PNSQC

OCTOBER 10-12 2022

Jack McDowell
Ying Ki Kwong

Artificial Intelligence is the New Astrology of Software Quality

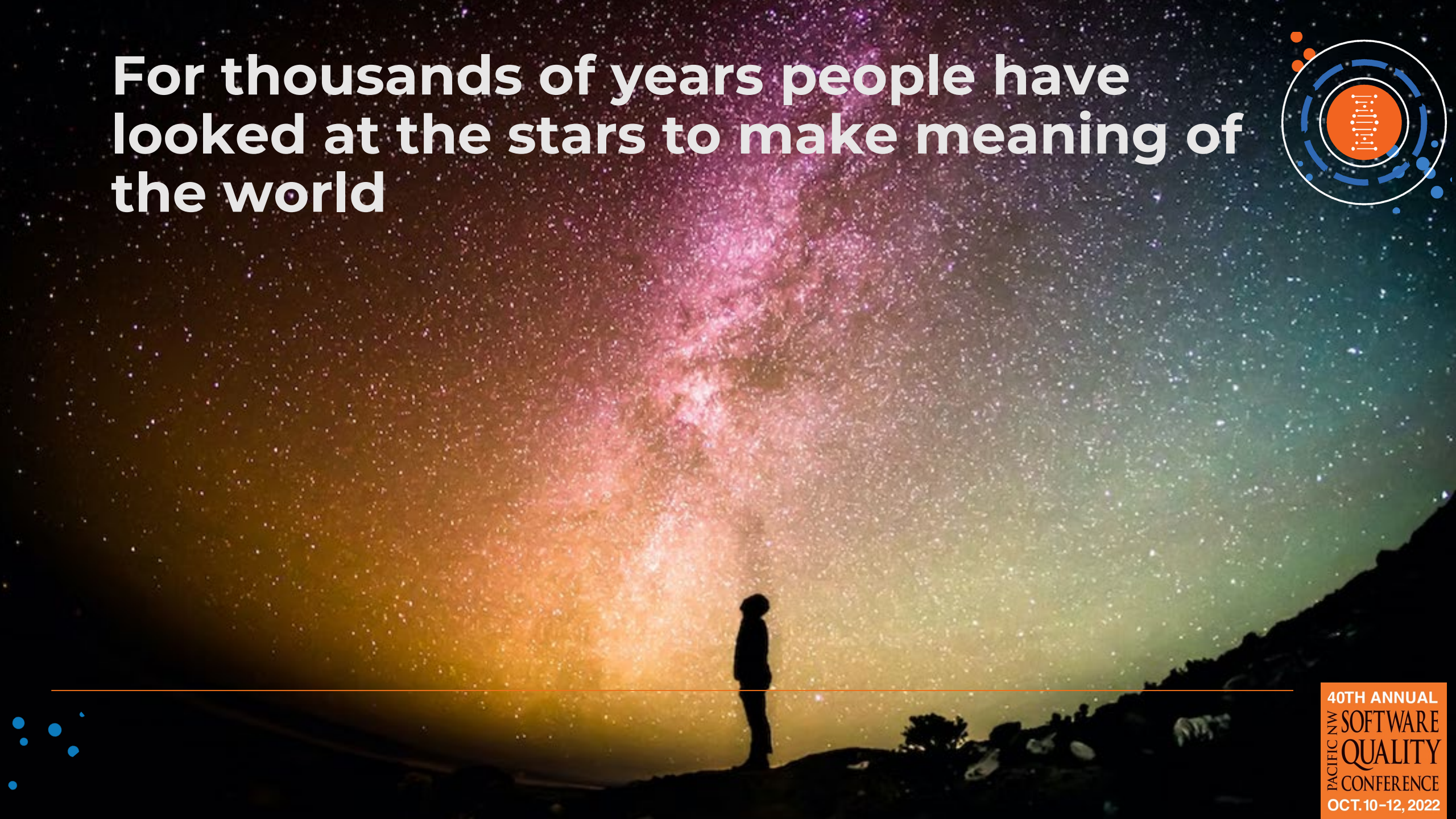
This presentation is *NOT* about...

- AI is wrong
- AI is not useful
- AI should never be used

This presentation is about...

