Florian Lier | 04/02/2019 | Bielefeld,CITEC

CITK 2019++ Strategie

Strategie

Zukünftige Ausrichtung des CITK auf Basis vergangener Aktivitäten und der aktuellen Situation u.A. Zeitraum, Entwickler, Userbase, Funding …

Abgeleitete Anforderungen

Neue Namensgebung (?) und die daraus abgeleiteten Anforderungen für den Umzug auf GitHub unter Berücksichtigung der zukünftigen Strategie

Arbeitspakete

Definieren von Arbeitspaketen/Schritten für u.A. den Umzug

English slides now

Post Gar-Installer ~2013/2014

Usually, nothing worked out of the box*. Hours and hours of compiling fixing bugs and writing glue code.

"… um yeah, it compiles on **my** machine!"

"No, can't fix the bug right now. I need to submit this paper next week."

"Yes, the paper is online, the code is not."

"Haha, you are trying to get **this** thing running?! We deprecated that stuff 2 years ago!"

"No. The person who has worked on that is now at [insert uni here]. Sorry mate."

"Documentation?! Look at the code!"

*We are talking about **prototype** systems, bleeding-edge tech, and the preliminary work of PhD candidates, mostly. I don't want to undermine the fantastic work I've also seen along the way!

"**The**" abstract 2014

Abstract. [PRE PRINT VERSION] Research on robot systems either integrating a large number of capabilities in a single architecture or displaying outstanding performance in a single domain achieved considerable progress over the last years. Results are typically validated through experimental evaluation or demonstrated live, e.g., at robotics competitions. While common robot hardware, simulation and programming platforms yield an improved basis, many of the described experiments still cannot be reproduced easily by interested researchers to confirm the reported findings. We consider this a critical challenge for experimental robotics. Hence, we address this problem with a novel process which facilitates the reproduction of robotics experiments. We identify major obstacles to experiment replication and introduce an integrated approach that allows (i) aggregation and discovery of required research artifacts, (ii) automated software build and deployment, as well as (iii) experiment description, repeatable execution and evaluation. We explain the usage of the introduced process along an exemplary robotics experiment and discuss our approach in the context of current ecosystems for robot programming and simulation.

Keywords: Software Engineering, Experimental Robotics, Development Process, Semantic Web, Continuous Integration, Software Deployment

Publications since then

Selected Publications

+ Invited Talks [Lincoln, Oxford, IIT, …]

2018 | Konferenzbeitrag | PUB-ID: 2916792

Geographically distributed deployment of reproducible HRI experiments in an interdisciplinary research context

Lücking P, Lier F, Bernotat J, Wachsmuth S, Sabanovic S, Eyssel FA (In Press)

In: 2018 ACM/IEEE International Conference on Human-Robot Interaction Companion. ACM/IEEE International Conference on Human-Robot Interaction Companion. Chicago: ACM/IEEE.

PUB | **DOI**

2017 | Kurzbeitrag Konferenz | PUB-ID: 2910475

Can we Reproduce it? Toward the Implementation of good Experimental Methodology in Interdisciplinary Robotics Research

Lier F, Lücking P, de Leeuw J, Wachsmuth S, Šabanović S, Goldstone R (2017)

In: ICRA 2017 Workshop on Reproducible Research in Robotics: Current Status and Road Ahead. Bonsignorio FP (Ed); Proceedings of IEEE International Conference on Robotics and Automation (ICRA), IEEE Xplore. IEEE.

PUB | **PDF**

2016 | Konferenzbeitrag | PUB-ID: 2901240

Towards Sustainable Robotics System Benchmarking

Weisz J, Huang Y, Lier F, Sethumadhavan S, Allen P (2016)

In: 2016 IEEE International Conference on Robotics and Automation. IEEE International Conference on Robotics and Automation (ICRA). .

PUB

2016 | Konferenzbeitrag | PUB-ID: 2904605

Towards Automated System and Experiment Reproduction in Robotics

Lier F, Hanheide M, Natale L, Schulz S, Weisz J, Wachsmuth S, Wrede S (2016)

In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Burgard W (Ed); International Conference on Intelligent Robots and Systems (IROS). Daejeon, Korea: IEEE.

