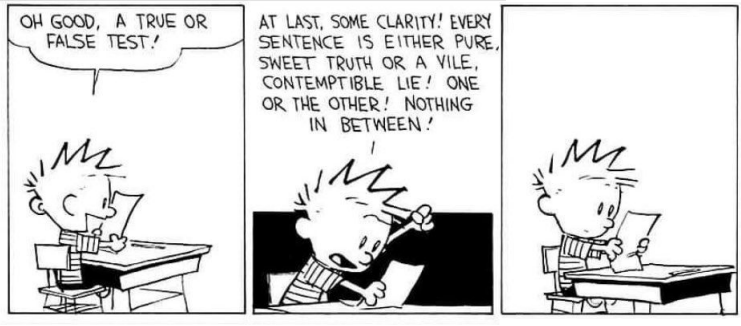
Correct answers: who needs them?
a story of numerical computing

RNT program, Isaac Newton Institute for Mathmatical Sciences, Cambridge, UK
25 May 2023
www.newton.ac.uk/seminar/39442

Andy Adler

Carleton University, Ottawa, Canada

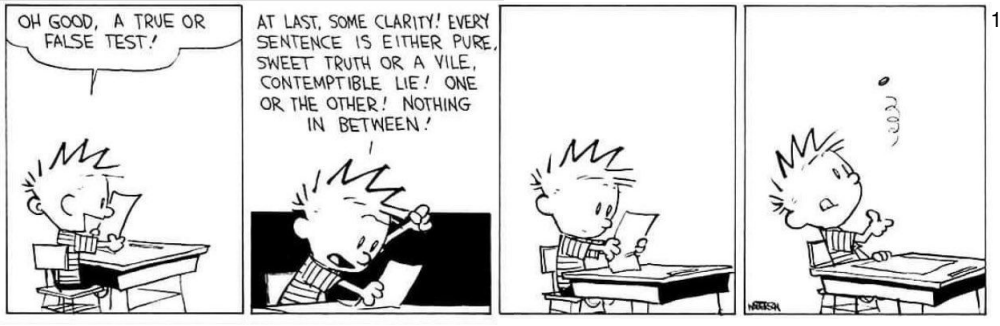
Correct?



1

[1] Watterson, “The Days Are Just Packed”, 1993

Correct?



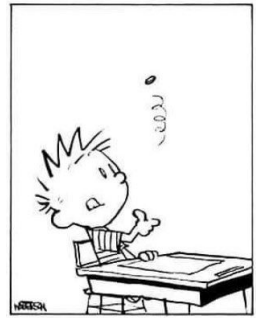
[1]

[1] Watterson, “The Days Are Just Packed”, 1993

Correct?

Answer Correct?

How do you know?



Correct?

Correct

$\triangleq$ 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

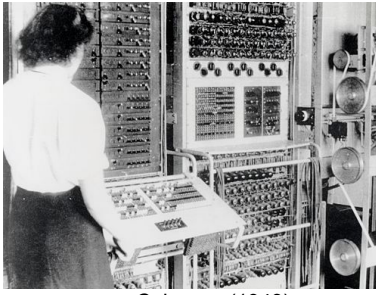
. . . distinct from a lucky guess, because of

Justification

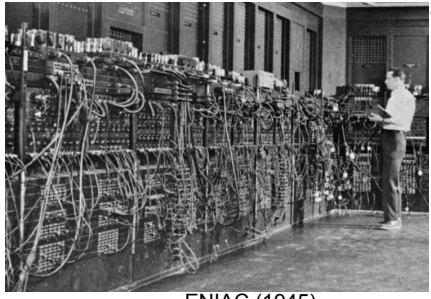
≜ reasons for a rationally admissible belief

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

Numerical computing



Colossus (1943)

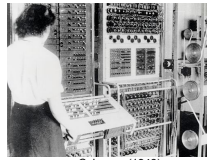


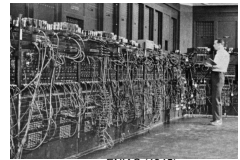
ENIAC (1945)[2]

- 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

[2] Colossus and ENIAC, Wikipedia.org

“Digital” vs Numerical computing

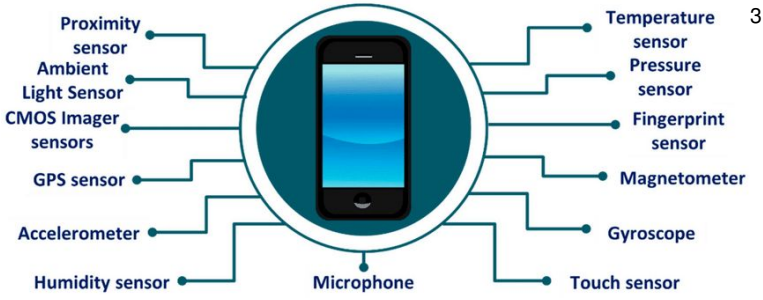

Colossus (1943)


ENIAC (1945)

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

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

Ubiquitous sensors



3

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

[3] Majumder & 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

Image stabilization

$\triangleq$ 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

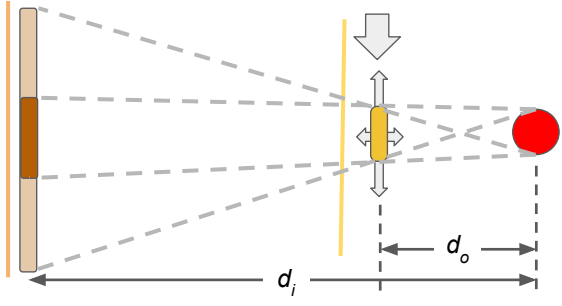
- Numerical algorithms are often used to solve inverse problems.
- These can fail when assumptions (*priors*) are not valid