- When AI is wrong
- When AI is useful
- When AI should be used
- When AI should *NOT* be used

For thousands of years people have looked at the stars to make meaning of the world



And as time went on, predictions became more complex

- More data became available
- More cases to observe the effects of the planets
- Greater computational power to calculate planetary positions in instants

Astrological predictions are observational predictions, formed by

- observing features and effects (for example, a happy marriage)
- observing the planetary positions of peoples birth charts (for example, what house a couple's planets are in)
- perhaps even observing the planetary positions of an auspicious date (to ensure a happy marriage)
- And from these observations we may say:

  Venus in the 7th house will produce a happy and harmonious marriage"

AI is also observational, but with big data and supercomputers

- observing features and effects as training data (for example, cat pictures)
- observing other data (for example, dog pictures) observing the planetary positions of peoples birth charts (for example, what house a couple's planets are in)
- training the AI until predictions are accurate (enough).
- And now our AI should be able to predict whether an image is a dog or a cat (within certain confidence interval)

Recent advances in AI rival aspect of human intelligence

Pikachu is battling a Cubone on the surface of the moon (Dall-E)




Jack McDowell

Artificial Intelligence is the New Astrology of Software Quality

40TH ANNUAL PACIFIC NW SOFTWARE QUALITY CONFERENCE
OCT. 10–12, 2022

But aspects of human intelligence don't equal human intelligence

A Comfortable Chair (Dall-E)




Jack McDowell

Artificial Intelligence is the New Astrology of Software Quality

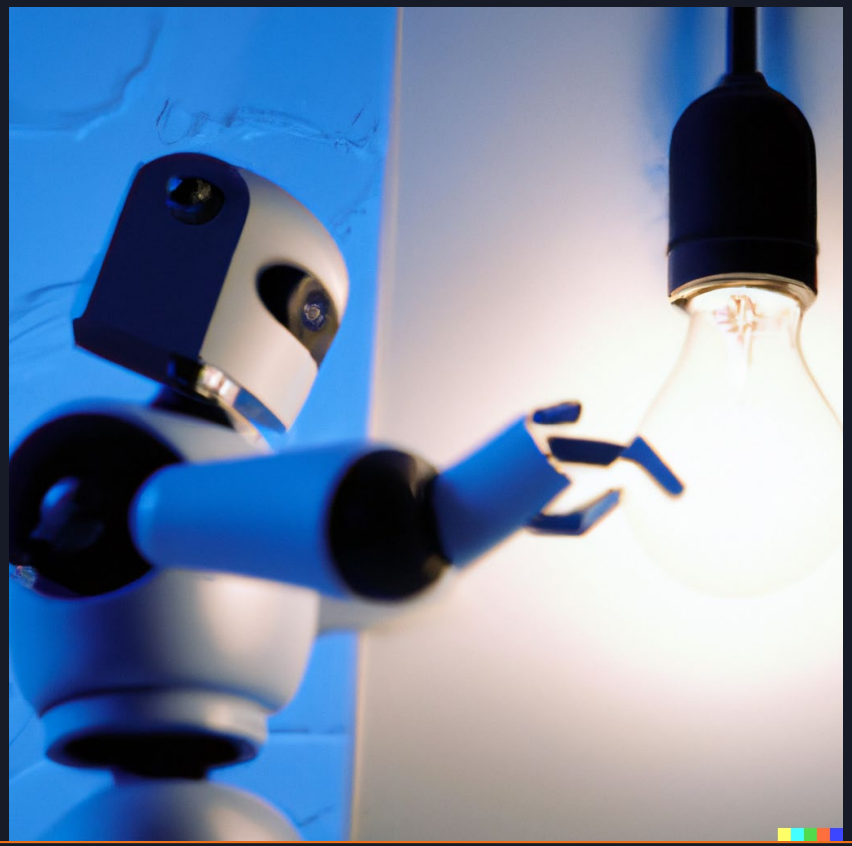
40TH ANNUAL PACIFIC NW SOFTWARE QUALITY CONFERENCE
OCT. 10–12, 2022

Artificial Intelligence Software: How does it differ from Functional Software?

- IFTT, with no statistical variation

As software becomes more complex, more logic is added to the program

Artificial Intelligence Software: How does it differ from Functional Software?