PUB | **PDF**

2014 | Kurzbeitrag Konferenz | PUB-ID: 2702526

The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments (POSTER)

Lier F, Wienke J, Nordmann A, Wachsmuth S, Wrede S (2014)

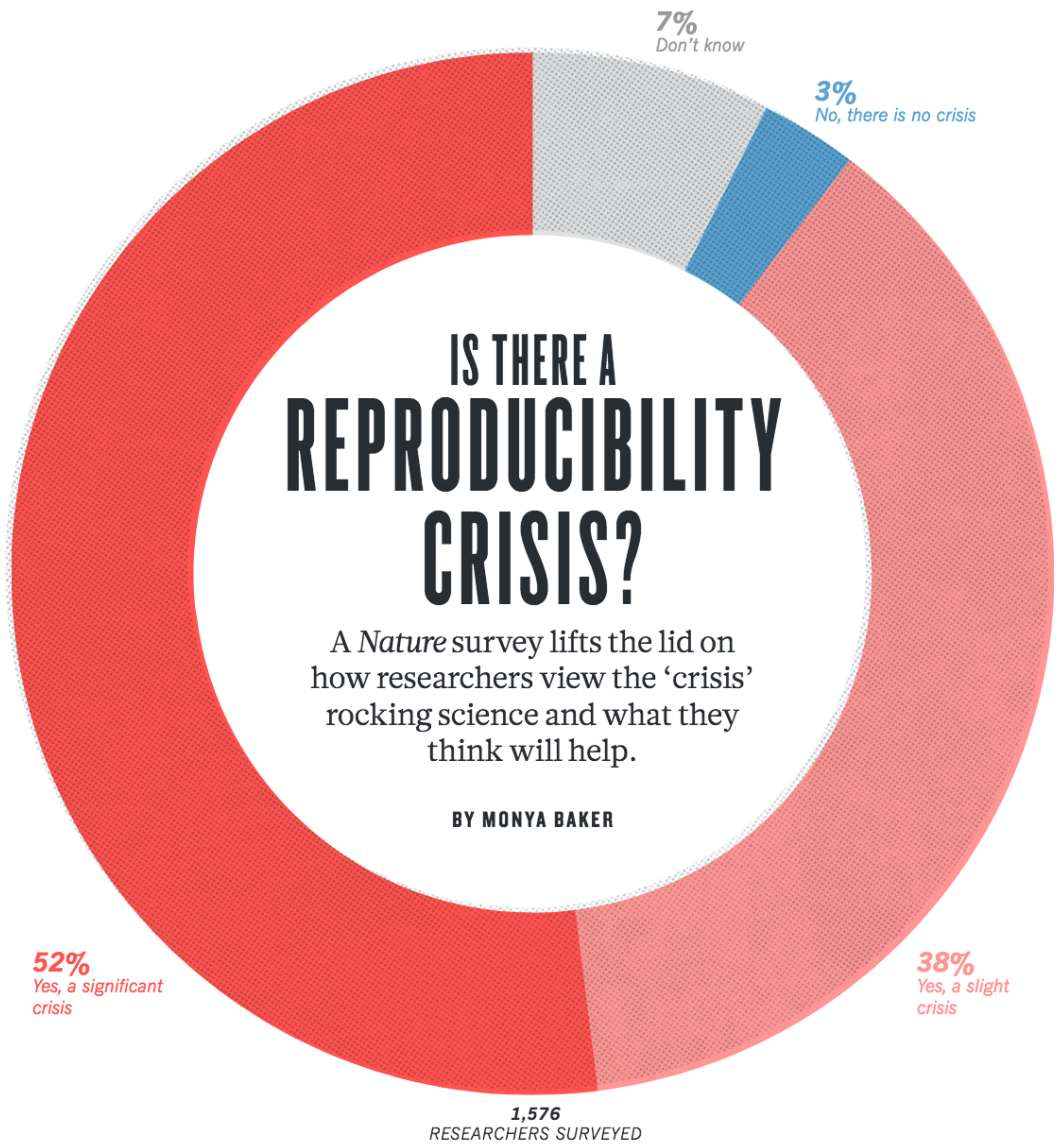
Presented at the International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR), Bergamo, Italy.