[5] expertphotography.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

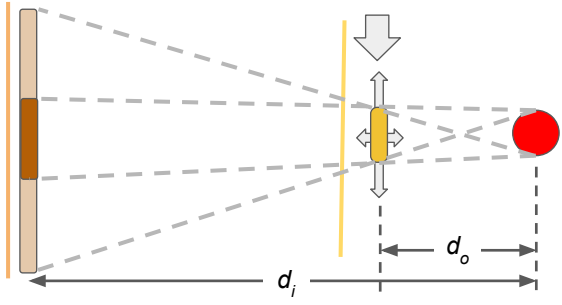
[6]Plato's cave, wikimedia.org, inspired by C Groetsch "Linear Inverse Problems" in O Schertzer Handbook of Mathematical Methods in Imaging

Inverse problems & Plato's cave



- Sensitivity
- Noise
- Null space

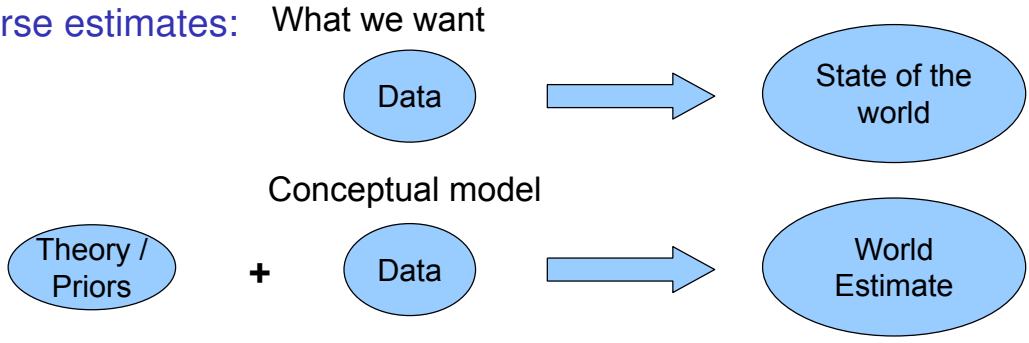
Inverse problems & Plato's cave



- Sensitivity
- Noise
- Null space

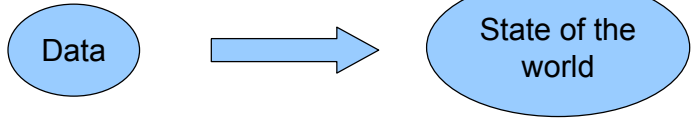
Adler, *Correct answers: who needs them?* . . .

Inverse estimates:

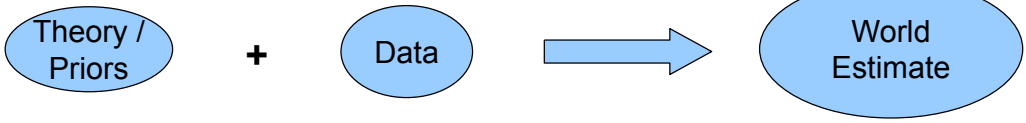


Inverse estimates:

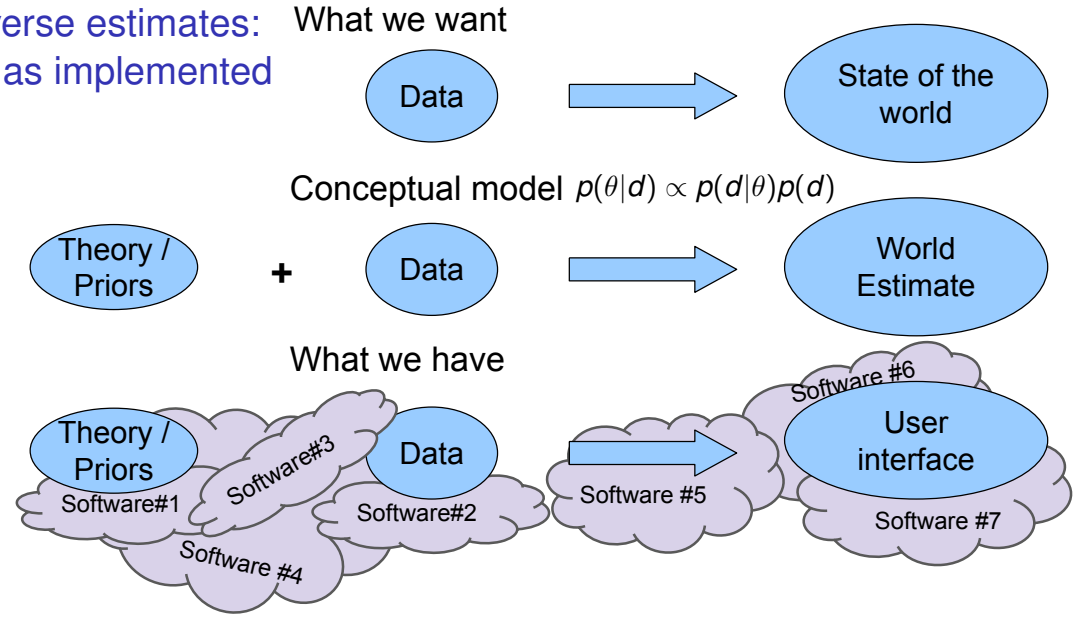
What we want



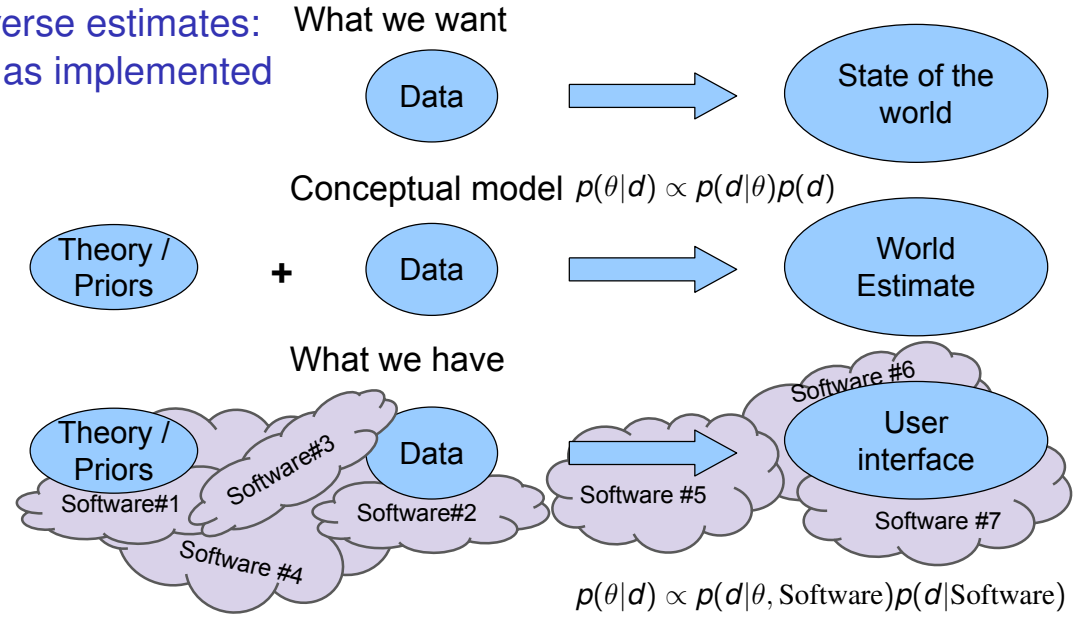
Conceptual model $p(\theta|d) \propto p(d|\theta)p(d)$



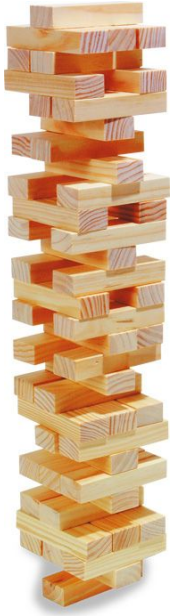
Inverse estimates: . . . as implemented



Inverse estimates: . . . as implemented



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 code [7]

- 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 SW[8]
- Part of open science . . . discussed later

[8] wikipedia.org/wiki/Open-source_software

[8] 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

[9]

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	dblat1.f	Level 1 BLAS Test Suite (prec. double)
…	…	…

Tests

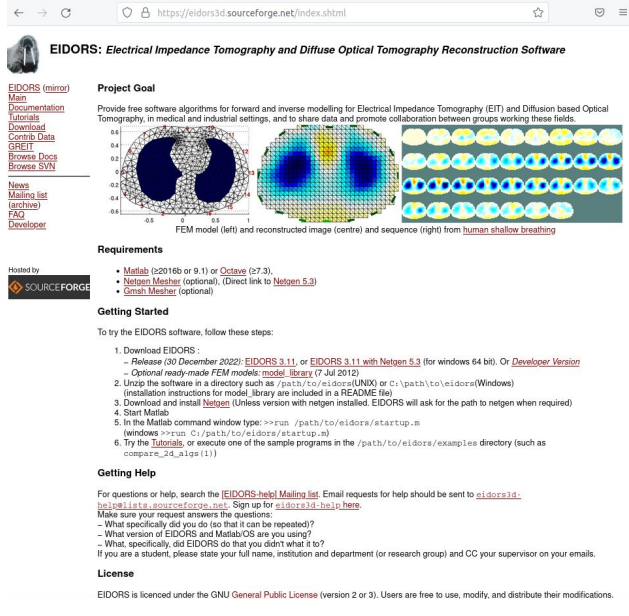
[9] wikipedia.org and netlib.org

EIDORS.org

Overview

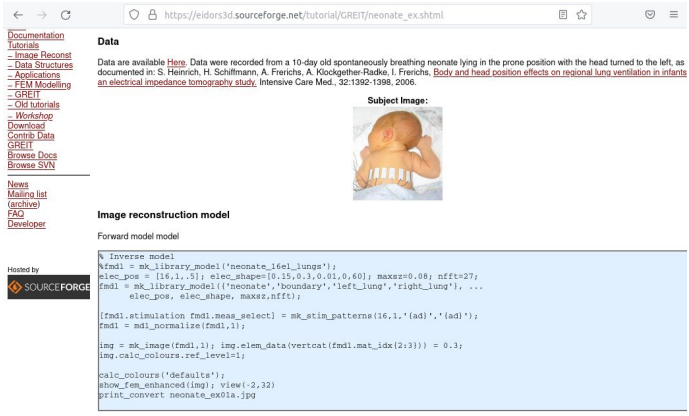
Uses other OSS

Build community
(Searchable answers)



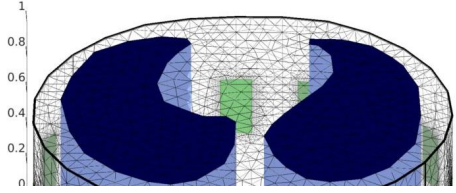
EIDORS.org

Tutorials!