- Prompt:

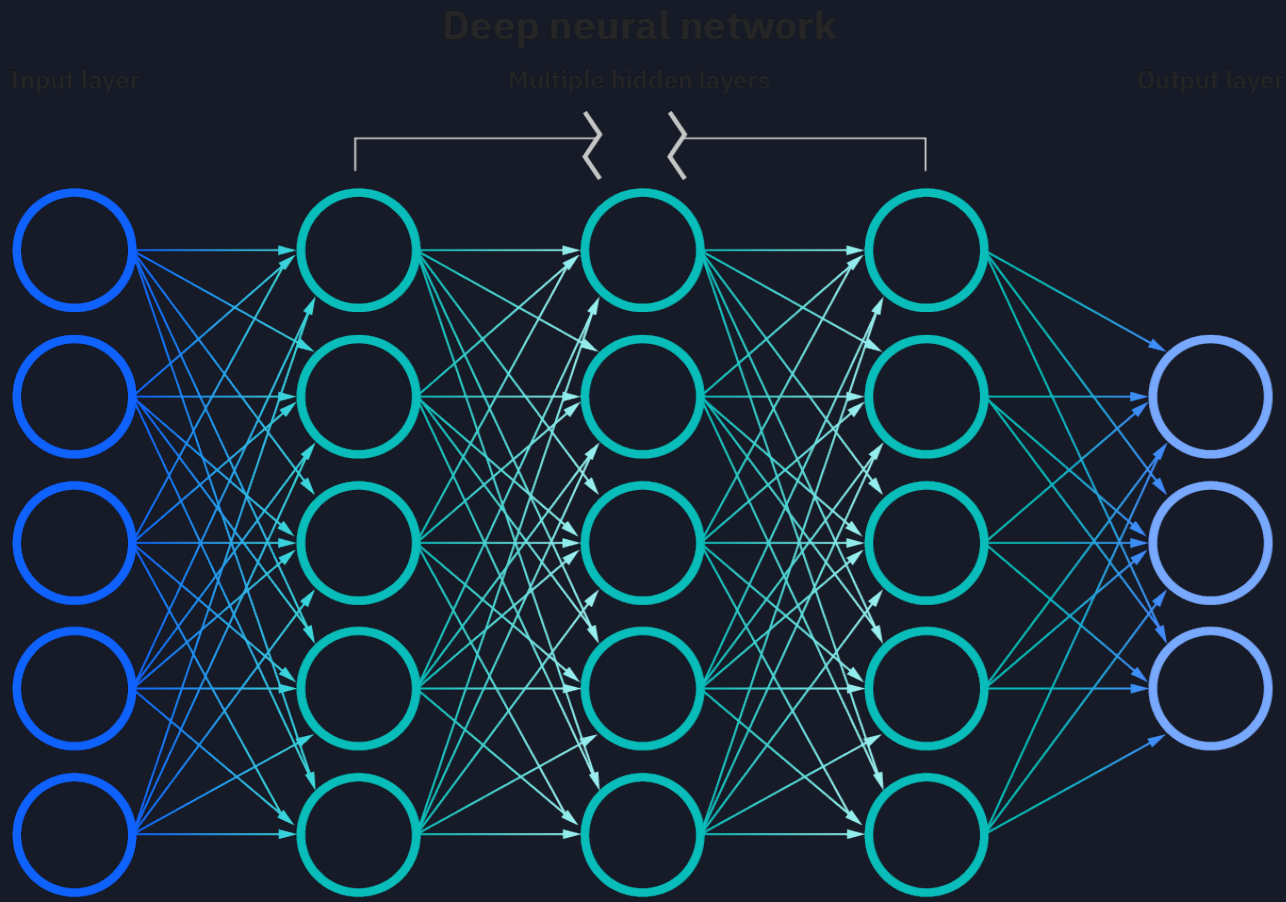
"A robot turning on a light bulb by pressing a switch"

- Microsoft Generated Alt Text:

"A close-up of a microscope"

Description automatically generated with low confidence

Artificial Intelligence Software: How does it differ from Functional Software?

