

Correct answers: who needs them?

a story of numerical computing

RNT program, Isaac Newton Institute for Mathematical Sciences, Cambridge, UK

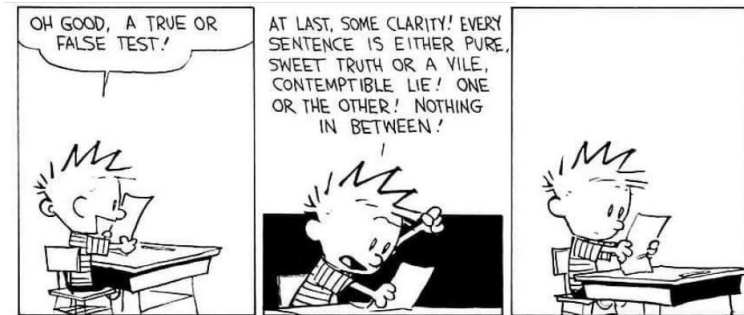
25 May 2023

www.newton.ac.uk/seminar/39442

Andy Adler

Carleton University, Ottawa, Canada

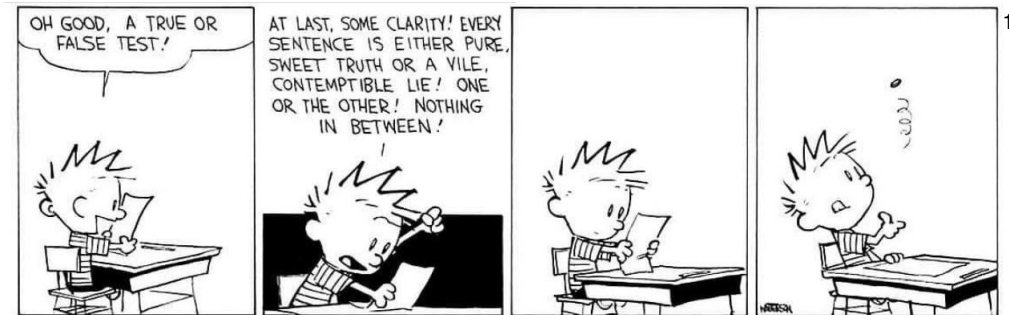
Correct?



1

¹Watterson, "The Days Are Just Packed", 1993

Correct?

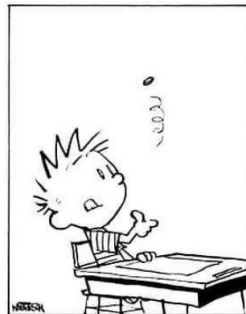


¹ Watterson, "The Days Are Just Packed", 1993

Correct?

Answer Correct?

How do you know?



Correct?

Correct

≡ free from error; in accordance with fact or truth

Correct?

Correct

\triangleq free from error; in accordance with fact or truth

... distinct from a lucky guess, because of

Justification

Correct?

Correct

≡ free from error; in accordance with fact or truth

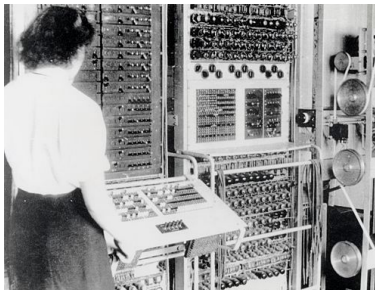
... distinct from a lucky guess, because of

Justification

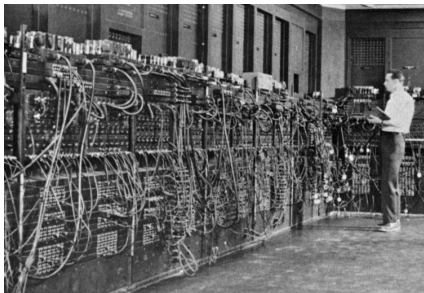
≡ reasons for a rationally admissible belief

... makes belief *knowledge* instead of opinion.

Numerical computing



Colossus (1943)

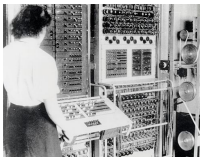


ENIAC (1945)

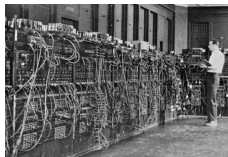
- **Colossus**: first electronic digital programmable computing device, used to break German ciphers during World War II
- **ENIAC**: designed to calculate artillery firing tables; first used for feasibility of thermonuclear weapons

²Colossus and ENIAC, Wikipedia.org

“Digital” vs Numerical computing



Colossus (1943)

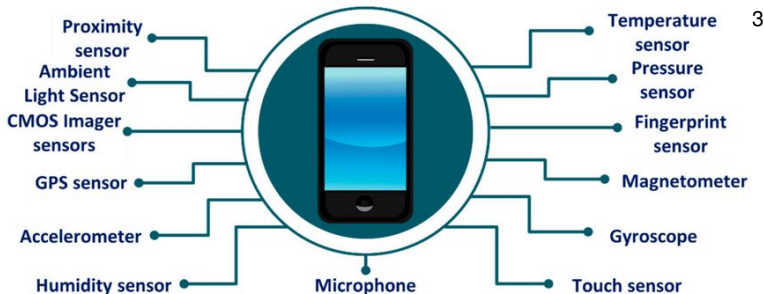


ENIAC (1945)

	<i>Digital</i>	<i>Numerical</i>
<i>Operations</i>	Integer	Floats
<i>Field</i>	Computer Science	Engineering, Applied Math
<i>Errors</i>	Yes/No	Distribution

Until recently, most computing was digital: office, banking, games, networking, publications, government, security

Ubiquitous sensors



- Sensor's give approximate data: must be interpreted.
- Cheap sensors + Powerful μC \rightarrow Great combination

³Majumder & Deen, "Smartphone Sensors for Health Monitoring and Diagnosis", Sensors 19:2164, 2019.

Why numerical algorithms

Scene optimization:

For example, if you take a photo of a person in front of a grassy field with blue sky in the frame, your phone's image signal processor could brighten up their face given they're probably the subject, boost the greens in the grass to make them look richer and enhance the blues in the sky ...

...[i.e. Google's] Pixel, a phone with no optical image stabilization, but with such good electronic stabilization that it outperformed much of the competition.⁴

⁴ [Techradar.com](https://www.techradar.com)

Image stabilization

≡ techniques to reduce blurring associated with the motion of a camera

- Optical image stabilization
- Lens-based
- Sensor-shift
- Digital image stabilization
 - take multiple exposures
 - reject blurred ones
 - transform and average the remaining ones

What went wrong?



