

Good enough

A risk-based canvas for validating AI in clinical trials

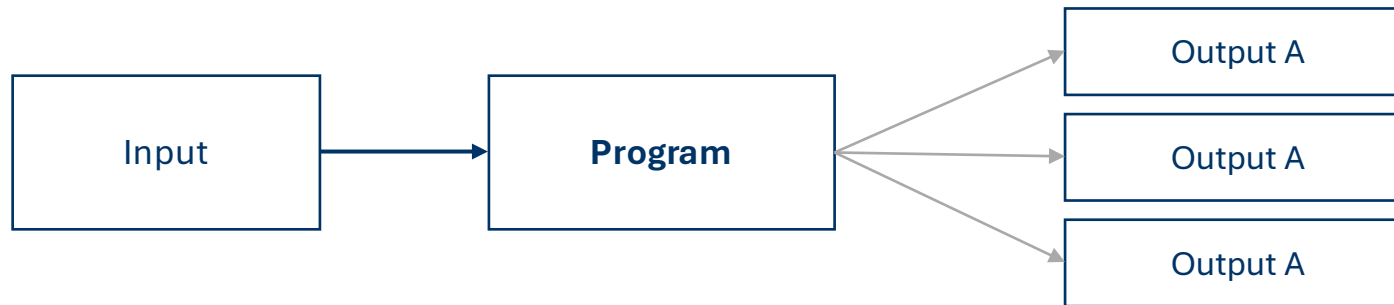
Nechama Katan · Wicked Problem Wizards

BRAVE Working Group · PHUSE Cambridge, October 8, 2026



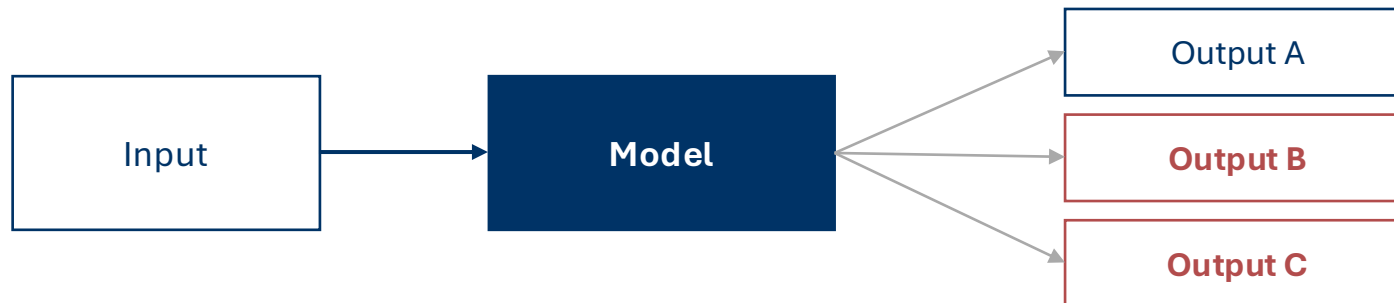
Same input, same output? Not anymore

Traditional software



Test once, trust every run

Generative model



A test that passes once does not prove the next run



BRAVE: find it, reduce it, justify it



Find it

Adaptive Risk Canvas

Q4 2026 • today

Reduce it

Controls that lower the score

March 2027

Justify it

Computer system monitoring

October 2027

Bridging Regulations and AI Validation Evidence • independent, cross-industry, pre-competitive • formerly the Adaptive Assurance Canvas

BRAVE working group licensed under Creative Commons (CC BY 4.0), permitting reuse and adaptation with attribution



The validation process in three parts

Risk Canvas

Find it

How clearly can you identify your current risk?

Adaptive Controls

Reduce it

How do I reduce risk before I have to validate?

Justified Position

Justify it

How do I know when I have done enough?