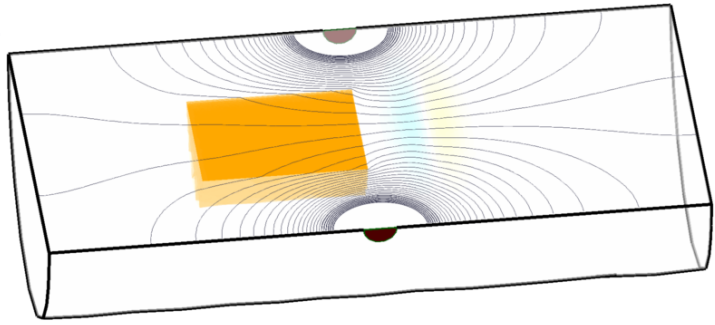
Users get started this way

(Often useful for authors)

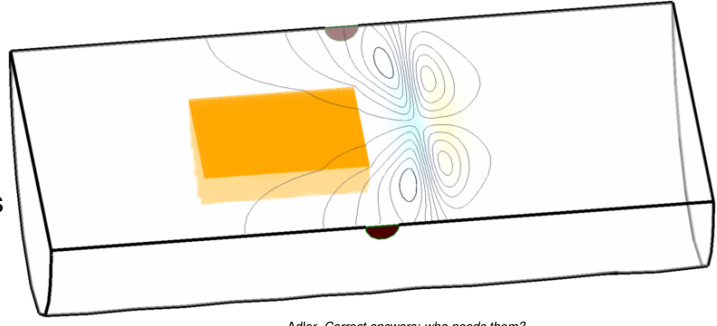


Idea: Combining Physics Acousto-Electric Tomography

Equipotenials

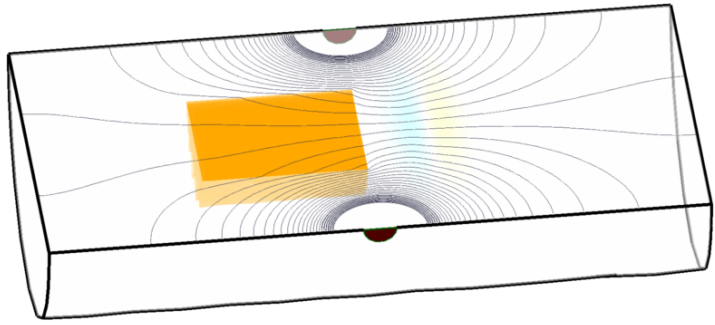


Δ Equipotenials

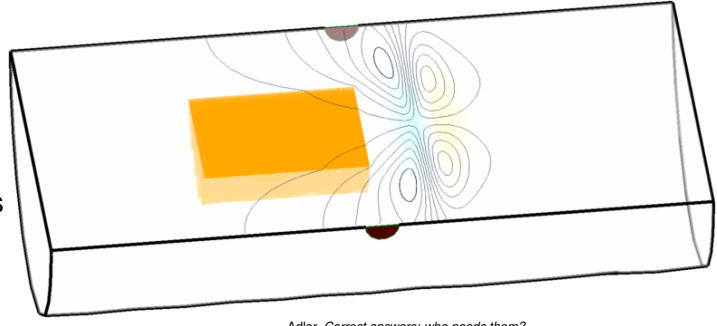


Idea: Combining Physics Acousto-Electric Tomography

Equipotenials

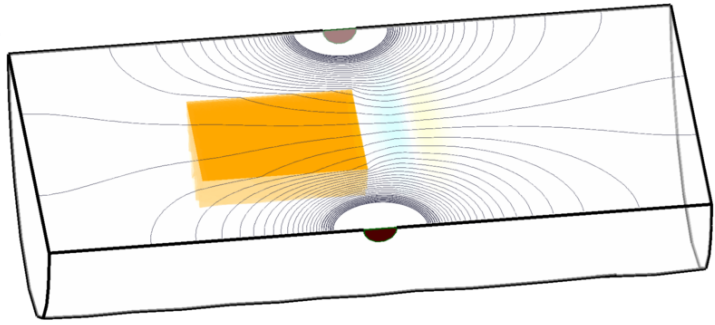


Δ Equipotenials

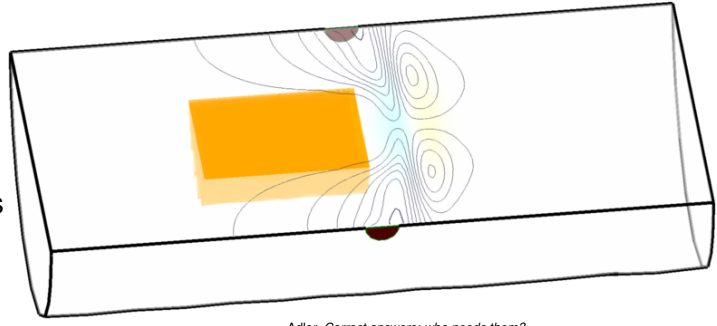


Idea: Combining Physics
Acousto-Electric Tomography

Equipotenials

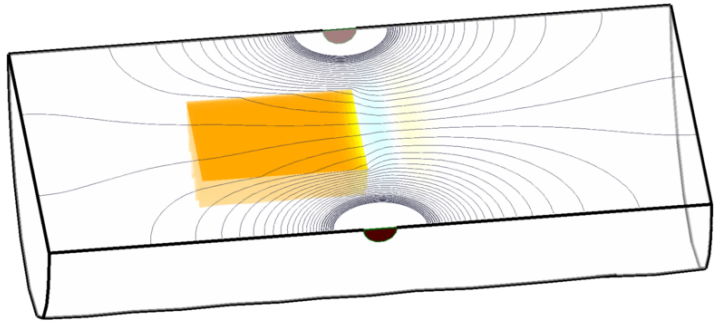


Δ Equipotenials

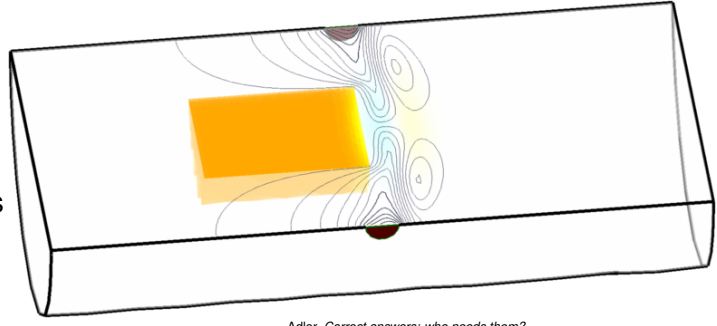


Idea: Combining Physics Acousto-Electric Tomography

Equipotenials

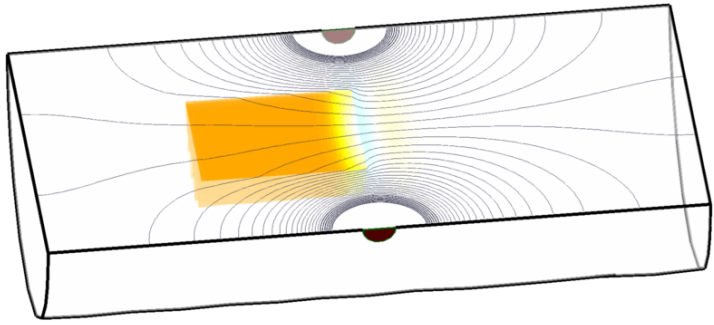


Δ Equipotenials

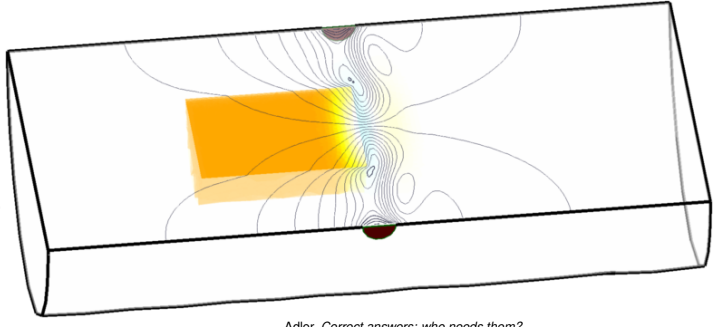


Idea: Combining Physics
Acousto-Electric Tomography

Equipotenials

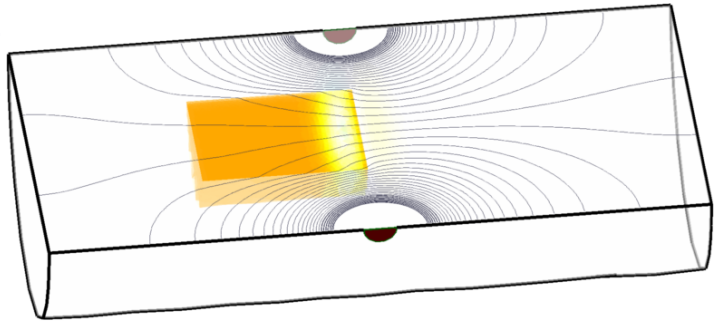


Δ Equipotenials

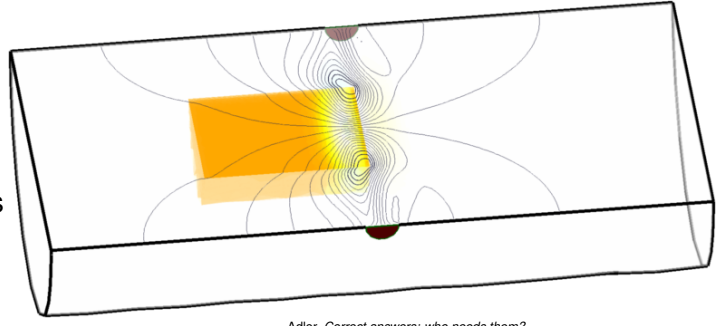