5

Panning the camera to follow a moving cyclist

- Numerical algorithms are often used to solve inverse problems.
- These can fail when assumptions (*priors*) are not valid

⁵expertphotography.com/image-stabilization. Advice is: Turn off the stabilization feature if you are intentionally moving the camera. Some image stabilization systems can figure out what you are doing and switch itself off. Otherwise, image stabilization works against your panning efforts.

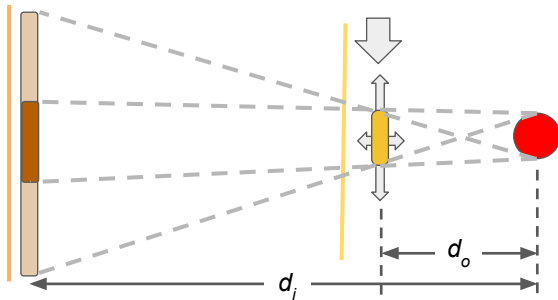
Inverse problems & Plato's cave

6



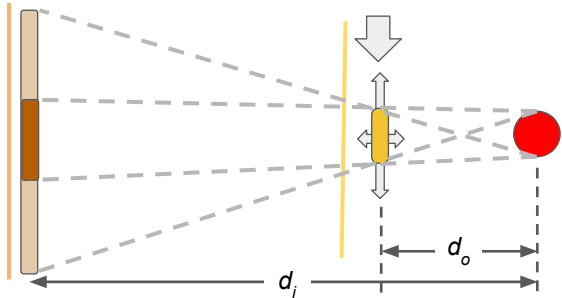
⁶Plato's cave, [wikipedia.org](https://www.wikipedia.org), inspired by C Groetsch "Linear Inverse Problems" in O Schertzer *Handbook of Mathematical Methods in Imaging*

Inverse problems & Plato's cave



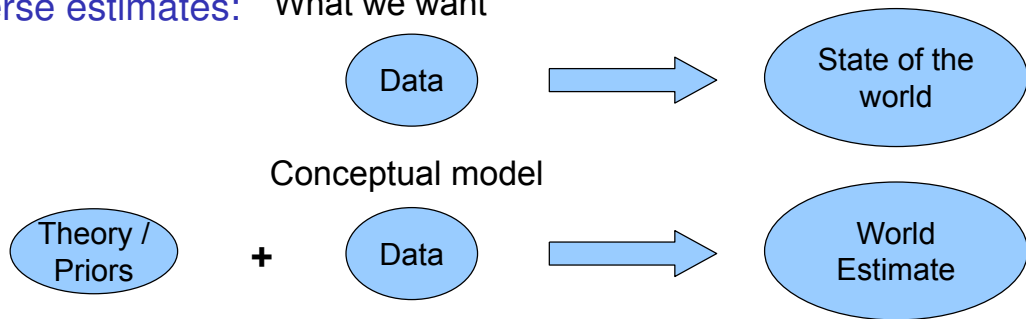
- Sensitivity
- Noise
- Null space

Inverse problems & Plato's cave

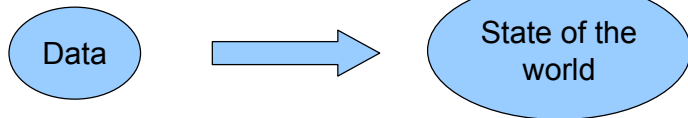


- Sensitivity
- Noise
- Null space

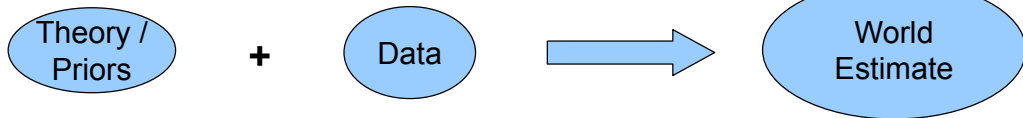
Inverse estimates: What we want



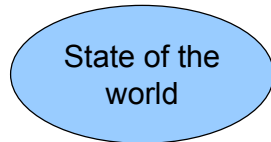
Inverse estimates: What we want



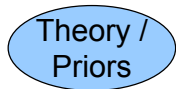
Conceptual model $p(\theta|d) \propto p(d|\theta)p(d)$



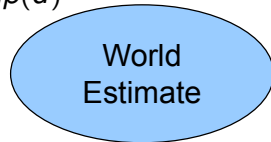
Inverse estimates: What we want
... as implemented



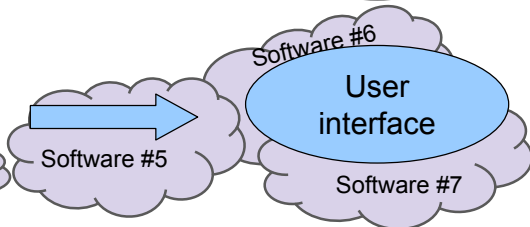
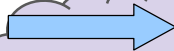
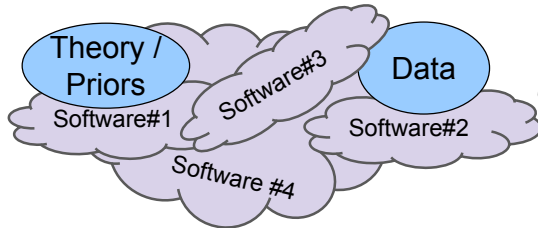
Conceptual model $p(\theta|d) \propto p(d|\theta)p(d)$



+



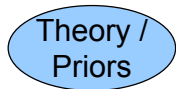
What we have



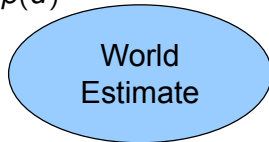
Inverse estimates: What we want
... as implemented



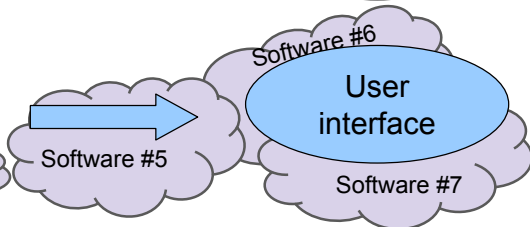
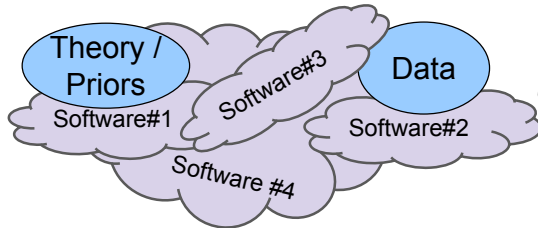
Conceptual model $p(\theta|d) \propto p(d|\theta)p(d)$



