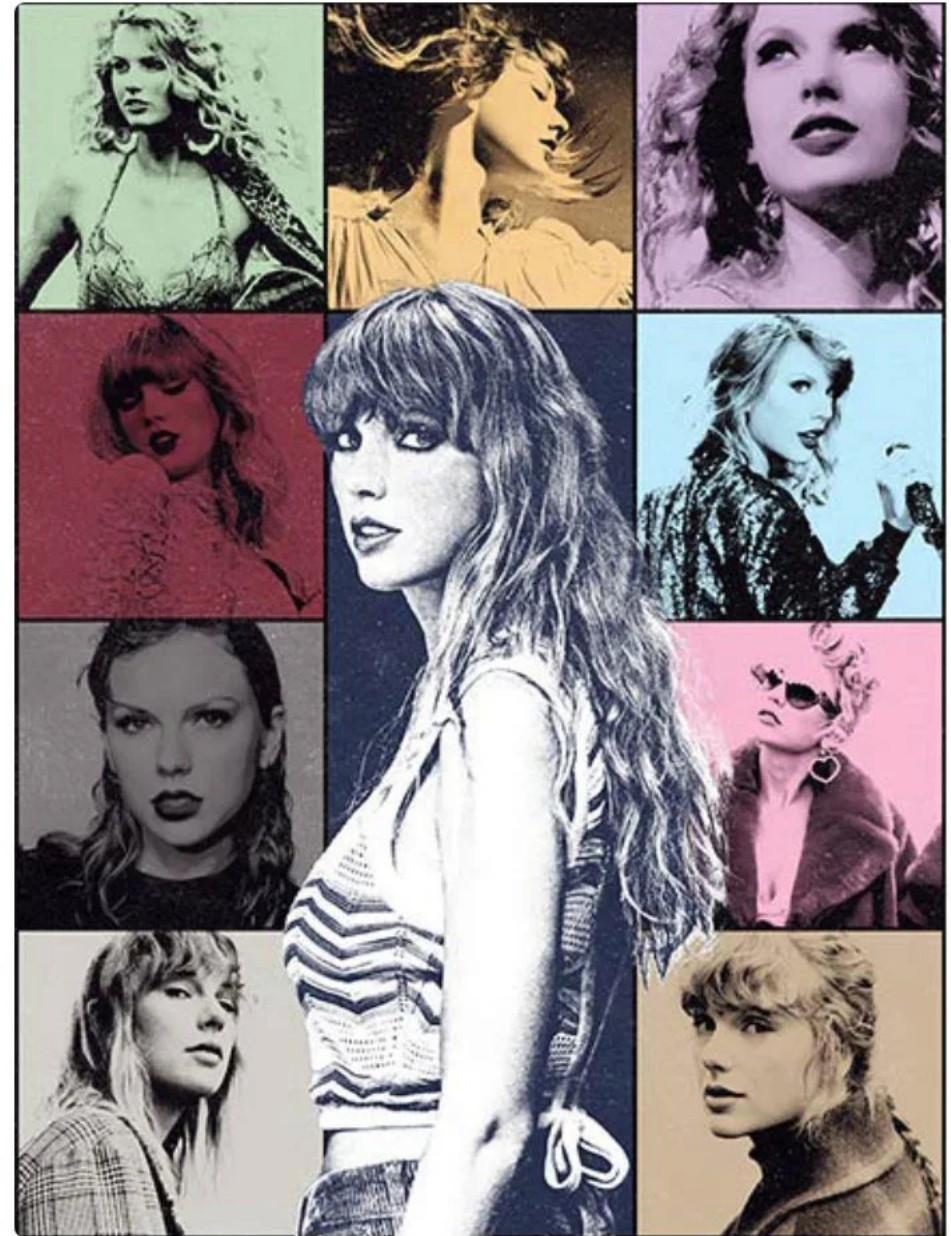
R-Evolution in Pharma

Ellis Hughes
Data Science Leader

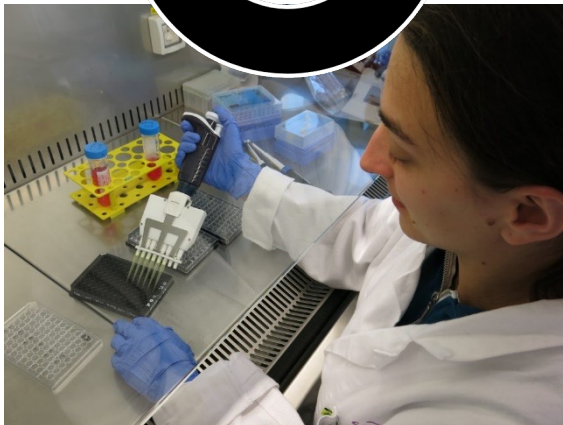
Taylor Swift

American Singer/Songwriter

- 2004 - Debuted a Country Music Singer
- 2014 - “Rebranded” in 2014 to “Pop” music, experimenting with different musical styles, changed streaming music industry
- 2021- Re-releasing songs with (Taylor’s Version), The Eras Tour
 - Tokyo Dome, February 2024



What does Taylor Swift have to do with the Pharmaceutical Industry?



- The Pharma Industry have been around for a long time, using tools we know worked
- 2010's – A shift to look at new, open source tools
- 2018 – Investigations and ground work to empower teams to use open source
- 2020 – Dedication from pharma industry to collaborate on open source tools

This is our version of the story

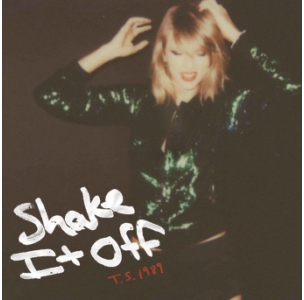


(GSK's Version)