Idea: Combining Physics Acousto-Electric Tomography

Equipotenials

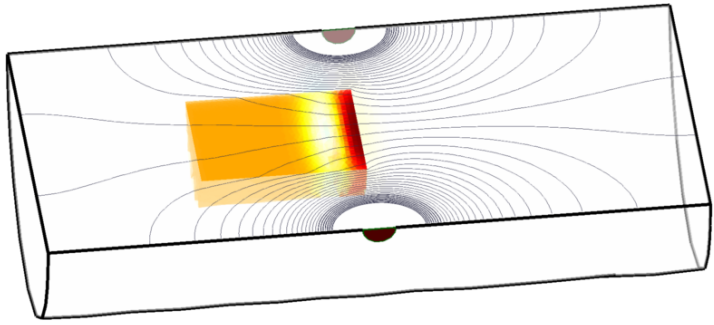


Δ Equipotenials

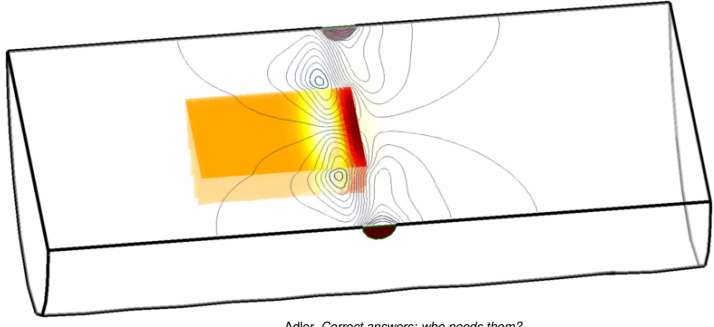


Idea: Combining Physics Acousto-Electric Tomography

Equipotenials

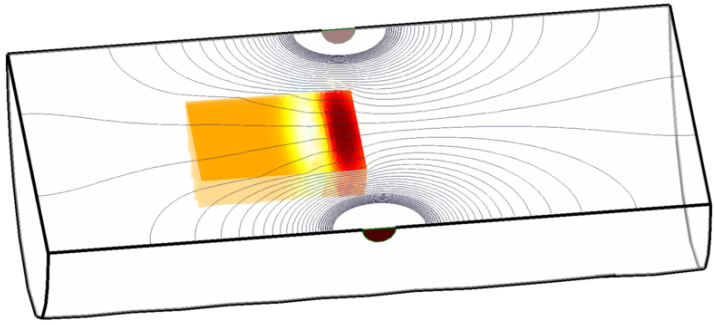


Δ Equipotenials

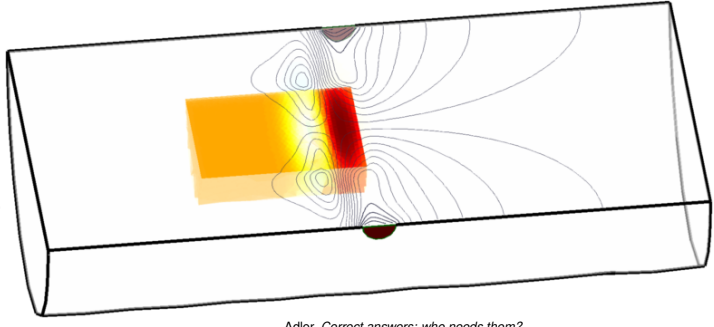


Idea: Combining Physics Acousto-Electric Tomography

Equipotenials

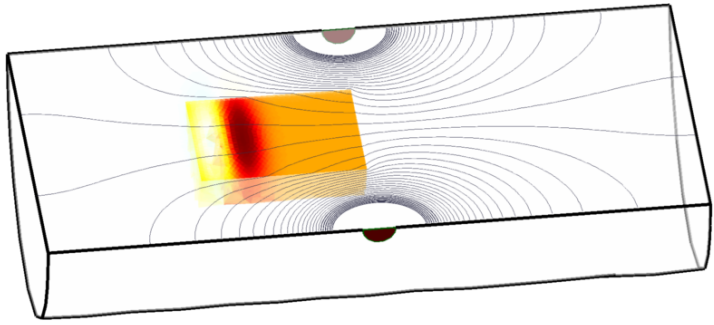


Δ Equipotenials

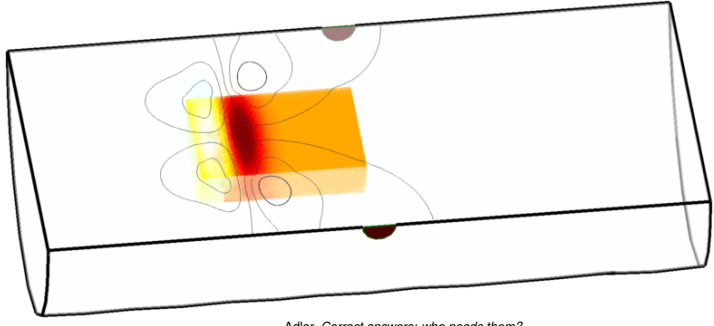


Idea: Combining Physics Acousto-Electric Tomography

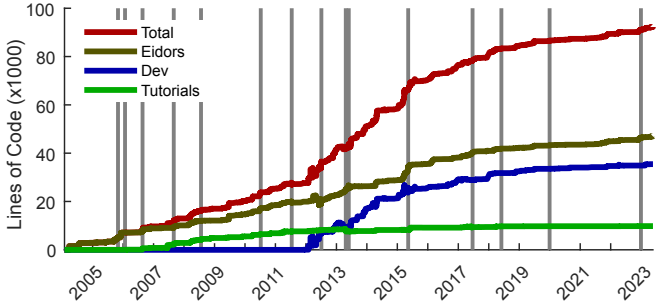
Equipotenials



Δ Equipotenials



EIDORS: experiences



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

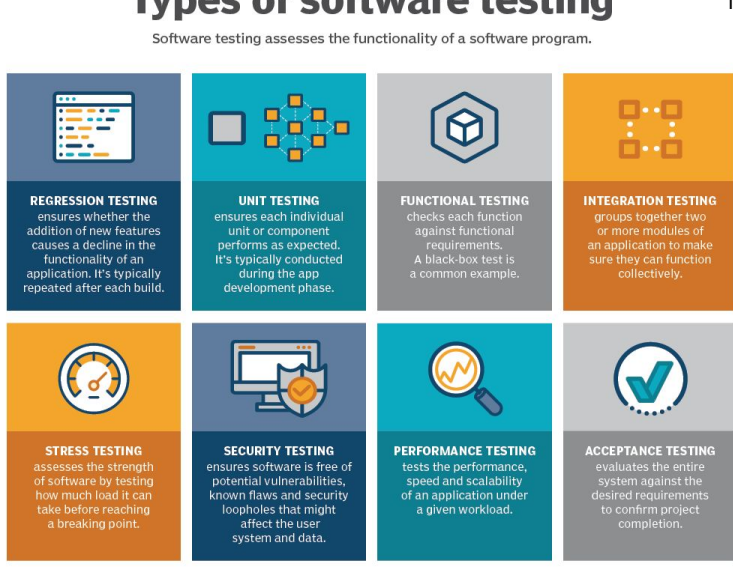
EIDORS: contributions and testing

- 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

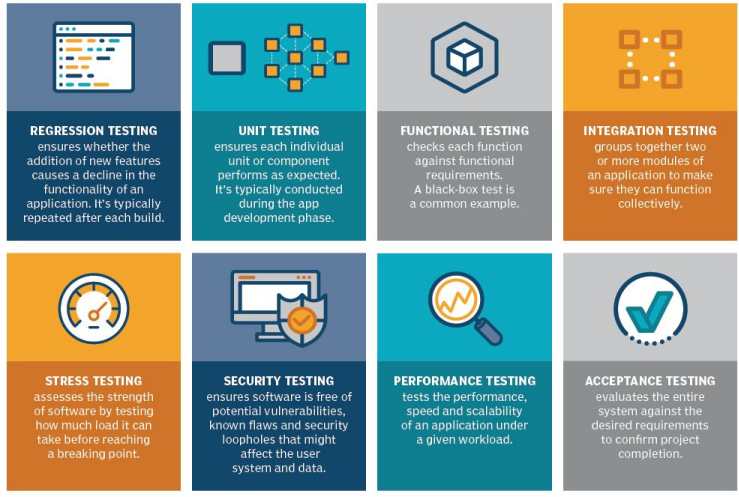


[10] www.techtarget.com/whatis/definition/software-testing

Software testing: a mature field in software engineering

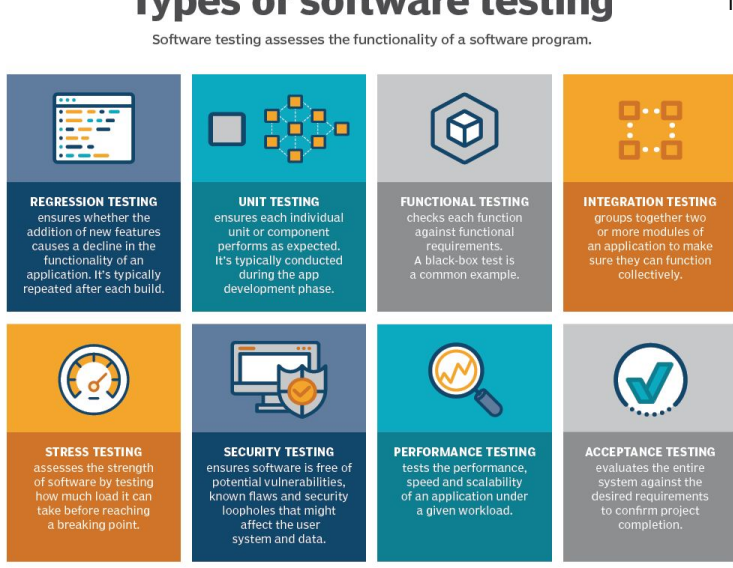