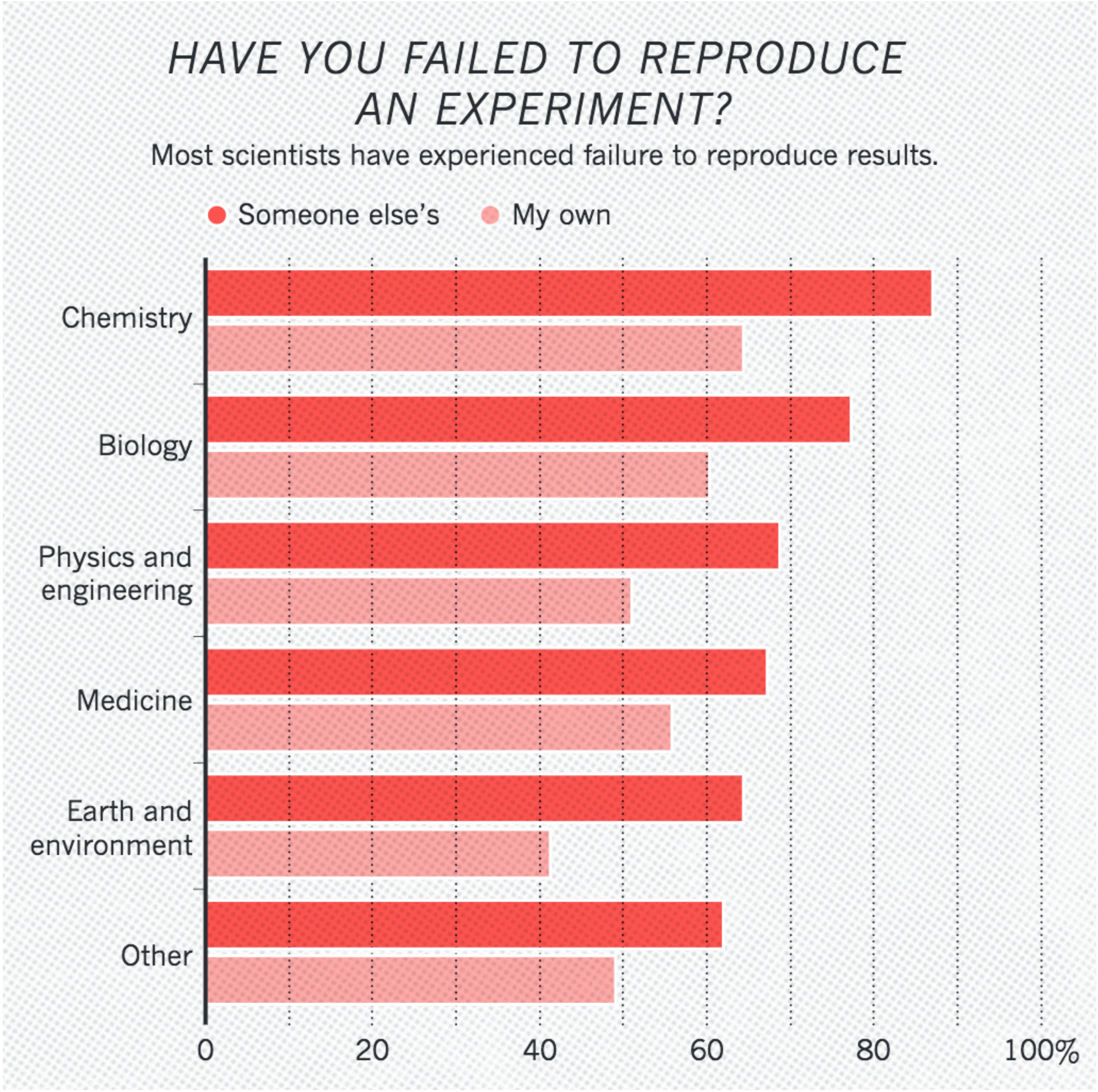
PUB | **PDF**

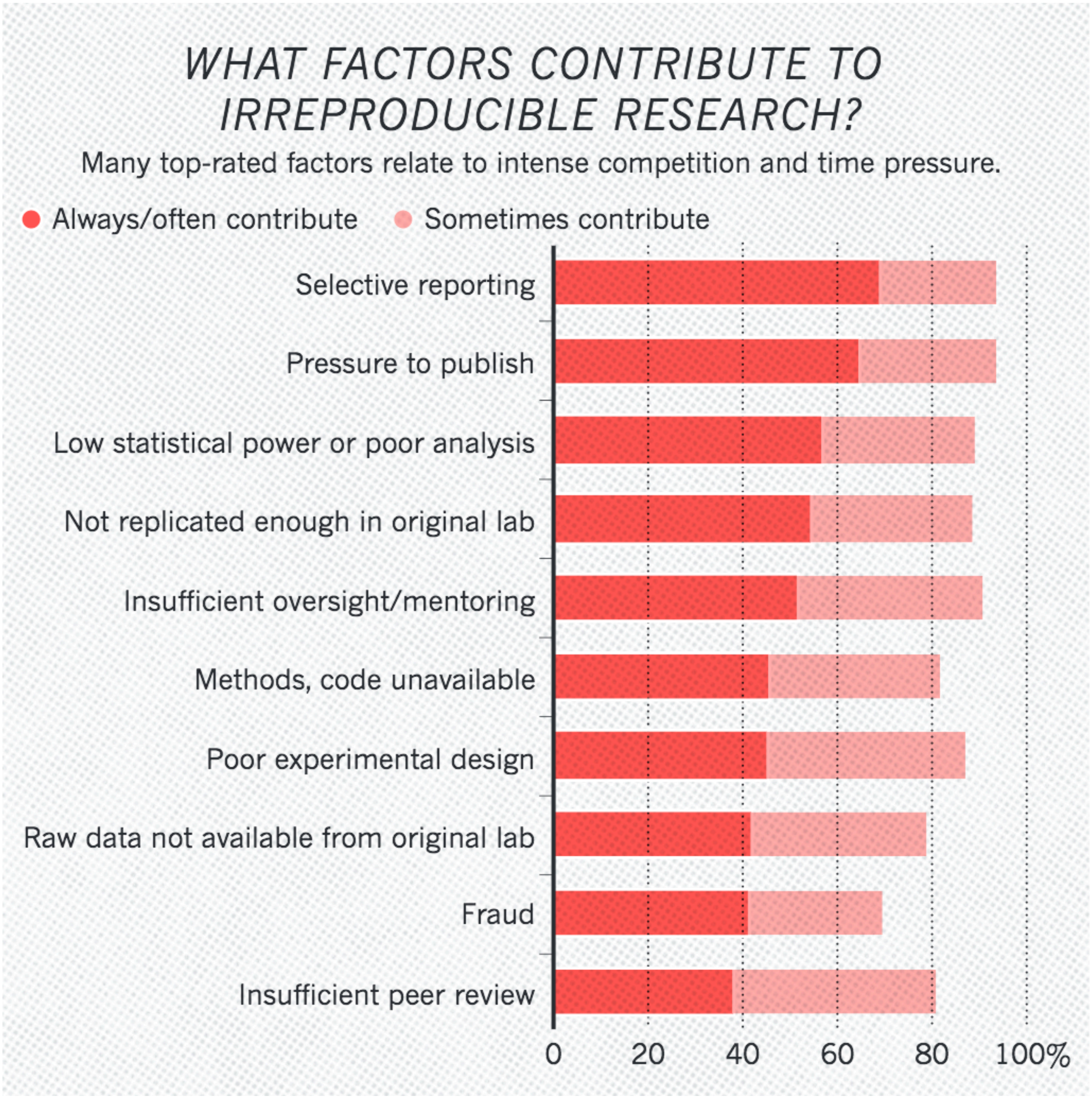
W e   w e r e   **r i g h t** ,   a n d   a h e a d   o f   t i m e …

THE DEBATE ABOUT A CRISIS

These figures emerged from Nature's survey of 1,576 researchers who took a brief online questionnaire on reproducibility in research (2016)







https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

Finally, our **big** **5 issues**

The big 5

i) **Information retrieval and aggregation:** Publications and associated artifacts relevant for reproduction, e.g., software components, data sets, documentation, related publications, and hardware used are often distributed over different locations, like digital libraries and other repositories. Hence, already the discovery, identification and aggregation of all required artifacts is often difficult. Furthermore, this kind of information is typically not available in a machine interpretable representation.

Lier et al. The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments

The big 5

ii) **Semantic relationships:** Often, crucial relationships between artifacts are unknown or underspecified, e.g., what specific *versions* (`master` or 1.33.7) of software components in combination with what data set, hardware, and experiment variant were in use for a particular study? It is obvious that a researcher can not publish every detail of a study in a regular paper. However, it is challenging to provide and make all these artifacts available and directly applicable.

Lier et al. The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments

The big 5

iii) **Software deployment:** Most current robotic systems are realized using a component-based architecture [90, 46, 25, 105], these components are usually not written in the same language. Consequently, they do not make use of the same build infrastructure[9], deployment mechanism, and execution environment. Therefore, it is an inherently complex and labor-intensive task to build and distribute a system in order to reproduce experimental results based on a software system. This becomes even more complex when experiments require software artifacts from more than one ecosystem: there is usually no cross-ecosystem integration model.