R-Evolution in Pharma

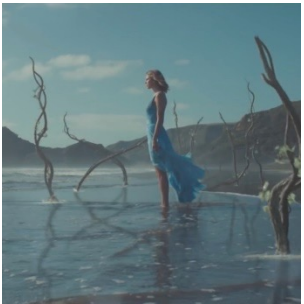
Ellis Hughes
Data Science Leader

Our Story in 3 Taylor Swift Songs



Shake it Off

- Shaking off old norms



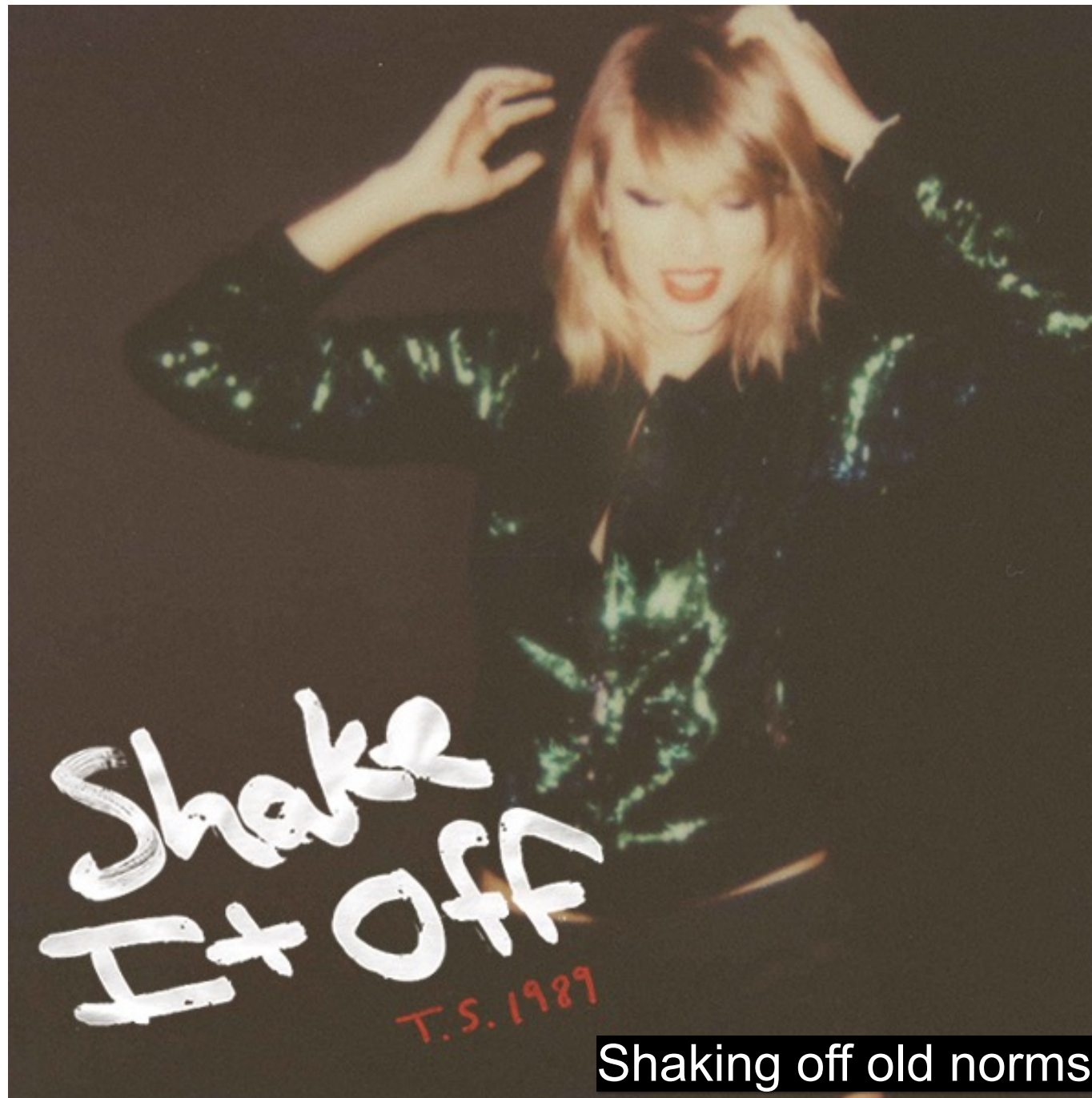
Out of the Woods

- Embracing new possibilities



Welcome to New York

- Creating innovative spaces



**In 2020 GSK Biostatistics decided to
make R a primary language for
statistical reporting**

Creating a New Normal

"I shake it off, I shake it off."



Adopt Multilingual
Approach for Clinical
Delivery



Begin to move
codebase to R and
piloting QC activities
using R



Start to use R as a
core tool for delivery



Begin to train and
embed R as a core
data science tool

Practical Experience

“Can't stop, won't stop moving”

- R4QC
 - Give teams ability to learn R while building out tools
 - Verify ability to recreate displays from SAS (gxp) using R
- R4Prod
 - One off before scaling use of R
 - Create baseline for future work
 - Alice Ehmann and Michael Rimler – R Got This! Delivering a complete Statistical Package using R
 - https://bit.ly/phuse_2022_r_got_this

Just In Time Training

“But I keep cruisin’ Can't stop, won't stop movin’”