EIDORS uses:

- Regression Tests:
  Same as last time?



[10] www.techtarget.com/whatis/definition/software-testing

Software testing: a mature field in software engineering



EIDORS uses:

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

[10] www.techtarget.com/whatis/definition/software-testing

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

- $\triangleq$ a crowdsourced collaboration . . . to repeat 100 published experimental and correlational psychological studies . . . led by the Center for Open Science[11]

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 implment 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 agree[12]

- *p*-values for more than 250,000 psychology papers checked; around half contained at least one incorrect *p*-value![13]

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

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

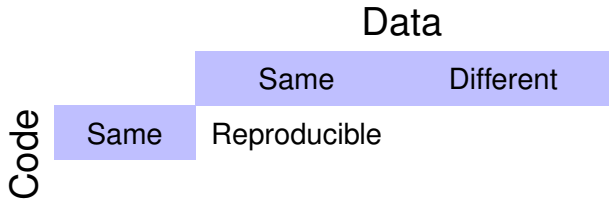
How about reproducing complicated results?

- 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

[14] Hardwicke *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

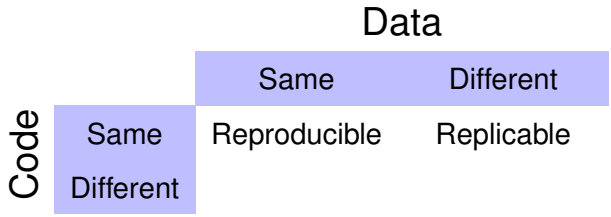
[15] Hardwicke *et al* "Analytic reproducibility in articles receiving open data badges at the journal Psychological Science: an observational study" (2021) Royal Soc. open sci 8201494201494

Open science: Reproducible science[16]



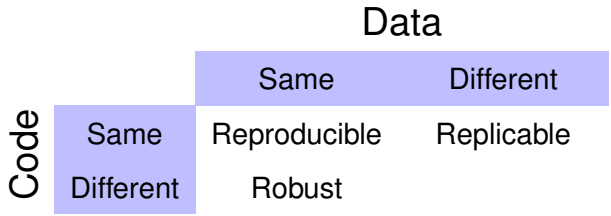
[16] Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science[16]



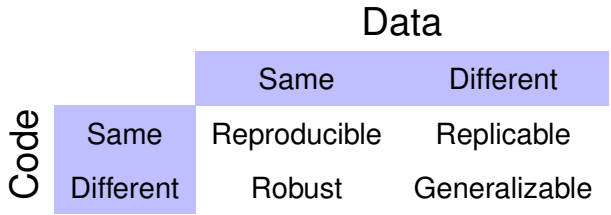
[16] Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science[16]



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

Open science: Reproducible science[16]



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

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

Summary so far

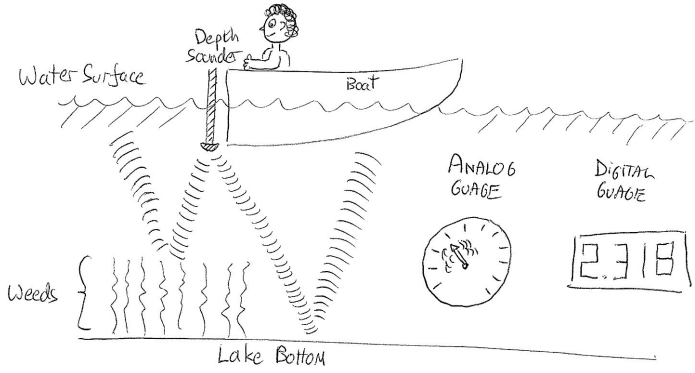
Numerical Computing

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

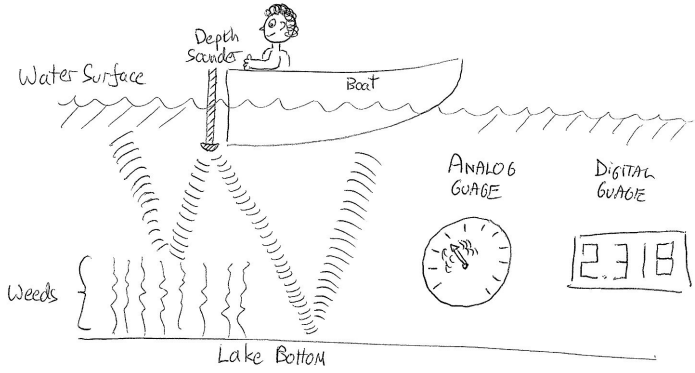
What does this mean for?