+



What we have



$$p(\theta|d) \propto p(d|\theta, \text{Software})p(d|\text{Software})$$

Software from a recent project



Python3 Packages:

```
localreg==0.5.0
matplotlib==3.5.1
neurokit2==0.2.4
numpy==1.21.5
pandas==2.0.1
Pillow==9.0.1
Pillow==9.5.0
pyparsing==2.4.7
scipy==1.8.0
skimage==0.0
wfdb==4.1.1
```

libpython3.10-stdlib:

```
libpython3.10-minimal
libbz2-1.0
libc6 (>= 2.34)
libcrypt1 (>= 1:4.1.0)
libdb5.3
libffi8 (>= 3.4)
liblzma5 (>= 5.1.1alpha+20120614)
libmpdec3
libncursesw6 (>= 6.1)
libnsl2 (>= 1.0)
libreadline8 (>= 7.0~beta)
libsqlite3-0 (>= 3.14.0)
libtinfo6 (>= 6)
libtirpc3 (>= 1.0.2)
libuuid1 (>= 2.20.1)
```

Numpy Packages: python3-numpy

```
python3-pkg-resources
python3 (<< 3.11)
python3 (>= 3.10~)
```

```
libblas3
libatlas3-base
libblas3
libblis3-openmp
libblis3-pthread
libblis3-serial
libopenblas0-openmp
libopenblas0-pthread
libopenblas0-serial
```

```
libc6 (>= 2.35)
```

```
liblapack3: <liblapack.so.3>
```

```
libatlas3-base
libc6 (>= 2.29)
libgcc-s1 (>= 4.0)
libgfortran5 (>= 8)
liblapack3
libopenblas0-openmp
libopenblas0-pthread
libopenblas0-serial
```

```
libc6 (>= 2.29)
```

```
libgcc-s1 (>= 4.0)
```

```
libgfortran5 (>= 8)
```

ATLAS

(Automatically
Tuned Linear
Algebra Software)

BLAS (Basic Linear
Algebra
Subprograms)

Open source software (OSS)

≡ computer software released under a license which grants users the rights to use, study, change, and distribute the software and its source code⁷

- core components of all network, security, image processing is OSS
- *example*: JPEG quality (0–100) is not part of the standard. Defined by Independent JPEG group's SW⁸
- Part of open science . . . discussed later

⁸[wikipedia.org/wiki/Open-source_software](https://en.wikipedia.org/wiki/Open-source_software)

⁸[ijg.org](https://www.ijg.org). First release 1991

BLAS: Basic Linear Algebra Subprograms

Basic Linear Algebra Subprograms

(**BLAS**) is a [specification](#) that prescribes a set of low-level routines for performing common [linear algebra](#) operations such as [vector](#) addition, [scalar multiplication](#), [dot products](#), linear combinations, and [matrix multiplication](#).

BLAS implementations are often optimized for speed on a particular machine,


It originated as a Fortran library in 1979 and its interface was standardized by the BLAS Technical (BLAST) Forum

file	blas.tgz	Fortran77 reference implementation
prec		single, double, complex, doublecomplex
lang		fortran
#		DOCUMENTATION AND TEST SUITES
file	blasqr.ps	quick reference guide for the BLAS.
file	blas3-paper.ps	details on the Level 3 BLAS
file	blas2test.f	original test driver for the real level two blas
file	sblat1.f	Level 1 BLAS Test Suite (prec. single)
file	dlat1.f	Level 1 BLAS Test Suite (prec. double)
...

Tests

⁹ [wikipedia.org](#) and [netlib.org](#)

← → ↺ <https://eidors3d.sourceforge.net/index.shtml> ☆ ☰



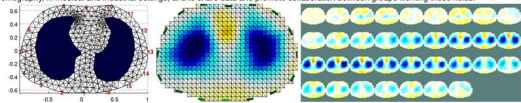
EIDORS: Electrical Impedance Tomography and Diffuse Optical Tomography Reconstruction Software

[EIDORS \(minor\)](#)
[Main](#)
[Documentation](#)
[Tutorials](#)
[Download](#)
[Contrib Data](#)
[GREIT](#)
[Browse Docs](#)
[Browse SVN](#)

[News](#)
[Mailing list \(archive\)](#)
[FAQ](#)
[Developer](#)

Project Goal

Provide free software algorithms for forward and inverse modelling for Electrical Impedance Tomography (EIT) and Diffusion based Optical Tomography, in medical and industrial settings, and to share data and promote collaboration between groups working these fields.



FEM model (left) and reconstructed image (centre) and sequence (right) from [human shallow breathing](#)

Requirements

- [Matlab](#) (≥2016b or 9.1) or [Octave](#) (≥7.3),
- [Netgen Mesher](#) (optional), (Direct link to [Netgen 5.3](#))
- [Gmsh Mesher](#) (optional)

Getting Started

To try the EIDORS software, follow these steps:

1. Download EIDORS :
 - Release (30 December 2022): [EIDORS 3.11](#), or [EIDORS 3.11 with Netgen 5.3](#) (for windows 64 bit). Or [Developer Version](#)
 - Optional ready-made FEM models: [model_library](#) (7 Jul 2012)
2. Unzip the software in a directory such as `/path/to/eidors(UNIX)` or `C:\path\to\eidors(Windows)` (installation instructions for `model_library` are included in a README file)
3. Download and install [Netgen](#) (Unless version with netgen installed, EIDORS will ask for the path to netgen when required)
4. Start Matlab
5. In the Matlab command window type: `>>run /path/to/eidors/startup.m`
(windows `>>run C:/path/to/eidors/startup.m`)
6. Try the [Tutorials](#), or execute one of the sample programs in the `/path/to/eidors/examples` directory (such as `compare_2d_algs(1)`)

Getting Help

For questions or help, search the [\[EIDORS-help\] Mailing list](#). Email requests for help should be sent to eidors3d-help@lists.sourceforge.net. Sign up for [eidors3d-help here](#).

Make sure your request answers the questions:

- What specifically did you do (so that it can be repeated)?
- What version of EIDORS and Matlab/OS are you using?
- What, specifically, did EIDORS do that you didn't want it to?

If you are a student, please state your full name, institution and department (or research group) and CC your supervisor on your emails.

License

EIDORS is licenced under the GNU [General Public License](#) (version 2 or 3). Users are free to use, modify, and distribute their modifications.

Overview

Uses other OSS


Build community
(Searchable answers)

Tutorials!