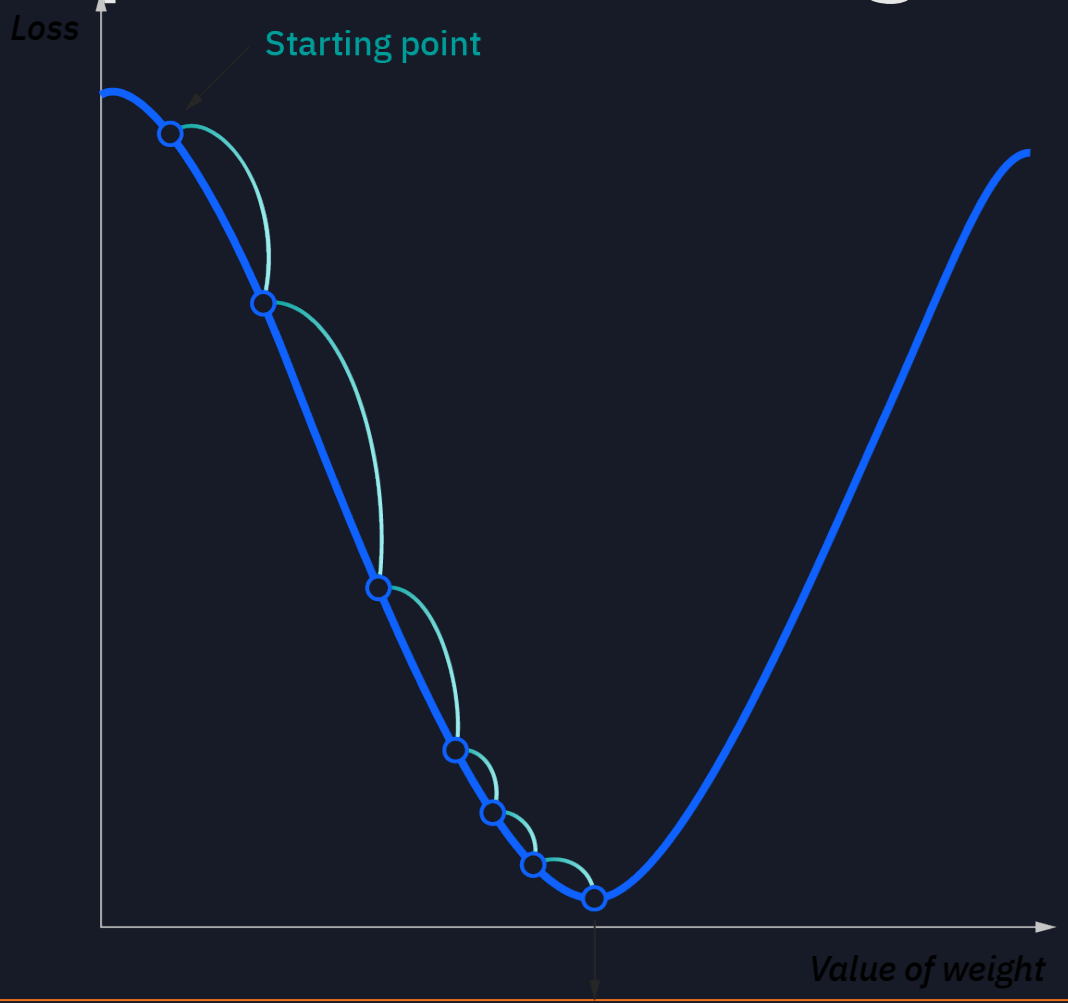


Jack McDowell

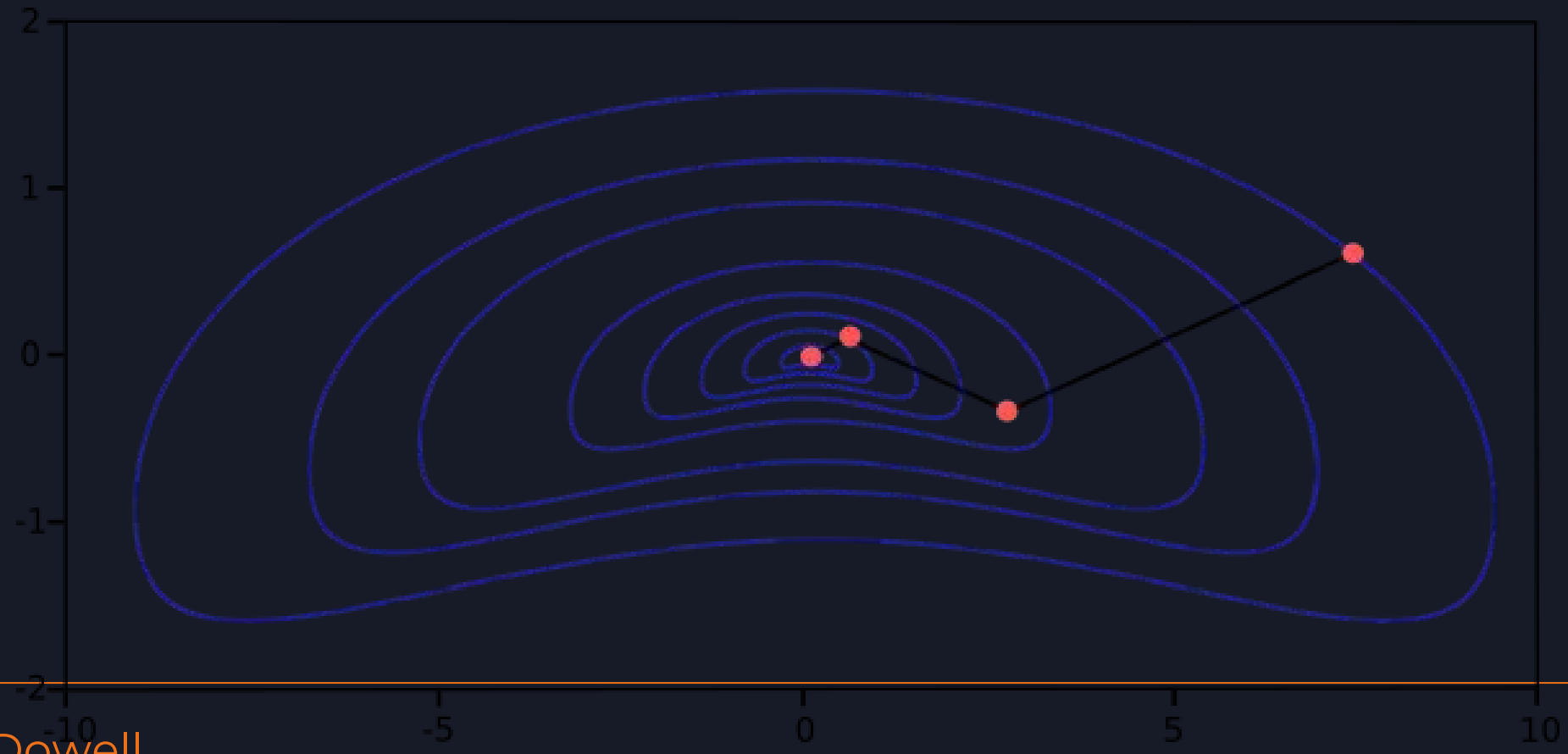
Artificial Intelligence is the New Astrology of Software Quality