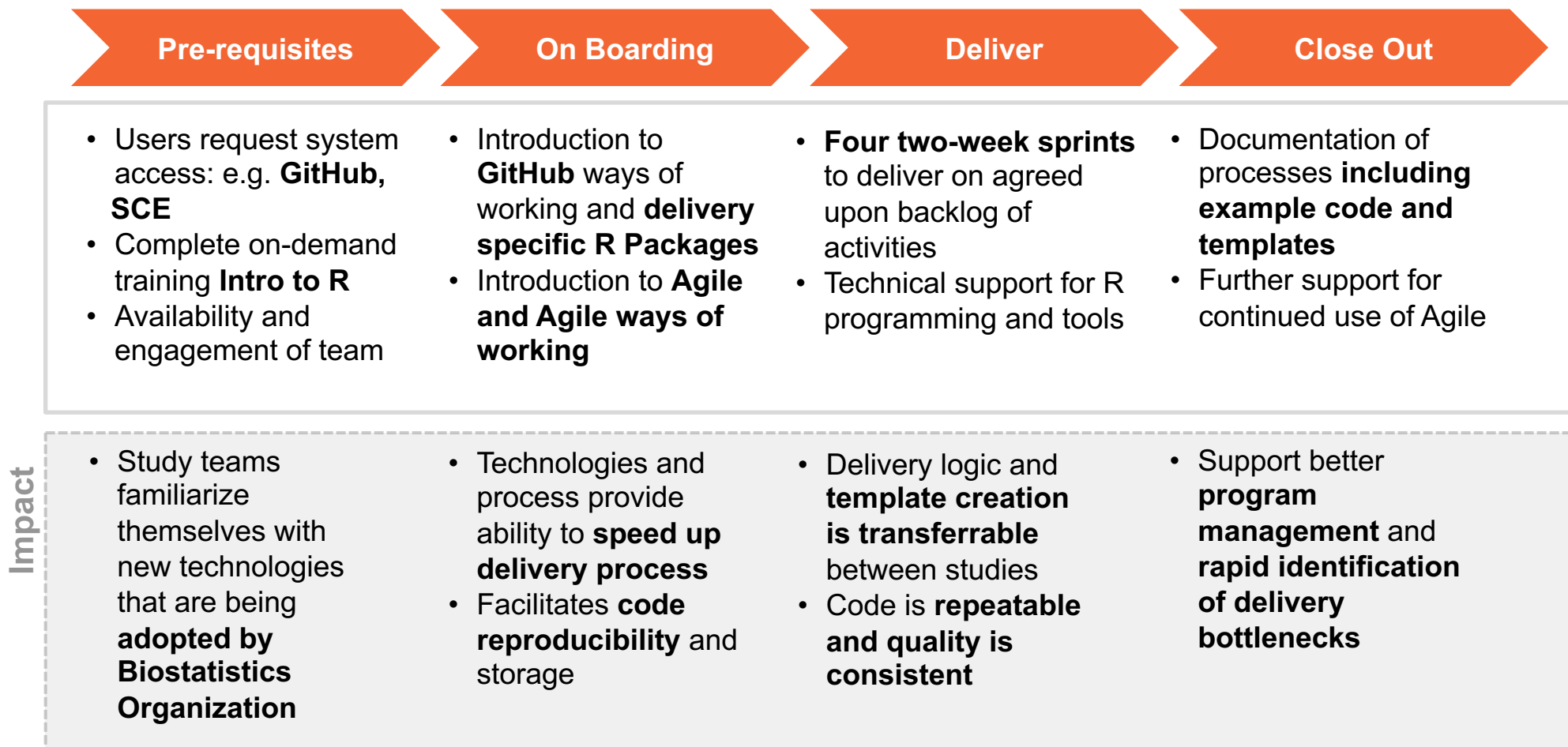
Learning From the Past

Training at Scale



Typical AccelerateR Engagement

How AccelerateR Utilises Tools, Technologies, and Processes



Training on Time

“I never miss a beat, I'm lightning on my feet”

- AccelerateR Delivering Training and Support In the Moment
- Teams learn more than R for success
 - git & github
 - New processes
 - New tools/ide's
- Ben Arancibia - The Need for Speed – AccelerateR-ing R Adoption in GSK (Posit::Conf(2023))





Embracing new possibilities

How do we know it's safe?

“But the monsters turned out to be just trees”

- FDA Statistical Clarifying Statement (2015) started the discussion
- What else and where else are there assumptions in the process
- How does this relate to outside the US?



Statistical Software Clarifying Statement

FDA does not require use of any specific software for statistical analyses, and statistical software is not explicitly discussed in Title 21 of the Code of Federal Regulations [e.g., in 21CFR part 11]. However, the software package(s) used for statistical analyses should be fully documented in the submission, including version and build identification.

