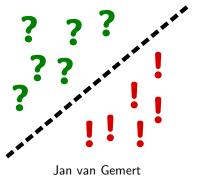
Metascience for ML

Preaching to the Choir

June 20, 2025



Whoami: Jan van Gemert



Assoc. prof; head Computer Vision lab @ PRB

Two main research themes:

- Fundamental empirical understanding-based deep learning research; (to)
- Find & evaluate powerful yet flexible physical priors for data-efficient visual recognition AI.

Whoami: Jan van Gemert



Assoc. prof; head Computer Vision lab @ PRB

Two main research themes:

- Fundamental empirical understanding-based deep learning research; (to)
- Find & evaluate powerful yet flexible physical priors for data-efficient visual recognition AI.

I signed up because...

I want to share my vision and learn from others how they do ML research.

Metascience for Machine Learning is...

A method for doing research.

I would like to contribute to metascience for ML by...

My own incomplete work-in-progress methodology :).

Whoami: Jan van Gemert



Assoc. prof; head Computer Vision lab @ PRB

Two main research themes:

- Fundamental empirical understanding-based deep learning research; (to)
- Find & evaluate powerful yet flexible physical priors for data-efficient visual recognition AI.

I signed up because...

I want to share my vision and learn from others how they do ML research.

Metascience for Machine Learning is...

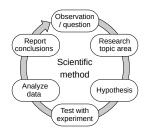
A method for doing research.

I would like to contribute to metascience for ML by...

My own incomplete work-in-progress methodology :).

Mine is not "The Way"; it's "A Way".

Deep learning powers AI; yet as a scientific field has growing pains [2,3,4]



- Improvement-driven (large compute/data);
- Trial and error (graduate student descent)
- Opportunistic (career driven);
- Reviewer damage (Benchmark fetish; Mathiness);
- Confusing speculation with explanation
- Not identifying the reasons for empirical gains.

3/7

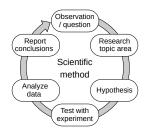
^{[1]:} https://en.wikipedia.org/wiki/Scientific_method

^{[2]:} Lipton et al. "Troubling Trends in Machine Learning Scholarship", 2018.

^{[3]:} Sculley, David, et al. "Winner's curse? On pace, progress, and empirical rigor." 2018.

^{[4]:} Togelius, Julian, et al. "Choose your weapon: Survival strategies for depressed AI academics" Proc. of the IEEE (2024).

Deep learning powers AI; yet as a scientific field has growing pains [2,3,4]



- Improvement-driven (large compute/data);
- Trial and error (graduate student descent)
- Opportunistic (career driven);
- Reviewer damage (Benchmark fetish; Mathiness);
- Confusing speculation with explanation
- Not identifying the reasons for empirical gains.

3/7

ML/DL does not have many empirical theories.

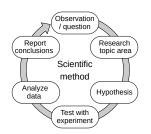
^{[1]:} https://en.wikipedia.org/wiki/Scientific_method

^{[2]:} Lipton et al. "Troubling Trends in Machine Learning Scholarship", 2018.

^{[3]:} Sculley, David, et al. "Winner's curse? On pace, progress, and empirical rigor." 2018.

^{[4]:} Togelius, Julian, et al. "Choose your weapon: Survival strategies for depressed AI academics" Proc. of the IEEE (2024).

Deep learning powers AI; yet as a scientific field has growing pains [2,3,4]



- Improvement-driven (large compute/data);
- Trial and error (graduate student descent)
- Opportunistic (career driven);
- Reviewer damage (Benchmark fetish; Mathiness);
- Confusing speculation with explanation
- Not identifying the reasons for empirical gains.

3/7

- \bullet ML/DL does not have many empirical theories. Some that I am aware of:
 - \circ Neural Scaling Laws; \circ Bias/variance \circ ML is like physics/neuroscience;
 - \circ Simple axioms explaining intelligence $\circ \dots$

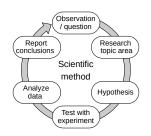
^{[1]:} https://en.wikipedia.org/wiki/Scientific_method

^{[2]:} Lipton et al. "Troubling Trends in Machine Learning Scholarship", 2018.

^{[3]:} Sculley, David, et al. "Winner's curse? On pace, progress, and empirical rigor." 2018.

^{[4]:} Togelius, Julian, et al. "Choose your weapon: Survival strategies for depressed AI academics" Proc. of the IEEE (2024).

Deep learning powers AI; yet as a scientific field has growing pains [2,3,4]



- Improvement-driven (large compute/data);
- Trial and error (graduate student descent)
- Opportunistic (career driven);
- Reviewer damage (Benchmark fetish; Mathiness);
- Confusing speculation with explanation
- Not identifying the reasons for empirical gains.

3/7

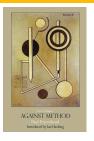
- ML/DL does not have many empirical theories. Some that I am aware of:
 Neural Scaling Laws;
 Bias/variance
 ML is like physics/neuroscience;
 - ∘ Simple axioms explaining intelligence ∘ ...
- Mores in the field: End-to-end learning; 'bold' numbers on common datasets; trail and error; openly sharing code/weights/data; all papers open on ArXiv.

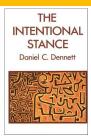
^{[1]:} https://en.wikipedia.org/wiki/Scientific_method

^{[2]:} Lipton et al. "Troubling Trends in Machine Learning Scholarship", 2018.