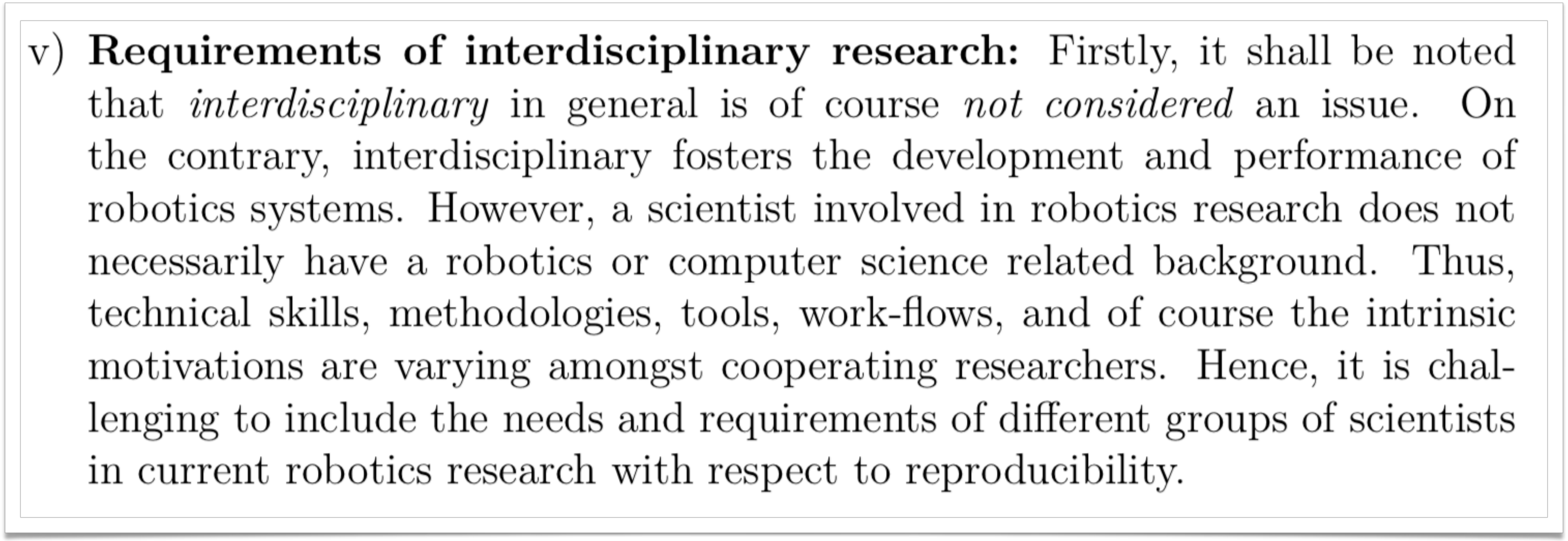
Lier et al. The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments

The big 5

iv) **Experiment testing, execution and evaluation:** Advanced robotics experiments require significant efforts spent on system development, integration testing, execution, evaluation and preservation of results — especially under consideration of the hardware-in-the-loop requirement. This is particular costly if many of these tasks are carried out manually, which is intriguing, as established methods from software engineering are available to *automate* these tasks, e.g., based on the continuous integration (CI) paradigm [31]. These techniques are not widely adopted by the large part of the robotics community for the iterative design, automated execution and ensured repeatability of robotics experiments.