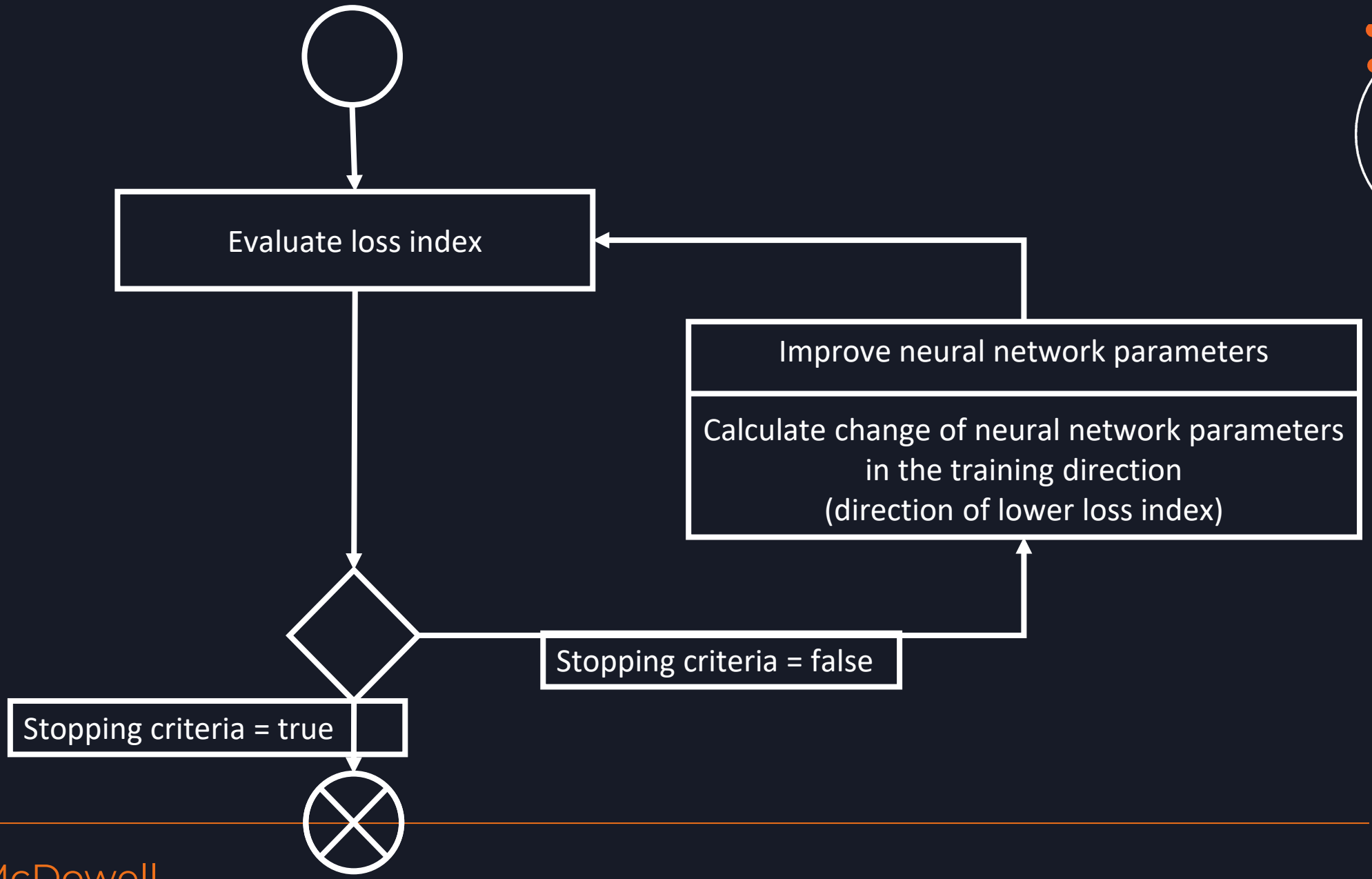
40TH ANNUAL PACIFIC NW SOFTWARE QUALITY CONFERENCE
OCT. 10–12, 2022

Reaching a point of convergence



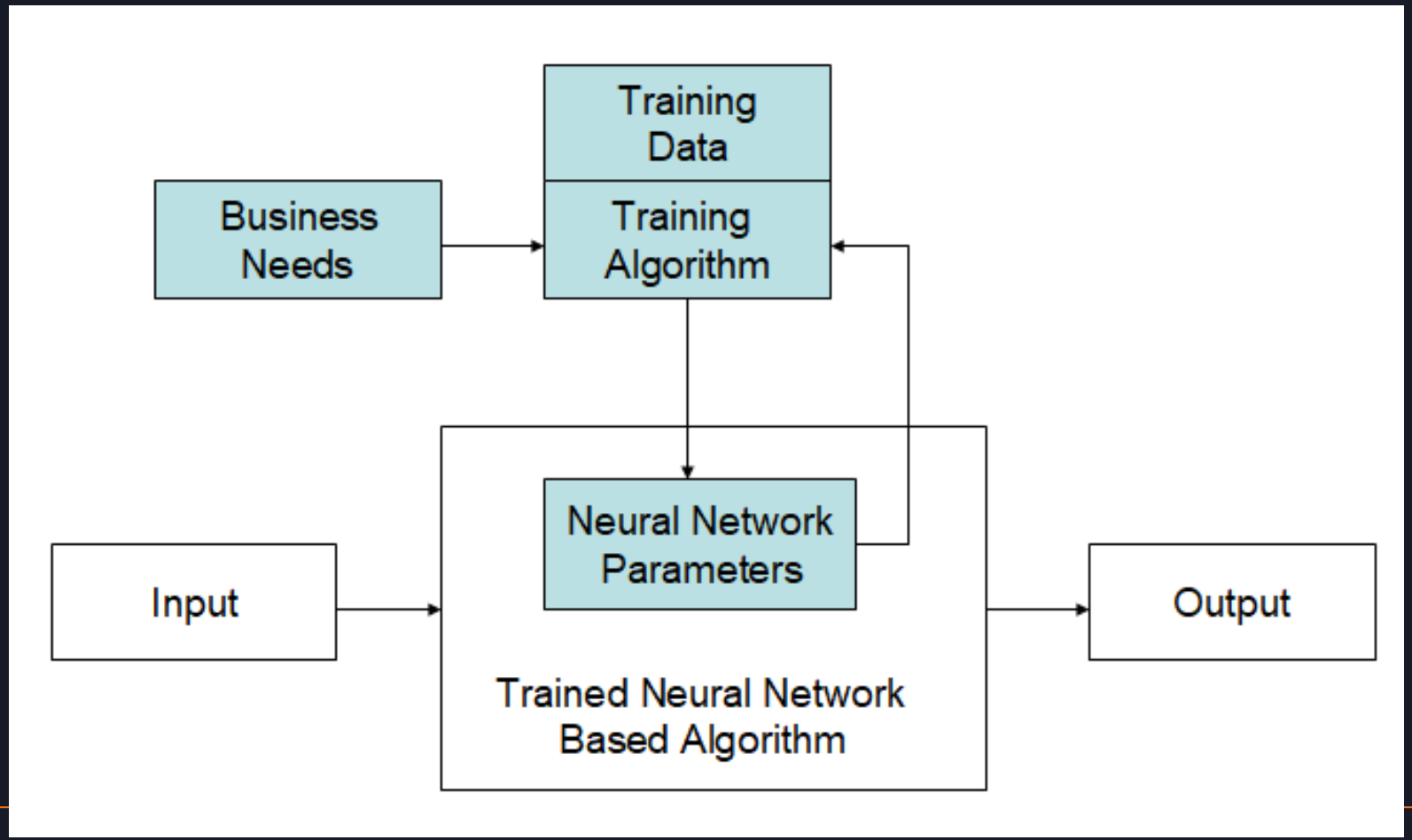
Reaching a point of convergence




Evaluate loss index
Improve neural network parameters
Calculate change of neural network parameters in the training direction (direction of lower loss index)
Stopping criteria = false
Stopping criteria = true

Comparing AI in Software Systems to Traditional Development

Comparing AI in Software Systems to Traditional Development



Artificial Intelligence in Practice

- AI is everywhere:
    - Voice assistant where AI interprets queries
    - Facial recognition
    - Fraud detection
    - Content moderation
    - People moderation
    - Network security
    - SpaceX Rockets

Key Challenges of validating AI systems

- Limited data training and validation. AI algorithms are only validated with limited input data under ad-hoc contexts.
- Data-driven learning features, static or dynamic, that negatively affect software outcomes, results, and actions.
- Inconsistent system outputs, responses, or actions due to uncertainty inherent in statistical models.

Where does AI work well?

- Scientific problems which can be falsified are perhaps the best ones suited to AI. This is because the AI can be given working parameters where out of bounds results are clear, and problems are repeatable experimentally.

- For example:
  - multiplying matrices
  - rocket telemetry
  - playing Go
  - playing Jeopardy!

Where AI could work well

- AI failures are related to poor data sets or poor models.
- Situations where training data biases can be removed, such as facial recognition.
- Low risk situations where an unexpected or unfair outcome can be challenged: e.g.
  - voice assistant: a voice query can be repeated,
  - facial ID: additional verification can be submitted,
  - suspended account: a support ticket can be logged and tracked out of the AI system
- Low risk of illegal or sensitive outcomes.

Would it be logical to train an AI with only Vulcans?