^{[3]:} Sculley, David, et al. "Winner's curse? On pace, progress, and empirical rigor." 2018.

^{[4]:} Togelius, Julian, et al. "Choose your weapon: Survival strategies for depressed Al academics" Proc. of the IEEE (2024).

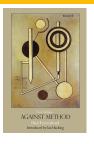


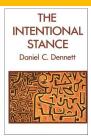


 There is not "one way" to do science. Science moves on, despite the methodology used^[5].

^{[5]:} Paul Feyerabend; "Against method", 1975

^{[6]:} Daniel Dennet; "Intentional Stance", 1989

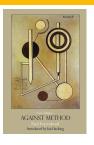


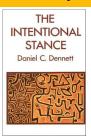


- There is not "one way" to do science. Science moves on, despite the methodology used^[5].
- Squabbling over definitions gets us nowhere (we can't even define a 'chair').
 If an LLM has 'intent', if it 'understands', if it is 'intelligent'. The meaning of these words is clear and useful for AI systems^[6].

^{[5]:} Paul Feyerabend; "Against method", 1975

^{[6]:} Daniel Dennet; "Intentional Stance", 1989

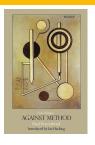


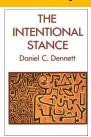


- There is not "one way" to do science. Science moves on, despite the methodology used^[5].
- Squabbling over definitions gets us nowhere (we can't even define a 'chair').
 If an LLM has 'intent', if it 'understands', if it is 'intelligent'. The meaning of these words is clear and useful for AI systems^[6].
- I'm in ML/DL science because I like the intellectual pursuit, not 'solve AI'. So, why should I tell 'system builders' they need to stop what they like?

^{[5]:} Paul Feyerabend; "Against method", 1975

^{[6]:} Daniel Dennet; "Intentional Stance", 1989





- There is not "one way" to do science. Science moves on, despite the methodology used^[5].
- Squabbling over definitions gets us nowhere (we can't even define a 'chair').
 If an LLM has 'intent', if it 'understands', if it is 'intelligent'. The meaning of these words is clear and useful for AI systems^[6].
- I'm in ML/DL science because I like the intellectual pursuit, not 'solve AI'.
 So, why should I tell 'system builders' they need to stop what they like?

Let people do research however they want (including yourself).

^{[5]:} Paul Feyerabend; "Against method", 1975

^{[6]:} Daniel Dennet: "Intentional Stance", 1989







scienceintegritydigest.com

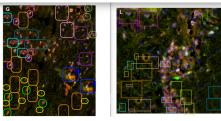
1: ML misconduct: tune on the testset; cherry picking; plagiarism, overclaiming; isn't as bad as the explict manipulation as done here.

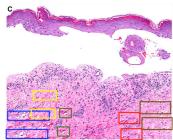
5/7



G O S udenceintegritydgest.com/sbox/ B t/r Œlisabethiik → 1	D & • U F + 0 - 2
Science Integrity Digest A blog about actoric integrity, by Elisabeth Bill, for Harbers-Bill LLC. Support my work at Poteron corea/elisabeth bill.	
Home About FAQ How-To-guides	
About	RECENT POSTS
Elisabeth life is a renowned raicrobiologist and science integrity advocate known for detecting	Sciencetwardians, where disgrantled authors complain about Publiver April 10, 2025
Elliabeth Illic is a recovered microbiologist and science integrity advocate known for detecting trage deplication in scientific publications. Through metiodous enalysis, she has exposed thousants of cases of micrordax, foreign temperary and accountability, Deophie facing	Sciencetsanilans, where disgrantied authors complain about Publiser
About Black Black is reversed existentisingles and observ lamply who was beaun for deserving strange displaced in its vicinity policies. The treaty terrelation models, the law reposed increased for control of the confidence of control office produced for control office of the control of control of the control of control of the	Science standards, where disgranded authors complains about Publiser April 10, 2025 Science Integrity Dignst – catching up

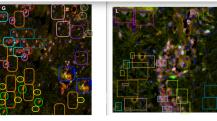
scienceintegritydigest.com

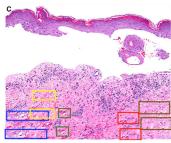












scienceintegritydigest.com

<u>Doesn't (only) preach "Do</u>n't do fraud; it's bad" 1; she does the work.

My work for fundamental empirical research in ML/DL



reproducedpapers.org



Online research guidelines



controlledexperimentsinml.org



MSc course

The last slide: end on a high note.

I don't believe:

- No single way to do science;
- No "definition squabbling" (we can't even define a chair).
- No preaching; let system builders build systems.

The last slide: end on a high note.

I don't believe:

- No single way to do science;
- No "definition squabbling" (we can't even define a chair).
- No preaching; let system builders build systems.

I believe:

- ML/DL work is open as a field, openly sharing code, weights, papers.
- ML/DL misconduct (tune on the testset; cherry picking; plagiarism, overclaiming) is not as bad as elsewhere; limited direct fraud
- that the scientific method will correct things eventually.
- in "Be and let Be". Let others do research their own way.
- in doing: help the ones that want to be helped.
- in moving constructively forward: Methodological development: reproducedpapers.org; controlledexperimentsinml.org; research guidelines; MSc course, this workshop, etc... (?)