As noted in the FDA guidance, *E9 Statistical Principles for Clinical Trials* (available at <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>), “The computer software used for data management and statistical analysis should be reliable, and documentation of appropriate software testing procedures should be available.” Sponsors are encouraged to consult with FDA review teams and especially with FDA statisticians regarding the choice and suitability of statistical software packages at an early stage in the product development process.

May 6, 2015

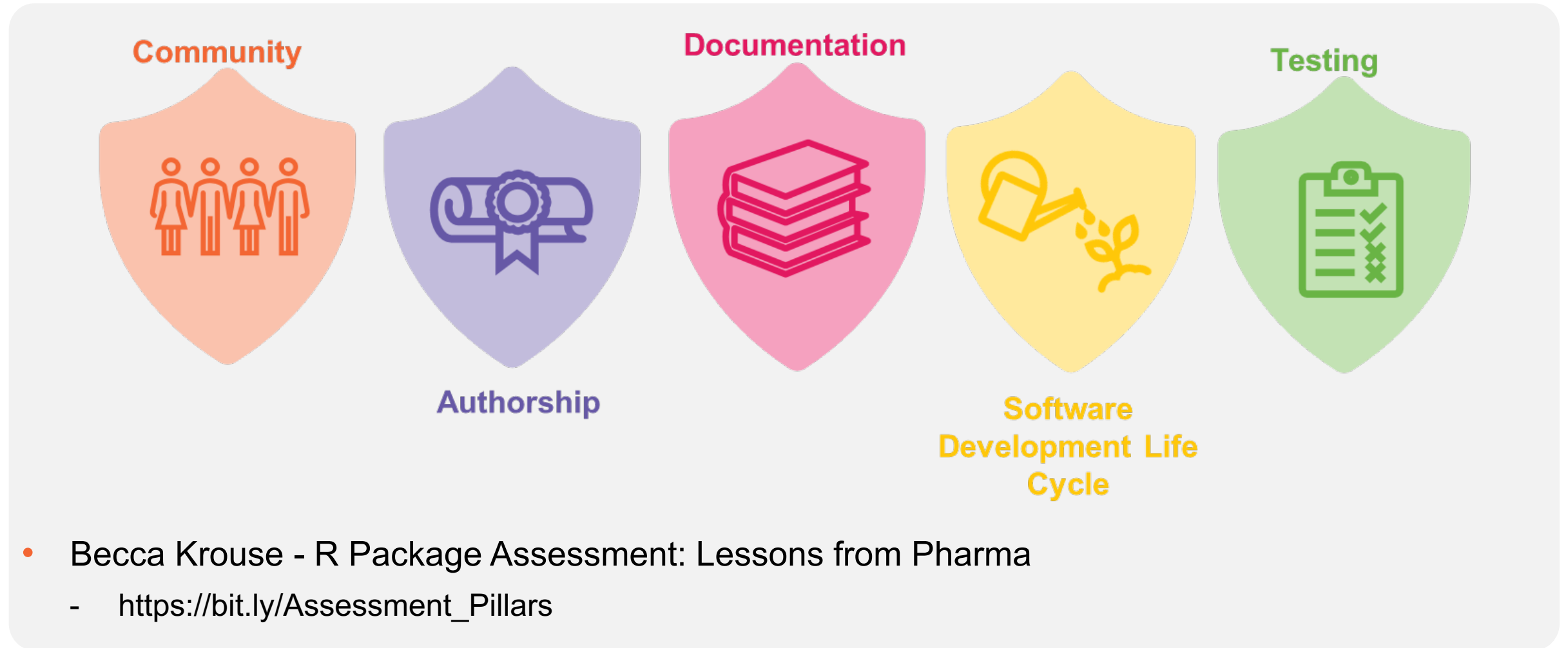
What are our responsibilities

“I walked out, I said "I'm setting you free"”