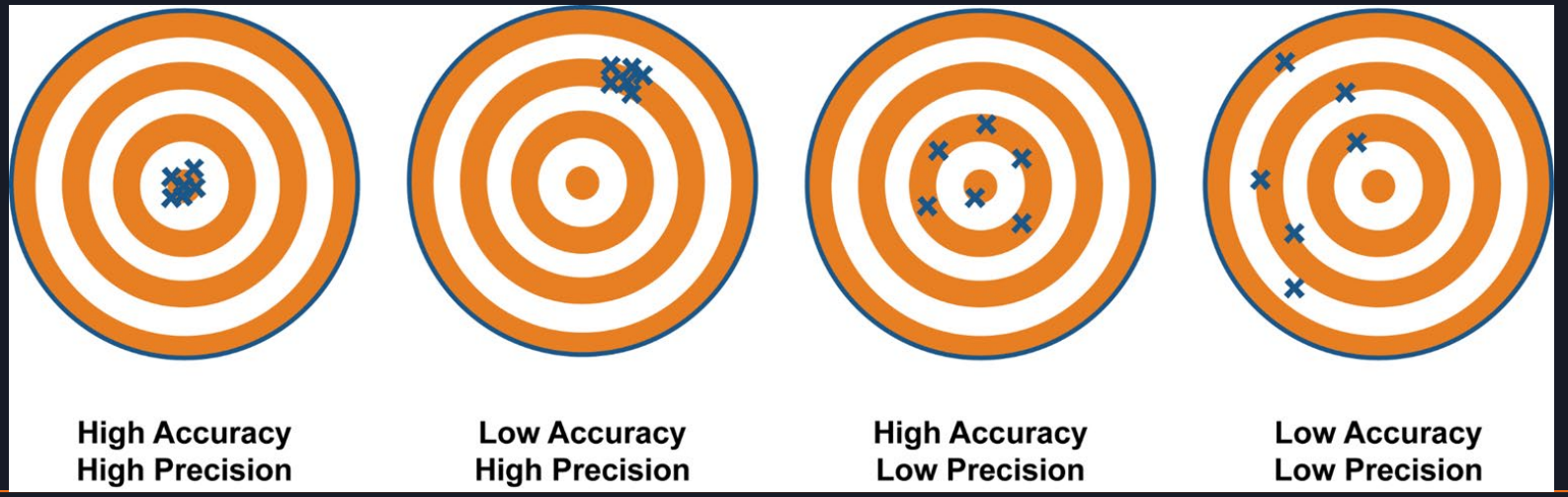
Jack McDowell

Artificial Intelligence is the New Astrology of Software Quality

40TH ANNUAL PACIFIC NW SOFTWARE QUALITY CONFERENCE
OCT. 10–12, 2022

Inherently Inaccurate Algorithms

- When algorithms cannot be improved with additional training data due to limits of knowledge (unknown information, unknown unknowns), and
- Where probabilistic algorithms are legally or morally problematic.



Limits of knowledge

Issues arise with AI when it clashes with limits of knowledge, such as when the training data does not exist or when an AI is asked to solve a new problem.

Think of a black swan: if an AI is presented only with white swans, how would it know it is a swan?

Legal and moral issues

Inherent to probabilistic models, is the legal question as to whether an AI algorithm's conclusion is sufficient to pass legal muster, and whether its findings would hold up in court.

For example, states such as Pennsylvania and Oregon used AI algorithms to assist in the determination of child placement in foster care. After complaints, some of these systems were abandoned.

How can we ensure that these systems are not producing illegal outcomes, which may injure individuals and cause liabilities?

Comparing traditional models to AI

Inmate risk assessment model:

Item	Risk Factor	Codes	Weight
1	Age at intake		0.0322
2	Sentence length (in years)		0.0663
3	Violent offense		0.4766
4	Gang affiliation		2.1492
5	Mental illness		0.5217
6	Custody rating (minimum = 1, medium = 2, close = 3)		0.6504

Suggested Nominal Risk Categories	
Total Score:	**Risk Category:**
6.00 to 0.01	Low-risk
0.00 to 2.38	Moderate-risk
2.39 to 4.00	High-risk

What would an AI program look like?

And how can we determine that it is fair, legal, and unbiased?

Legal and moral issues

- How can we trust that the subject of an AI algorithm's decision fits within the confidence interval and probability of the model, instead of being an outlier?
- How can we unpack the algorithm's determination, given the black box nature of the neural network?
- Can we trust, as a society, an AI algorithm to determine such unique decisions as determining whether a child should be removed from a household and placed in foster care?
- Does big data analytics translate to single cases?

Legal and moral issues

“Surfacing and responding to algorithmic bias upfront can potentially avert harmful impacts to users and heavy liabilities against the operators and creators of algorithms, including computer programmers, government, and industry leaders” (Lee et al. - Brookings Institute).

Parting thoughts

- Traditional software design is based on algorithms whose underlying logic is a priori pre-determined. As such, the resultant logic can be unpacked and analyzed formally to support software verification & validation.

- AI-based software design is based on algorithm whose underlying "logic" is derived from a posteriori (statistical) training procedures that can be repeated in time when more training data become available. As such, the resultant logic cannot be easily unpacked and analyzed formally to support verification & validation beyond statistical correlations.

Typical issues with AI models

- Models may not be good or are poorly trained.
- Models may be well trained but poorly understood.
- Inherent biases are embodied in the training data.
- Lack of accountability with accuracy of models and when models should / should not work.
- Lack of formal understanding about the models beyond a posteriori statistical correlation.

Final thoughts

"It is important for algorithm operators and developers to always be asking themselves: Will we leave some groups of people worse off as a result of the algorithm's design or its unintended consequences?" (Lee, Brookings Institute)

"Decisions impacting millions of people should be fair, transparent and contestable. These new technologies must help us address the major challenges in our world today, such as increased inequalities and the environmental crisis, and not deepening them." (Ramos, UNESCO)

THANK YOU

PNSQC

OCTOBER 10-12 2022

Jack McDowell

jackmcdowell@gmail.com

Artificial Intelligence is the New Astrology of Software Quality