Lier et al. The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments

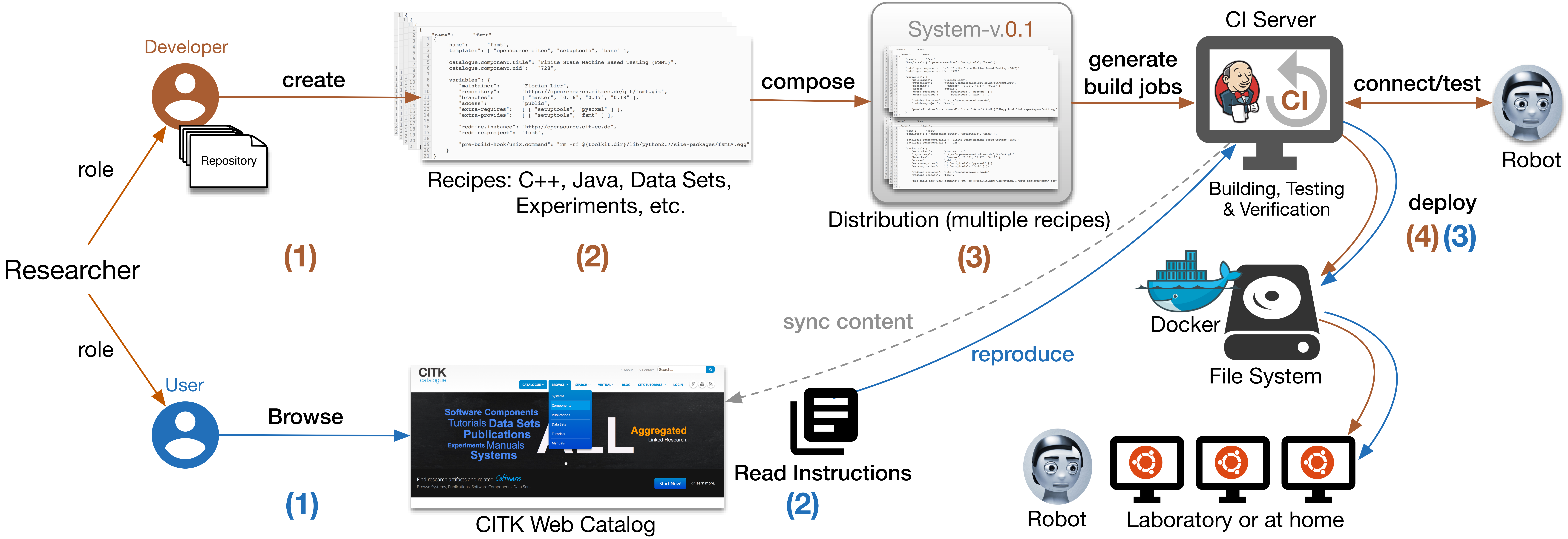
The big 5



v) **Requirements of interdisciplinary research:** Firstly, it shall be noted that *interdisciplinary* in general is of course *not considered* an issue. On the contrary, interdisciplinary fosters the development and performance of robotics systems. However, a scientist involved in robotics research does not necessarily have a robotics or computer science related background. Thus, technical skills, methodologies, tools, work-flows, and of course the intrinsic motivations are varying amongst cooperating researchers. Hence, it is challenging to include the needs and requirements of different groups of scientists in current robotics research with respect to reproducibility.

Lier et al. The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments

C I T K   A p p r o a c h

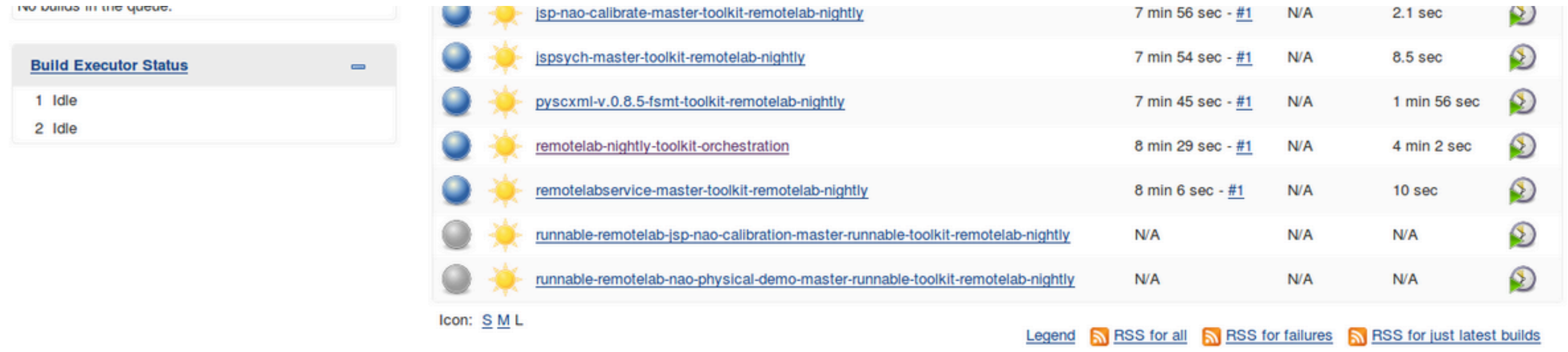


I will skip some technical details for the sake of time. Please talk to me after the presentation.