- ICH E6
 - Ensure and document that the electronic data processing **system(s) conforms to the sponsor's established requirements** for completeness, accuracy, reliability, and consistent intended performance (**i.e., validation**)....takes into consideration the **intended use of the system** and the **potential of the system to affect human subject protection and reliability** of trial results. (*ICH E6(R2) Section 5.5.3*)
- ICH E9
 - The **credibility of the numerical results** of the analysis depends on the **quality and validity of the methods and software (both internally and externally written)** used both for data management (data entry, storage, verification, correction, and retrieval) and **for processing the data statistically** (*ICH E9 Section 5.8 H*)

Pillars of Trust

“Looking at it now It all seems so simple”





Creating innovative spaces

Why construct a GXP Environment

“Kaleidoscope of loud heartbeats under coat”

- Ensure version of software
 - Reproducibility
 - Traceability
 - Control
- Access to other tools within the company
 - Databases
 - Folder Systems



What makes it into your environment

“Welcome to New York, it's been waitin' for you”

- Open Source tooling is developing constantly
 - External packages developed by community
 - Internal packages created to support specific processes
- Endorsed vs. Assessed packages
 - Production use Assessed Only
 - QC
 - Non-GXP work
 - R&D

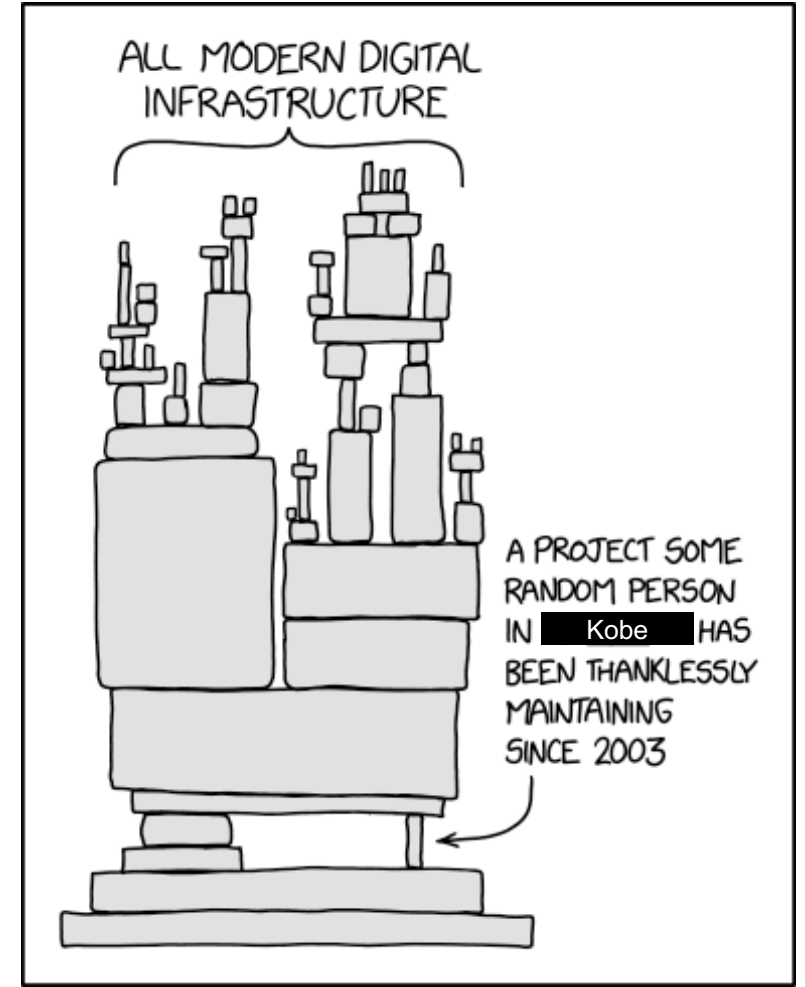
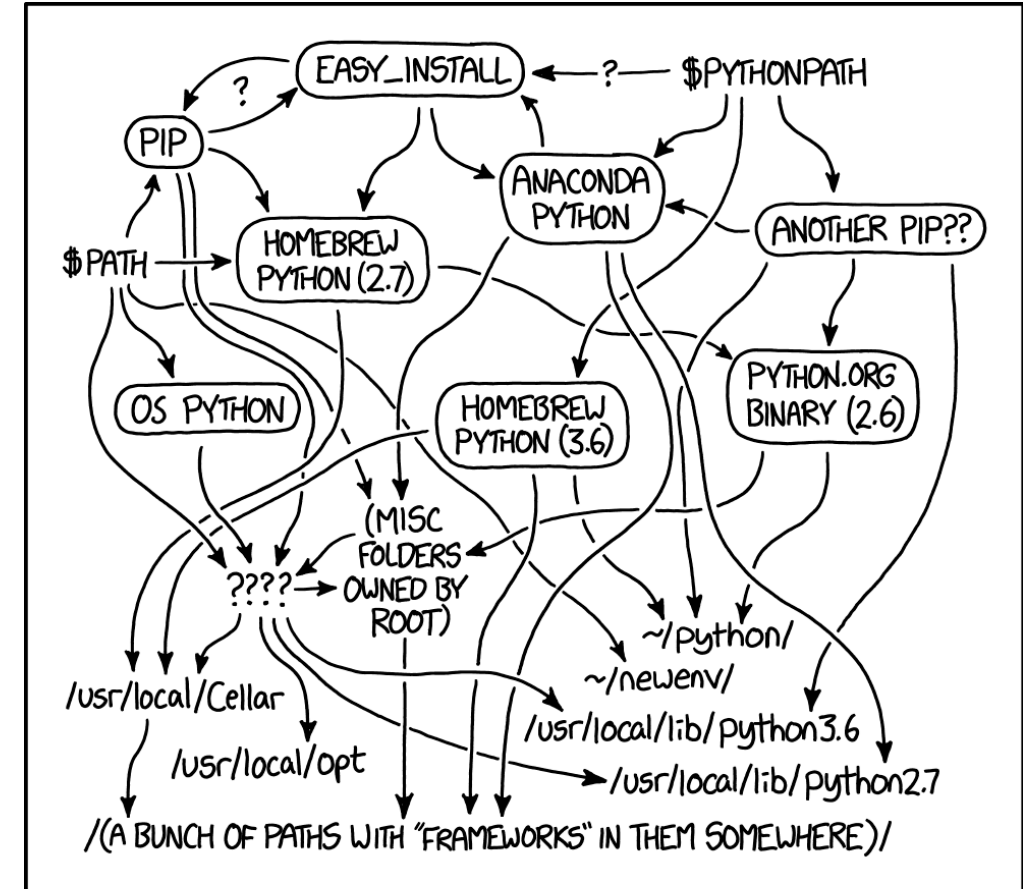