The working group

June 2026

started

Weekly

meetings

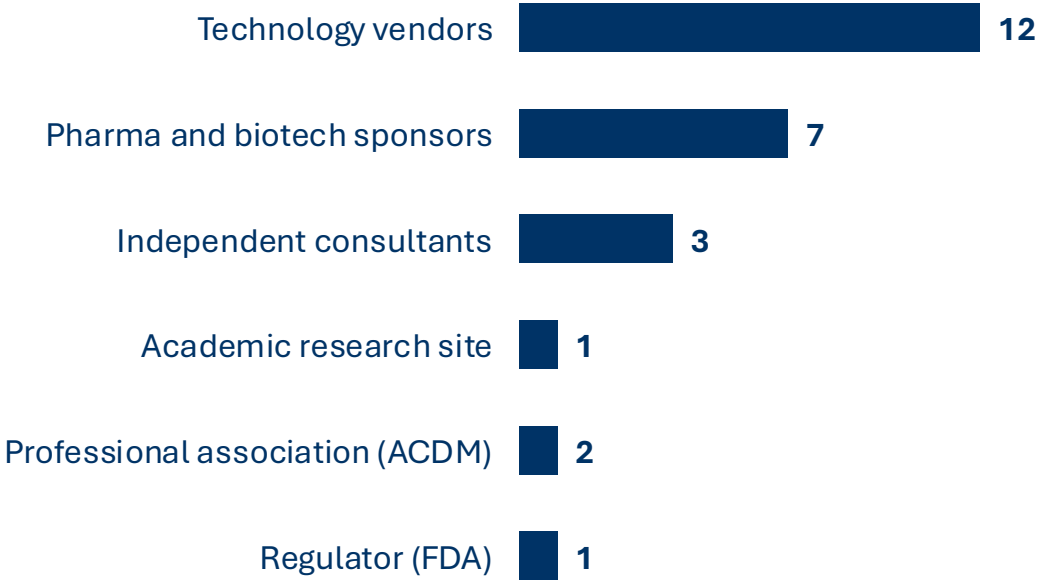
3

deliverables: find, reduce, justify

FDA

a staff member attends

Organizations represented (33 members)