Lier et al. The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments, 2014

Lier et al. Towards Automated System and Experiment Reproduction in Robotics, 2016

Example Recipe **Example Distribution**



There will be **TWO** grey jobs:

- a) "runnable-remotelab-jsp-nao-calibration-master-runnable-toolkit-remotelab-nightly"
- b) "runnable-remotelab-nao-physical-demo-master-runnable-toolkit-remotelab-nightly"

These jobs will be used later on to:

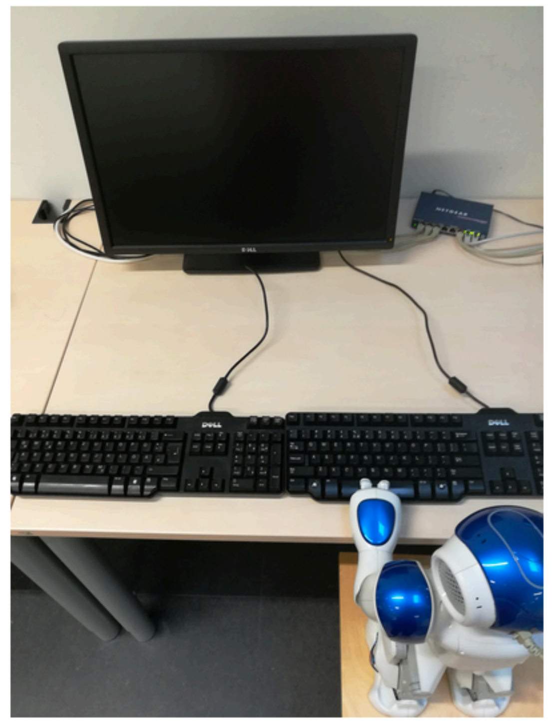
- a) actually **CALIBRATE** the robot
- b) **RUN** your experiment!

How *cool* is that? We will now setup the physical part of the experiment.

Step 4: Physical Experiment Setup

Because two NAOS were available at Bielefeld, a symmetrical setup was originally set up in an otherwise empty office.

Having two NAOs available is not given in every laboratory, thus we will describe a setup using just **one** NAO in the following. The setup is easily adjustable by moving the robot from one side to another depending on the chosen position (see Step 6) of the subject (human).



```
head -1 ord_<any_id>.csv > merged/merged.csv
```

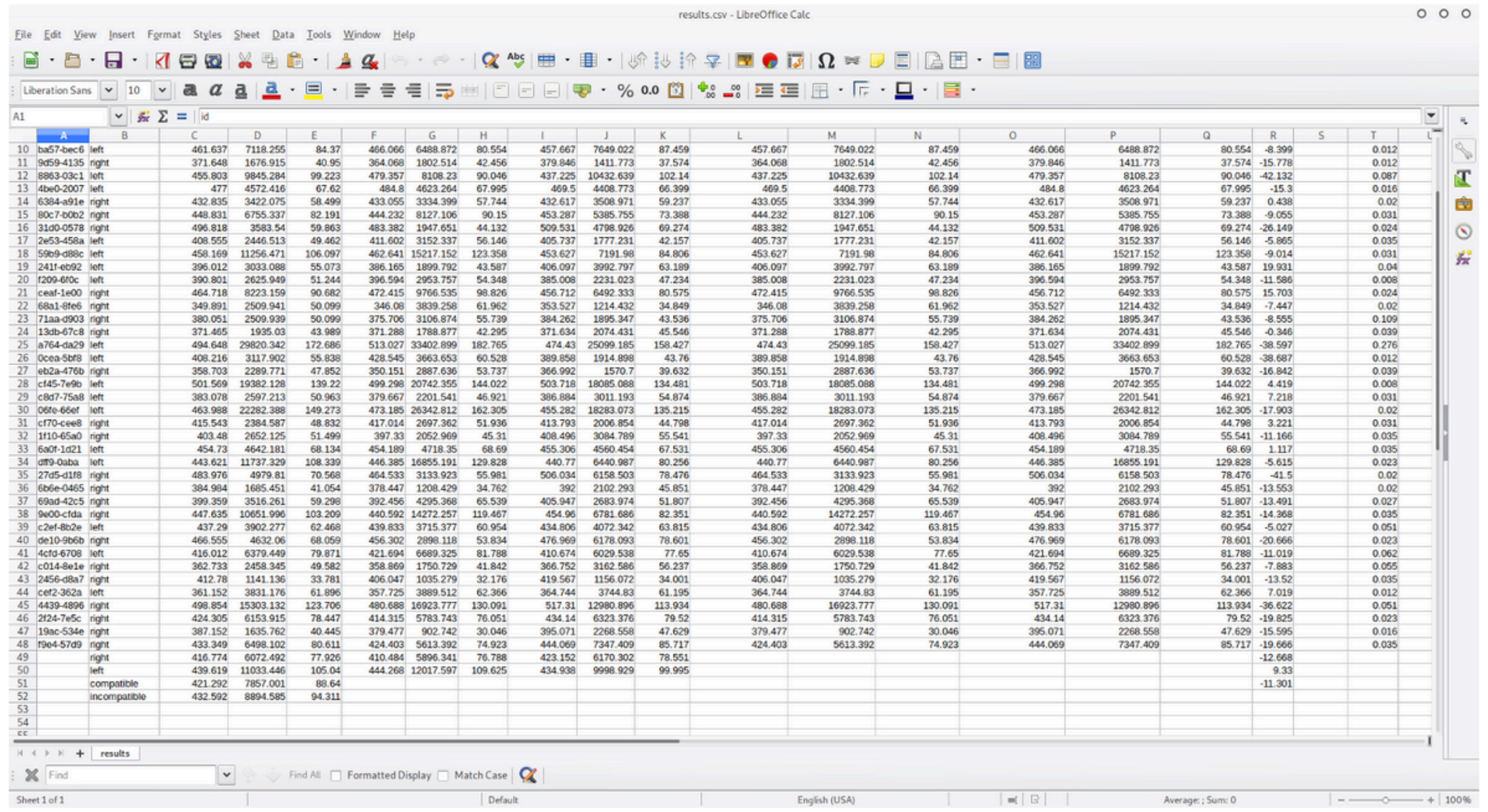
9. Finally, to merge all files into one type:

```
for filename in $(ls ord_*.csv); do sed 1d $filename >> merged/merged.csv; done
```