← → ↻ https://eidors3d.sourceforge.net/tutorial/GREIT/neonate_ex.shtml ⓘ ☆ ⌵ ☰

[Documentation](#)
[Tutorials](#)
[- Image Reconst](#)
[- Data Structures](#)
[- Applications](#)
[- FEM Modelling](#)
[- GREIT](#)
[- Old tutorials](#)
[- Workshop](#)
[Download](#)
[Contrib Data](#)
[GREIT](#)
[Browse Docs](#)
[Browse SVN](#)

[News](#)
[Mailing list](#)
[\(archive\)](#)
[FAQ](#)
[Developer](#)

Hosted by 

Data

Data are available [Here](#). Data were recorded from a 10-day old spontaneously breathing neonate lying in the prone position with the head turned to the left, as documented in: S. Heinrich, H. Schiffmann, A. Frerichs, A. Klockgether-Radke, I. Frerichs, [Body and head position effects on regional lung ventilation in infants an electrical impedance tomography study](#), Intensive Care Med., 32:1392-1398, 2006.

Subject Image:




Image reconstruction model

Forward model model

```
% Inverse model
%fmd1 = mk_library_model('neonate_16el_lungs');
elec_pos = [16,1,.5]; elec_shape=[0.15,0.3,0.01,0.60]; maxsz=0.08; nfft=27;
fmd1 = mk_library_model({'neonate','boundary','left_lung','right_lung'}, ...
    elec_pos, elec_shape, maxsz,nfft);

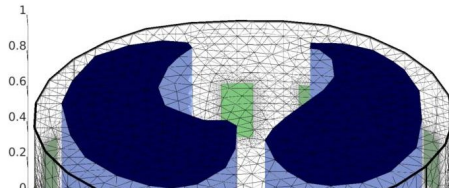
[fmd1.stimulation fmd1.meas_select] = mk_stim_patterns(16,1,'{ad}','{ad}');
fmd1 = mdl_normalize(fmd1,1);

img = mk_image(fmd1,1); img.elem_data(vertcat(fmd1.mat_idx(2:3))) = 0.3;
img.calc_colours.ref_level=1;

calc_colours('defaults');
show_fem_enhanced(img); view(-2,32)
print_convert neonate_ex01a.jpg
```

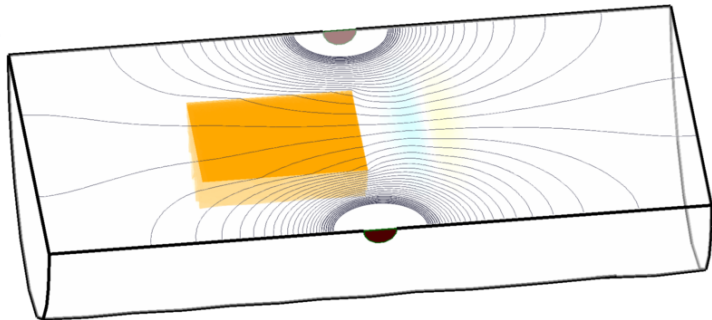
Users get started this way

(Often useful for authors)