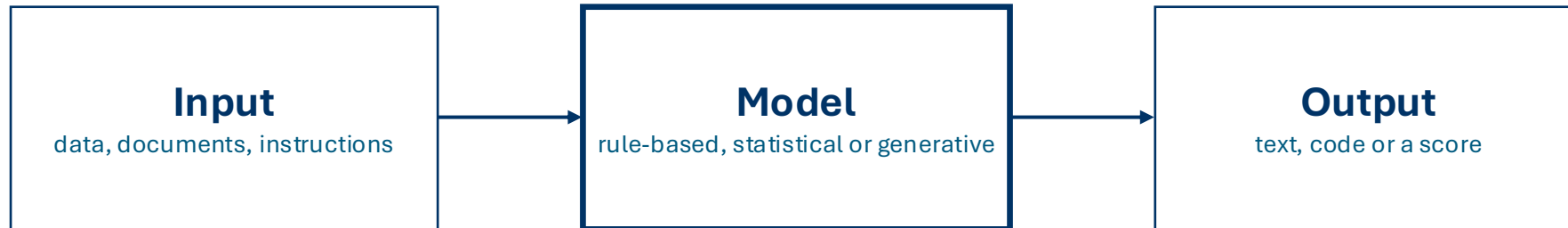


Members in North America, Europe and India



Most of "the AI" is ordinary software

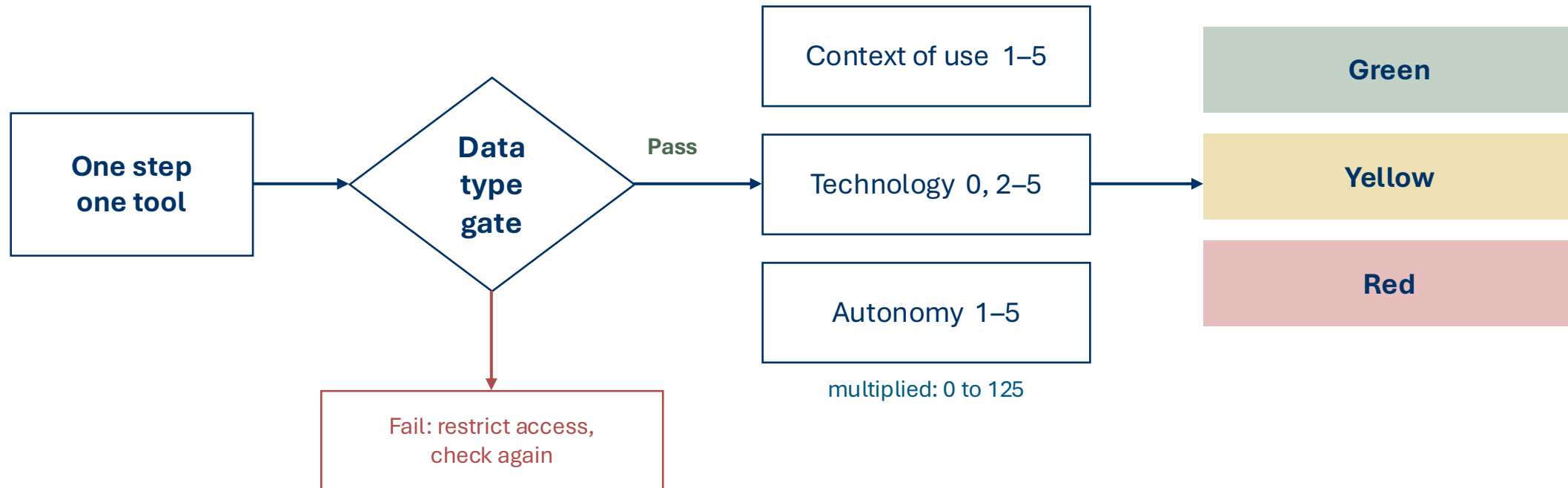
Harness: ordinary software, validated like traditional software



The harness decides what goes in, what the output becomes, and what gets logged



How the canvas scores one step



First: the data type gate

	No outside path, or controlled	Uncontrolled outside path
Safety and care data	Pass	Fail
Identifiable participant data	Pass	Fail
Unblinded trial data	Pass	Fail
Coded or de-identified data	Pass	Pass
Non-personal data	Pass	Pass

Highest harm at the top



Three axes, five levels each

Context of use

impact if the output is incorrect

5	Patient safety
4	Data integrity
3	Data cleaning
2	Operational processes
1	Personal productivity

Technology

how easily the output can be checked

5	Agentic
4	Free-text generative
3	Constrained generative
2	Statistical, machine learning
0	Explicit / static

Autonomy

who reviews, and when

5	Acts without review
4	Acts, then a person reviews
3	Person approves final output only
2	Person approves each output
1	Read-only

Context of use follows the FDA's "decision consequence" (AI draft guidance, January 2025)



One grid per context of use

Patient safety					
Acts without review	0	50	75	100	125
Acts, then a person reviews	0	40	60	80	100
Person approves the final output only	0	30	45	60	75
Person approves each output	0	20	30	40	50
Read-only	0	10	15	20	25
	Explicit / static	Statistical, machine learning	Constrained generative	Free-text generative	Agentic

Data integrity					
Acts without review	0	40	60	80	100
Acts, then a person reviews	0	32	48	64	80
Person approves the final output only	0	24	36	48	60
Person approves each output	0	16	24	32	40
Read-only	0	8	12	16	20
	Explicit / static	Statistical, machine learning	Constrained generative	Free-text generative	Agentic

Data cleaning					
Acts without review	0	30	45	60	75
Acts, then a person reviews	0	24	36	48	60
Person approves the final output only	0	18	27	36	45
Person approves each output	0	12	18	24	30
Read-only	0	6	9	12	15
	Explicit / static	Statistical, machine learning	Constrained generative	Free-text generative	Agentic

Operational processes					
Acts without review	0	20	30	40	50
Acts, then a person reviews	0	16	24	32	40
Person approves the final output only	0	12	18	24	30
Person approves each output	0	8	12	16	20
Read-only	0	4	6	8	10
	Explicit / static	Statistical, machine learning	Constrained generative	Free-text generative	Agentic

Personal productivity					
Acts without review	0	10	15	20	25
Acts, then a person reviews	0	8	12	16	20
Person approves the final output only	0	6	9	12	15
Person approves each output	0	4	6	8	10
Read-only	0	2	3	4	5
	Explicit / static	Statistical, machine learning	Constrained generative	Free-text generative	Agentic

Green A validation approach for this use already exists.

Yellow The use can be validated once controls are added, or once the step is redesigned to lower its score.

Red Validating the use is high risk, and the use likely cannot be validated as designed.