10. Now you got all data in one file - Congratulations!
11. To create a .csv with basic statistics type:

```
python2 nao_jse_reader.py -f merged/merged.csv
```

12. It will create the results.csv where you can see the mean/var/stdev for every ID and at the bottom the mean for the the two groups *compatible* and *incompatible*
13. Now you can do an in-depth analysis with e.g. R or SPSS!



Literature

- [1] https://www.ncbi.nlm.nih.gov/pubmed/22866762
- [2] http://journal.frontiersin.org/article/10.3389/fpsyg.2014.00974/full
- [3] https://pub.uni-bielefeld.de/publication/2910475
- [4] https://pub.uni-bielefeld.de/publication/2904605
- [5] https://link.springer.com/article/10.3758/s13428-014-0458-y

Tools vs. Use-Case vs. Name

CI+DEV	**REPRO**
Jenkins tar gz *	* Jenkins tar gz
Recipes (*.proj + *.distro) *	* Recipes (*.proj + *.distro + *.exp + *.hw)
Generator *	* Generator
(CITKAT) *	* FSMT
	* CITKAT
	* JSPsych

Idea: The CITK is a collection of tools + dev process (aligned with CI paradigm) that fosters reproducibility of software intensive experiments as introduced earlier (as of now).

Problem: The name CITK is *misleading* (multiple reviews and comments confirmed)

Situation: Find a better name (?) move the toolset to GitHub in order to:

a) Improve the quality of systems and recipes

b) Higher visibility in the community, set target domain, publications of multiple aspects

c) Improve development process (PR vs. force push, forks, dedicated roles via teams)

d) Consolidation / structure (folders or hierarchy)

The ultimative (old) goal

EXAMPLE(!)

```
$ r2dtk --bootstrap jenkins
$ r2dtk --bootstrap recipes
$ r2dtk --run jenkins
$ r2dtk --install-experiment meka_grasping
$ r2dtk --run-experiment meka_grasping
$ r2dtk --bootstrap catalog
$ r2dtk --publish-catalog meka_grasping
```

But **WHAT** will be our focus? Because the focus will have **implications** for the transition to GitHub

Implications

CI+DEV

CITK is "Bielefeld" *
Split recipes, templates, distros *
Introduce new dev pipeline: branches, forks, new repos *

REPRO

* Include experiments + hardware desc.
* Carefully version systems and experiments for citations + generate catalog (DOI!)
* Easy setup for non-tech researchers, e.g,
  ! clone recipes, clone templates, download generator x
  …

Some stats since 2015

Distributions	153
Components, Datasets, etc.	1169
Revisions	26640
Cooperations / Collaborators **IIT, Columbia U, UTGV, Indiana U, …**	> 10
Publications	11
Developers	2.5