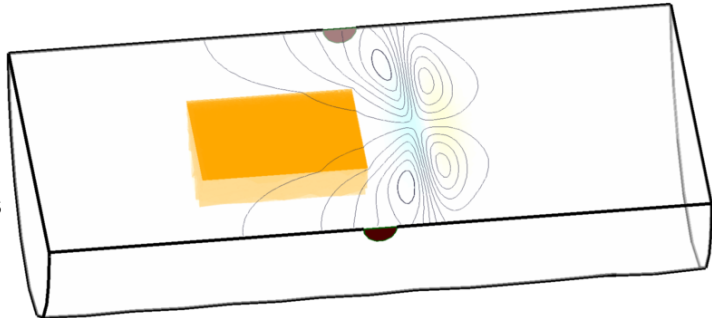


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

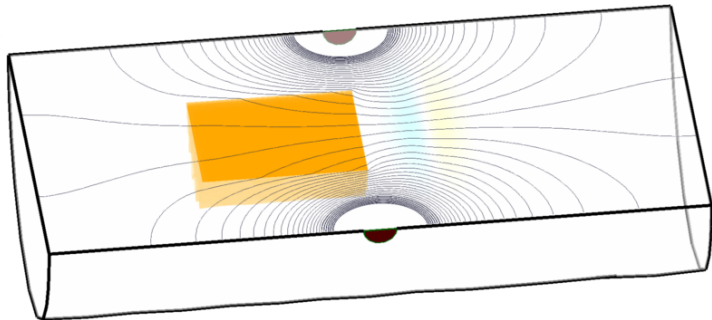


Δ Equipotentials

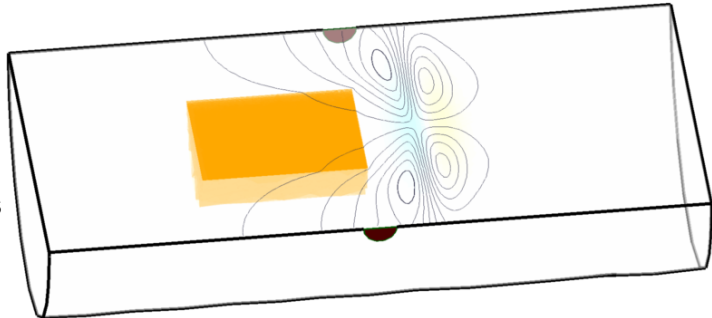


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

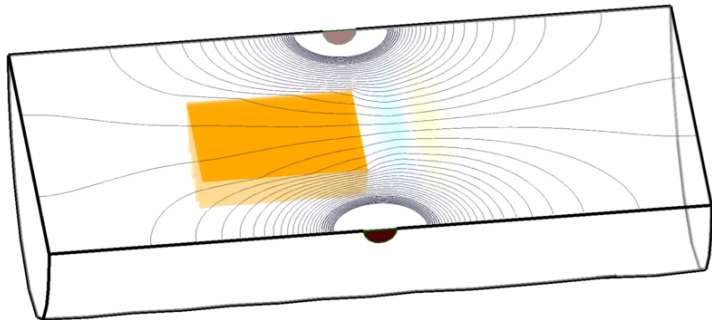


Δ Equipotentials

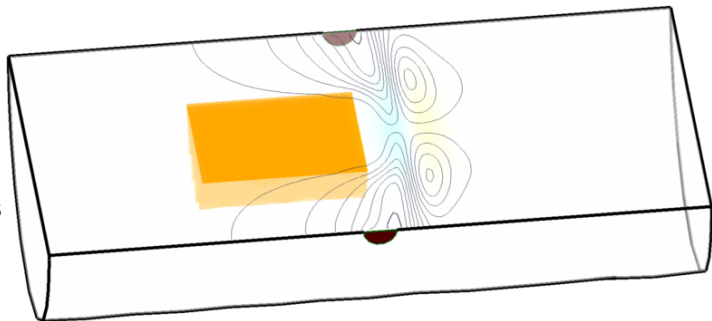


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

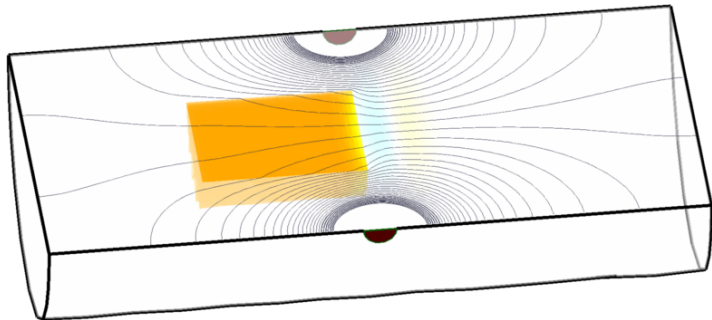


Δ Equipotentials

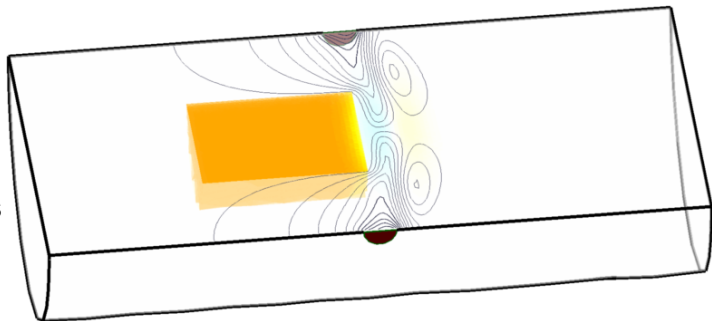


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

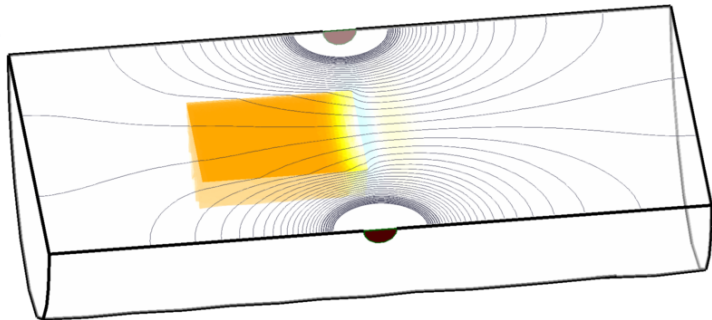


Δ Equipotentials

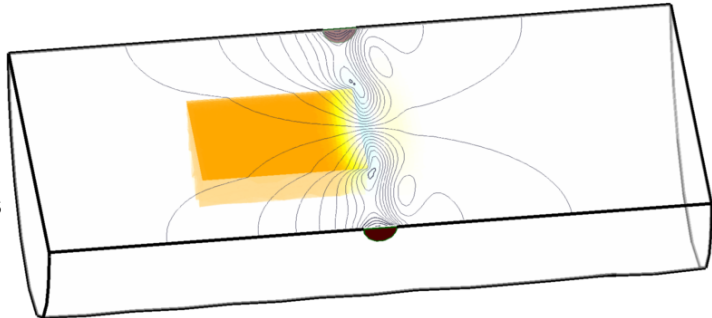


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

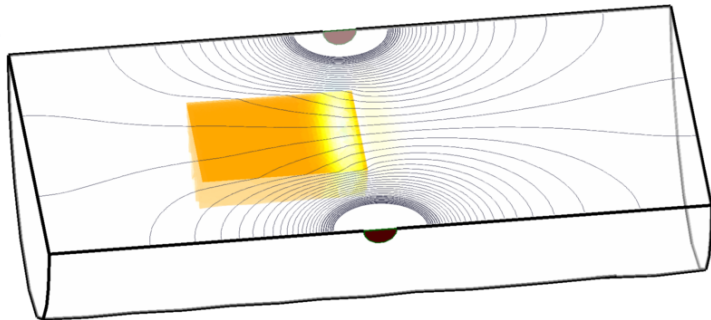


Δ Equipotentials

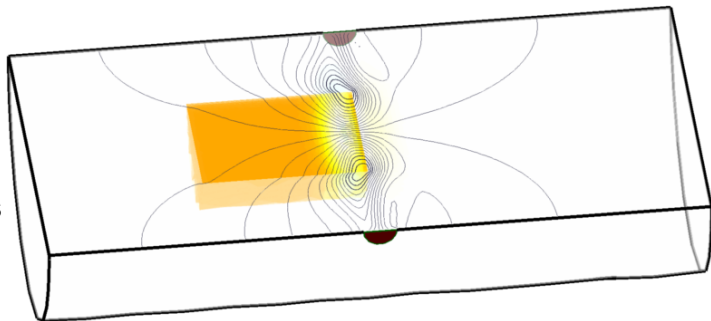


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

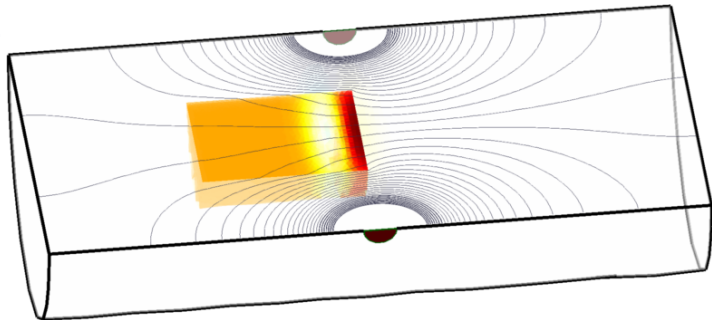


Δ Equipotentials

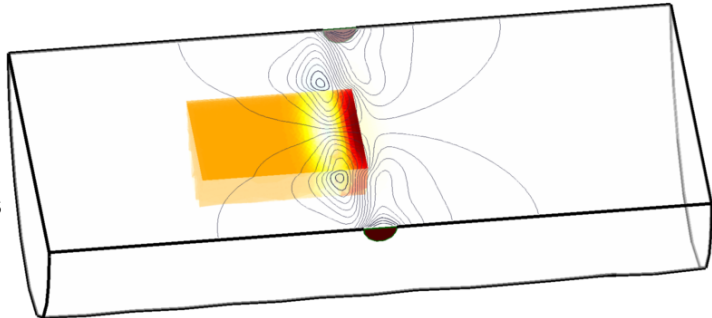


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

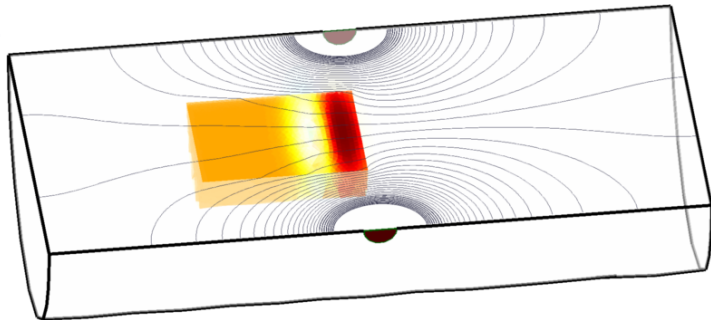


Δ Equipotentials

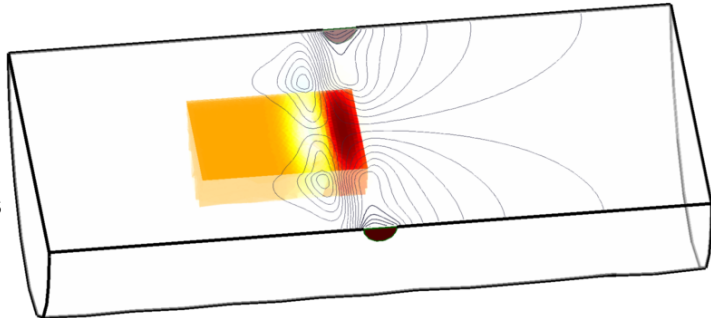


Idea: Combining Physics Acousto-Electric Tomography

Equipotentials

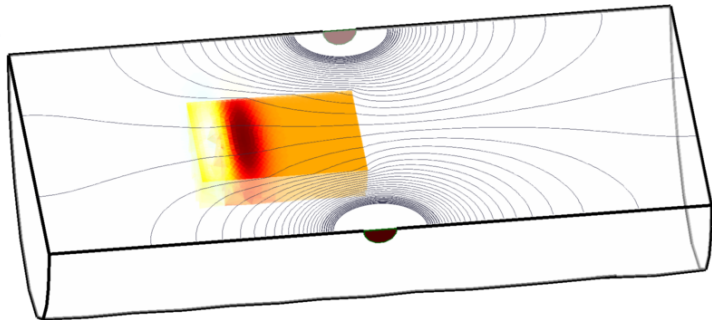


Δ Equipotentials

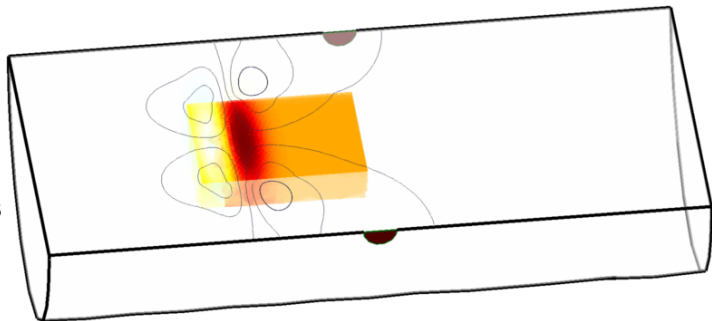


Idea: Combining Physics Acousto-Electric Tomography

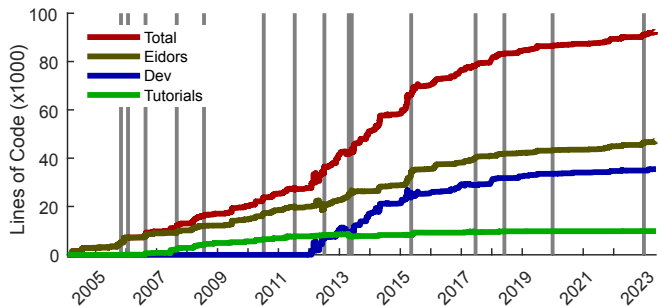
Equipotentials



Δ Equipotentials



EIDORS: experiences



- Contributions from community
- Managing growth
- Testing!!
- Toolchain rot (sourceforge, mailing lists, matlab, perl, html)

EIDORS: contributions and testing

- Contributions are great!
- Contributors will solve their own problem, but often won't test all cases
- Who checks?

On Thu, Feb 16, 2023 XXXX wrote:

> Dear Prof. Adler,

>

> I am a phd-student, who is dealing with EIT.