- Data
- Users
- Science

Data Quality



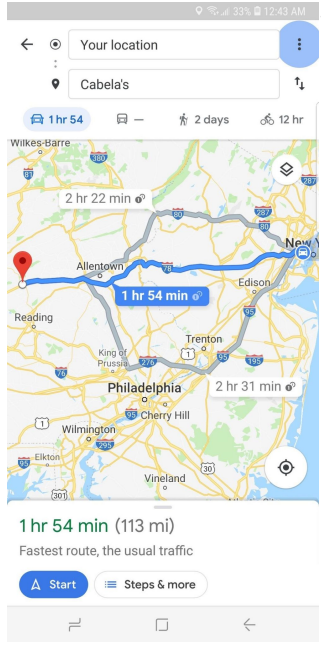
Data Quality



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

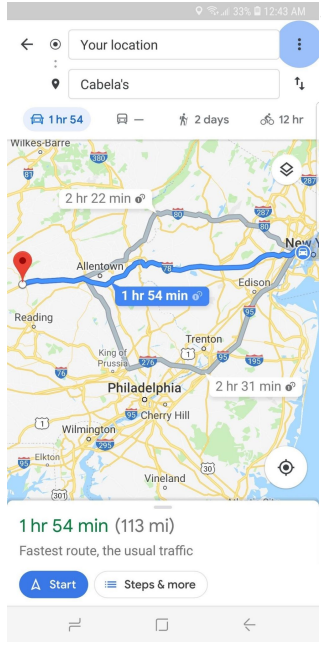
Thoughts on Numerical Algorithms: . . . working with a user



[17] smartphones.gadgethacks.com

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

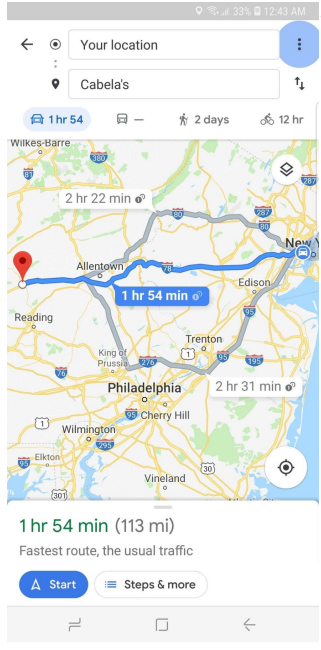
- Good that they say: “the usual traffic”
  - But is this estimate for me
  - or is it an average for all drivers?



[17] smartphones.gadgethacks.com

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



[17] smartphones.gadgethacks.com

Thoughts on Open Science

Reproducibility is the
sine qua non of science

Why is it so hard?

Answer Correct?

How do you know?

[18] Reynolds, "The Git Hater's Guide"

Thoughts on Open Science

Answer Correct?

How do you know?

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

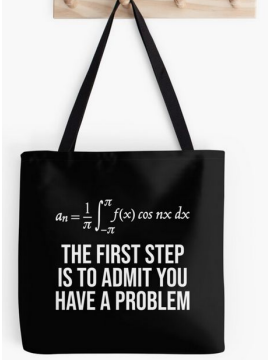
[18] Reynolds, "The Git Hater's Guide"

The road goes on

Answer Correct?

- Sensor revolution is not slowing down
  $\rightarrow$ we need more and more data analysis *How do you know?*
- ML/AI has all of these challenges . . . on steroids
- Attitudes: “move fast and break things”
- Systems are too complicated to prove correctness

Correctness: is there hope?



[redbubble.com]

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

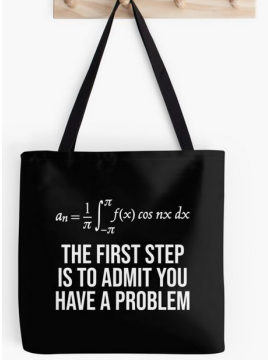
[20] The Turing Way handbook to reproducible, ethical and collaborative data science

[21] Alston & 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]

[19] Baker, 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

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

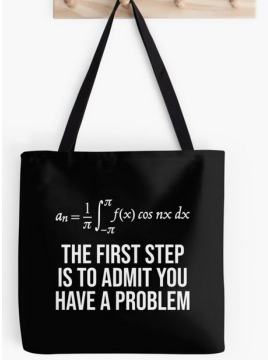
Correctness: is there hope?

. . . 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)



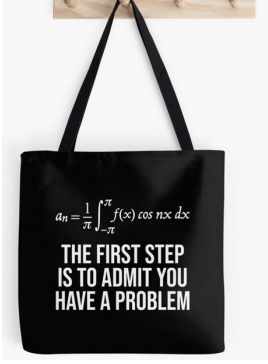
[redbubble.com]

[19] Baker, 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

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

Correctness: is there hope?



[redbubble.com]

. . . 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

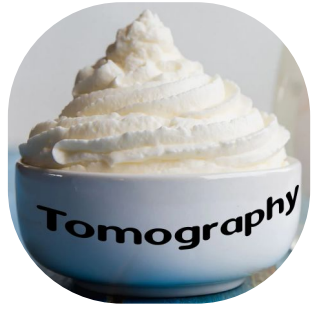
- The perfect is the enemy of the good. 2nd best time to start is today.[20]
- The most common reader of your code is you
- You want others to use your ideas (and cite your papers)[21]

[19] Baker, 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

[21] Alston & 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".