Image Source: <https://xkcd.com/2347/>

Creating Proof

“The lights are so bright, but they never blind me, me”

- Computer systems have many dependencies, proof of correct behavior in system is necessary
 - *ICH E9 Section 5.8 H*
- Using programmatic tools, easily execute tests and record results
- Where the rubber meets the road – Qualifying an R Environment
 - https://bit.ly/phuse_2023_environments



MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED
THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

Image Source: <https://xkcd.com/1987/>

A large, flowing, orange graphic that starts from the top left and curves towards the right, creating a sense of movement and energy.

Now its your turn to write your own story

GSK



sk.com

Benefits of the Open Source

“And I don't know how it gets better than this”

- Development
 - New Statistical Techniques often developed in Open Source
 - Open Source is already embedded and well tested
 - Share burdens in the “non competitive space”
- Workforce
 - Recruiting, Graduates know Open Source
 - New Mindsets
- Thomas Neitmann – Why Roche is Investing into Building Open-Source Software for Clinical Reporting
 - https://bit.ly/phuse_2023_OS_investments

Participating in the Revolution

“And you know I wanna ask you to dance right there”

- Pharmaverse
- Open Source in Pharma
- R Validation Hub
- R Consortium
 - Submissions to FDA
 - Engaging with more Health Authorities (PMDA)
- Collaborate
 - {admiral}
- Create and Contribute



R-Evolution in Pharma

(YOUR Version)