> Recently I published an article ...

>

> I would like to ask you whether you think it is a good idea to modify the code of my paper such that it can be integrated directly into EIDORS?

...



```
function data = fwd_solve_2p5d_1st_order(fwd_model, img)
% FWD_SOLVE_2P5D_1ST_ORDER: data= fwd_solve_2p5d_1st_order( img)

% -----
function do_unit_test
% -----

unit_test_cmp('h2a 2p5d values TEST', uvh, [ 0.0060; 0.0080; 0.0110; 0.0160; 0.0320 ], 1./scale);
tol = norm(vh.meas)*0.025; % 2.5% error tolerance
unit_test_cmp('2D (h2a) vs analytic TEST', norm(vh.meas - v2.meas), 0, -tol);
unit_test_cmp('2.5D (h2a + 2.5D, k=0) vs 2D TEST', norm(v2.meas - v2p5k0.meas), 0, tol);
unit_test_cmp('2.5D (h2a + 2.5D trapz, 2*tol) vs analytic TEST', norm(vh.meas - v2p5t.meas), 0, 2*tol);
unit_test_cmp('2.5D (h2a + 2.5D quadv**) vs analytic TEST', norm(vh.meas - v2p5q.meas), 0, tol);

plot([vh.meas, v2p5q.meas, v3.meas, v2.meas], ':');
legend('analytic', 'FEM 2.5D', 'FEM 3D', 'FEM 2D', 'Location', 'Best');
% -----
```

Software testing: a mature field in software engineering

Types of software testing

10

Software testing assesses the functionality of a software program.



REGRESSION TESTING

ensures whether the addition of new features causes a decline in the functionality of an application. It's typically repeated after each build.



UNIT TESTING

ensures each individual unit or component performs as expected. It's typically conducted during the app development phase.



FUNCTIONAL TESTING

checks each function against functional requirements. A black-box test is a common example.



INTEGRATION TESTING

groups together two or more modules of an application to make sure they can function collectively.



STRESS TESTING

assesses the strength of software by testing how much load it can take before reaching a breaking point.