BRAVE working group licensed under Creative Commons (CC BY 4.0), permitting reuse and adaptation with attribution



Worked example: AI medical coding

Context of use

impact if the output is incorrect

5	Patient safety
4	Data integrity
3	Data cleaning
2	Operational processes
1	Personal productivity

Technology

how easily the output can be checked

5	Agentic
4	Free-text generative
3	Constrained generative
2	Statistical, machine learning
0	Explicit / static

Autonomy

who reviews, and when

5	Acts without review
4	Acts, then a person reviews
3	Person approves final output only
2	Person approves each output
1	Read-only

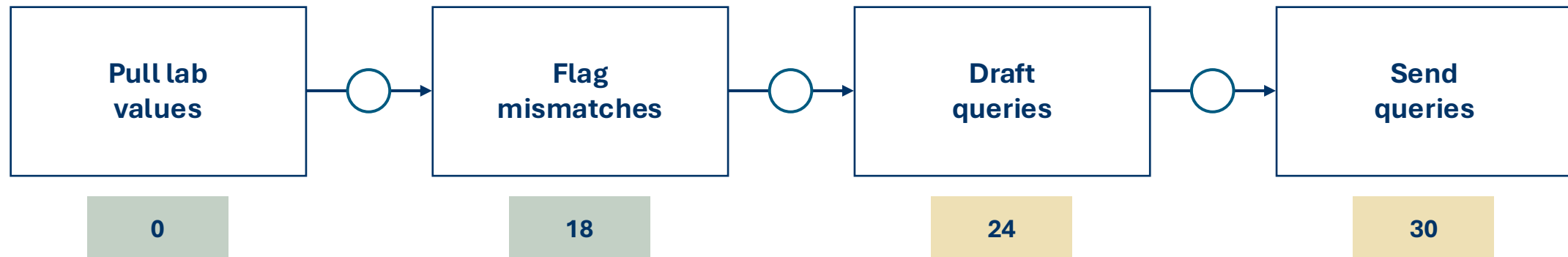
Gate: coded data, no outside path → pass

$$4 \times 3 \times 2 = 24$$

Yellow



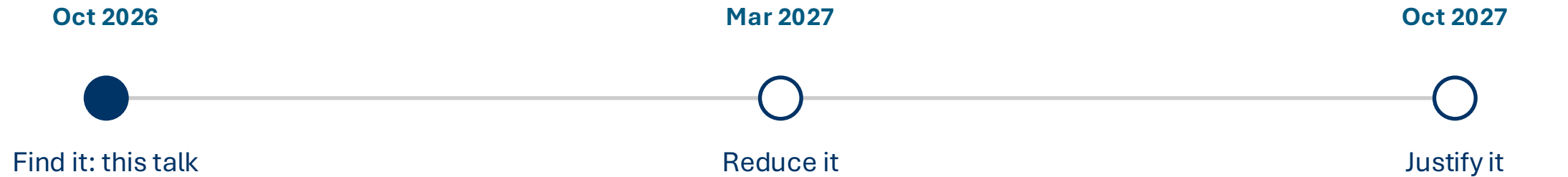
Score each step, check each hand-off



○ hand-off: is the audit trail kept by the system, or by a person?



How to get involved – This Slide needs help



Join the working group

Thursdays, 10 AM Eastern

Send us a use case

We use them to calibrate the canvas

Try the canvas

[link to tool or file]

[contact]



CSV

Built for tools that do what they're told.

CSA

Adds risk-proportionality, but assumes you can characterize the risk upfront.

CSM

Computer System Monitoring: a proposed working term, not yet a standard.

CSM: validation that never stops.

What it is

Ongoing performance evaluation of a deployed AI system to confirm it keeps meeting its intended use, beyond the point-in-time check CSV provides.

How it works

- Automated test harnesses and CI/CD, with regression testing
- Runs continuously, from development into production
- Comparing to human review

Why it matters

CSV and CSA confirm a system behaves at one point in time. CSM confirms it keeps doing so.

The role it fills

A formal complement to existing frameworks, essential for any AI that acts autonomously in a GxP environment.



Thank you

- Nechama Katan
- [linkedin.com/in/Nechama](https://www.linkedin.com/in/Nechama)



- [**www.wickedproblemwizards.com**](http://www.wickedproblemwizards.com)