SECURITY TESTING

ensures software is free of potential vulnerabilities, known flaws and security loopholes that might affect the user system and data.



PERFORMANCE TESTING

tests the performance, speed and scalability of an application under a given workload.



ACCEPTANCE TESTING

evaluates the entire system against the desired requirements to confirm project completion.

Software testing: a mature field in software engineering

EIDORS uses:

- Regression Tests:
Same as last time?

Types of software testing

Software testing assesses the functionality of a software program.



REGRESSION TESTING

ensures whether the addition of new features causes a decline in the functionality of an application. It's typically repeated after each build.



UNIT TESTING

ensures each individual unit or component performs as expected. It's typically conducted during the app development phase.



FUNCTIONAL TESTING

checks each function against functional requirements. A black-box test is a common example.



INTEGRATION TESTING

groups together two or more modules of an application to make sure they can function collectively.



STRESS TESTING

assesses the strength of software by testing how much load it can take before reaching a breaking point.



SECURITY TESTING

ensures software is free of potential vulnerabilities, known flaws and security loopholes that might affect the user system and data.



PERFORMANCE TESTING

tests the performance, speed and scalability of an application under a given workload.



ACCEPTANCE TESTING

evaluates the entire system against the desired requirements to confirm project completion.

Software testing: a mature field in software engineering

EIDORS uses:

- Regression Tests:
Same as last time?
- Unit Tests
Right answer?

Types of software testing

Software testing assesses the functionality of a software program.



REGRESSION TESTING

ensures whether the addition of new features causes a decline in the functionality of an application. It's typically repeated after each build.



UNIT TESTING

ensures each individual unit or component performs as expected. It's typically conducted during the app development phase.



FUNCTIONAL TESTING

checks each function against functional requirements. A black-box test is a common example.



INTEGRATION TESTING

groups together two or more modules of an application to make sure they can function collectively.



STRESS TESTING

assesses the strength of software by testing how much load it can take before reaching a breaking point.



SECURITY TESTING

ensures software is free of potential vulnerabilities, known flaws and security loopholes that might affect the user system and data.



PERFORMANCE TESTING

tests the performance, speed and scalability of an application under a given workload.



ACCEPTANCE TESTING

evaluates the entire system against the desired requirements to confirm project completion.

Review: Where are we?

Numerical Computing

- much more common with all the sensors
- solve difficult parameter estimation
- largely uses open algorithms, tools and software
- testing is difficult

Review: Where are we?

Numerical Computing

- much more common with all the sensors
- solve difficult parameter estimation
- largely uses open algorithms, tools and software
- testing is difficult

Questions:

- How are we doing?
- How common are errors?

Reproducibility Project: Psychology

△ a crowdsourced collaboration . . . to repeat 100 published experimental and correlational psychological studies . . . led by the Center for Open Science¹¹

Comments:

- Overall reproducibility is poor
- Let's not criticise psychology . . . they're brave
- Open science is the right way

¹¹ en.wikipedia.org/wiki/Reproducibility_Project

Can we implement a formula correctly?

- In APA reporting format, inferential statistics must report: the test statistic, the degrees of freedom for the test, and the p -value.
- values redundant with one another.
- can check consistency by evaluating whether they agree¹²
- p -values for more than 250,000 psychology papers checked; around half contained at least one incorrect p -value!¹³

¹²Frank *et al* *Experimentology: An Open Science Approach to Experimental Psychology Methods*, experimentology.io/

¹³Nuijten *et al*, "The prevalence of statistical reporting errors in psychology (1985–2013)", (2016) *Behavior Research Methods*. 48:1205–1226

How about reproducing complicated results?

- Hardwicke *et al*^{14,15} identified datasets with reusable data (not all complete and comprehensible)
- attempted to reproduce the main statistical results from 60 articles
- Very labour-intensive (5–10 hours of work each).
- Results
 - about 1/3 reproducible without help from the original authors
 - 62% reproduced after – sometimes extensive – correspondence
 - Many remaining appeared to have some irreproducible results

¹⁴Hardwicke *et al* “Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal *Cognition*” (2018) Royal Soc. open sci 5180448180448

¹⁵Hardwicke *et al* “Analytic reproducibility in articles receiving open data badges at the journal *Psychological Science: an observational study*” (2021) Royal Soc. open sci 8201494201494

Open science: Reproducible science¹⁶

		Data	
		Same	Different
Code	Same	Reproducible	
	Different		

¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science¹⁶

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different		

¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science¹⁶

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different	Robust	

¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science¹⁶

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different	Robust	Generalizable

- Data availability: zenodo, dataverse
- Code availability: github, or with data

¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Summary so far

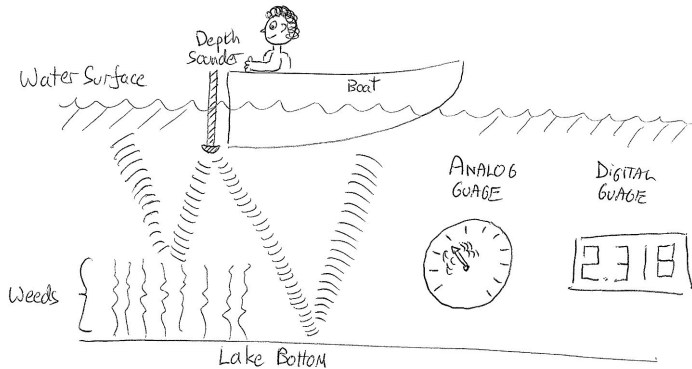
Numerical Computing

- increasingly common
- open algorithms, tools and software
- testing is difficult

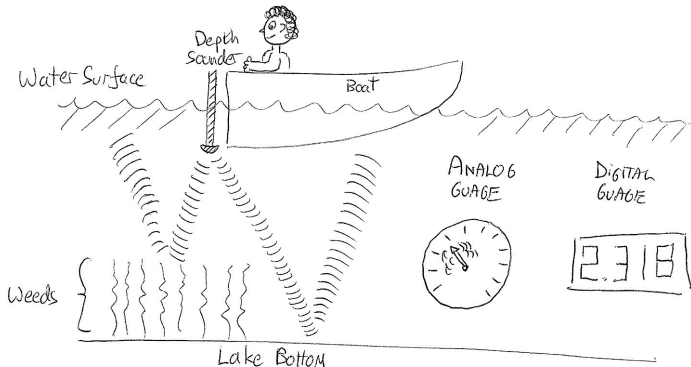
What does this mean for?

- Data
- Users
- Science

Data Quality



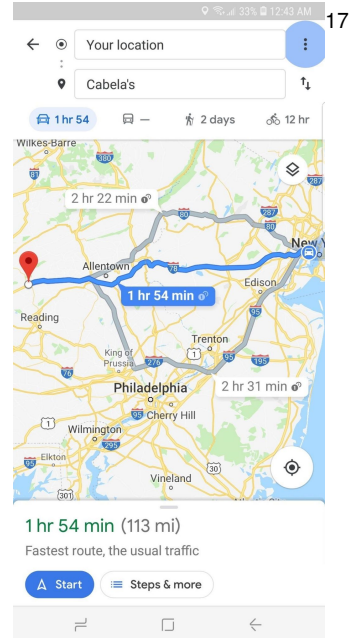
Data Quality



Data Quality

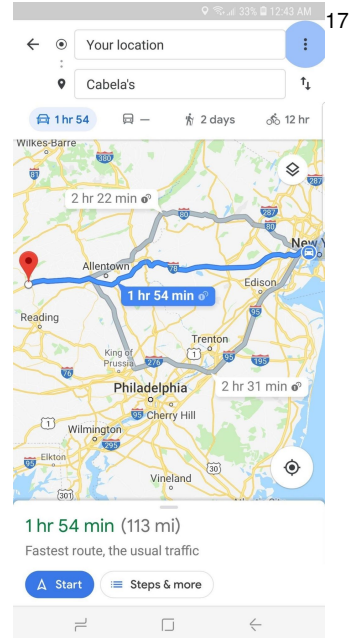
- Lots of bad data (e.g. Cut-and-paste errors: [\[Globe&Mail\]](#))
- ...
- How to write an prior for an “off-by-one paste” error
- How to determine if a pretty picture is valid

Thoughts on Numerical Algorithms: ... working with a user



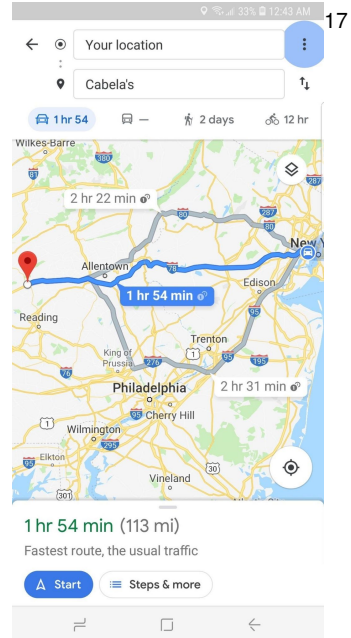
Thoughts on Numerical Algorithms: ... working with a user

- Good that they say: “the usual traffic”
 - But is this estimate for me
 - or is it an average for all drivers?



Thoughts on Numerical Algorithms: ... working with a user

- Good that they say: “the usual traffic”
 - But is this estimate for me
 - or is it an average for all drivers?
- How to communicate settings to a users for
 - Smart-phone camera
 - Intra-cardiac defibrillator
 - Self-driving car
 - Autonomous kill-drone



Thoughts on Open Science

Reproducibility is the
sine qua non of science

Why is it so hard?

Answer Correct?

How do you know?



¹⁸Reynolds, "The Git Hater's Guide"

Thoughts on Open Science

Reproducibility is the
sine qua non of science

Why is it so hard?

- Incentives
- Perfectionism → “I’ll upload this after I clean up the code”
- Hard-to-use tools → “Your Mindset: Git hates you. Git is really looking forward to destroying your code. Git is hoping you’ll let it destroy other people’s code . . .”¹⁸
- Rotting of tools and links
- Culture and Training

Answer Correct?

How do you know?



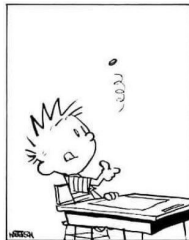
¹⁸Reynolds, “The Git Hater’s Guide”

The road goes on

Answer Correct?

- Sensor revolution is not slowing down
→ we need more and more data analysis
- ML/AI has all of these challenges . . . on steroids
- Attitudes: “move fast and break things”
- Systems are too complicated to prove correctness

How do you know?



Correctness: is there hope?



[redbubble.com]

¹⁹Baker, 1,500 scientists lift the lid on reproducibility Nature 533:452–454 (2016)

²⁰The Turing Way handbook to reproducible, ethical and collaborative data science

²¹Alston & Rick (2021) A Beginner's Guide to Conducting Reproducible Research

Correctness: is there hope?

... for industry

- big users of tooling, methodologies from OSS/Academia
- we have power



[redbubble.com]

¹⁹Baker, 1,500 scientists lift the lid on reproducibility Nature 533:452–454 (2016)

²⁰The Turing Way handbook to reproducible, ethical and collaborative data science

²¹Alston & Rick (2021) A Beginner's Guide to Conducting Reproducible Research

Correctness: is there hope?

... for industry

- big users of tooling, methodologies from OSS/Academia
- we have power

... for grant agencies, universities

- support research like the reproducibility project
- reward (promotions, grants) OS/open science contributions
- support **Research Software Engineering**
- funds to finish some better tools: (e.g. **gitless**)



¹⁹Baker, 1,500 scientists lift the lid on reproducibility Nature 533:452–454 (2016)

²⁰The Turing Way handbook to reproducible, ethical and collaborative data science

²¹Alston & Rick (2021) A Beginner's Guide to Conducting Reproducible Research

Correctness: is there hope?

... for industry

- big users of tooling, methodologies from OSS/Academia
- we have power

... for grant agencies, universities

- support research like the reproducibility project
- reward (promotions, grants) OS/open science contributions
- support **Research Software Engineering**
- funds to finish some better tools: (e.g. **gitless**)

... for researchers

- The perfect is the enemy of the good. 2nd best time to start is today.²⁰
- The most common reader of your code is you
- You want others to use your ideas (and cite your papers)²¹

¹⁹Baker, **1,500 scientists lift the lid on reproducibility** Nature 533:452–454 (2016)

²⁰**The Turing Way** handbook to reproducible, ethical and collaborative data science

²¹Alston & Rick (2021) **A Beginner's Guide to Conducting Reproducible Research**



Correct answers: who needs them?

a story of numerical computing



Rich and non-linear tomography?

Abstract: Numerical computing was motivated by rocket trajectories, and early algorithm work was marked by a focus on evaluating correctness. These techniques are now used to process nearly every image, sound and measurement from now ubiquitous sensors. Most work integrates numerous toolkits, many of which are open source. This wide use is accompanied by reduced focus on correctness. We discuss trends in validation of algorithms, along with the author's experience managing an open-source imaging software toolkit (EIDORS.org). Finally, we consider some higher-level issues: correctness vs "move fast and break things".