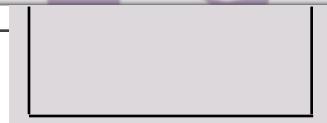


“The passion caused by the great and sublime in nature, when those causes operate most powerfully, is Astonishment; and astonishment is that state of the soul, in which all its motions are suspended, with some degree of horror. [...] When danger or pain press too nearly, they are incapable of giving any delight, and are simply terrible; but *at certain distances, and with certain modifications*, they may be, and they are delightful, as we every day experience.”

Edmund Burke, *Philosophical Enquiry into the Origin of our Ideas of the Sublime and the Beautiful*, 1757

“I say unto you: one must still have chaos in oneself to be able to give birth to a dancing star. I say unto you: you still have chaos in yourselves.”

- Nietzsche, *Thus spoke Zarathustra*, 1883



AI/ML IRL

Joshua Eckroth
Chief Architect / Assistant Professor of Computer Science
i2k Connect / Stetson University

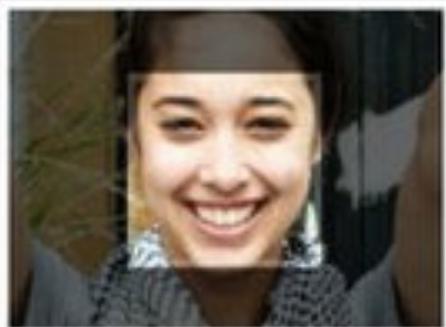




What can I help you with?







Who is this?



Who is this?



Who is this?



Francis Luu

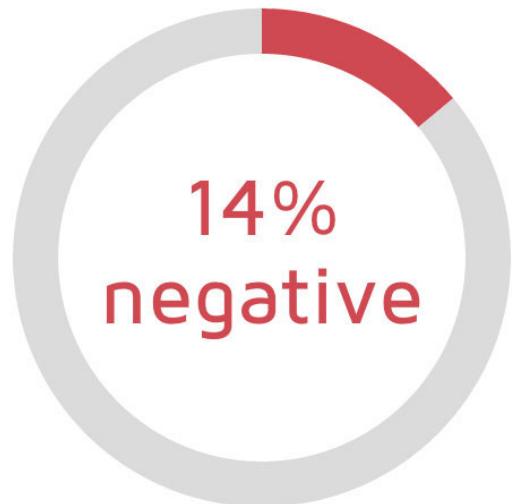
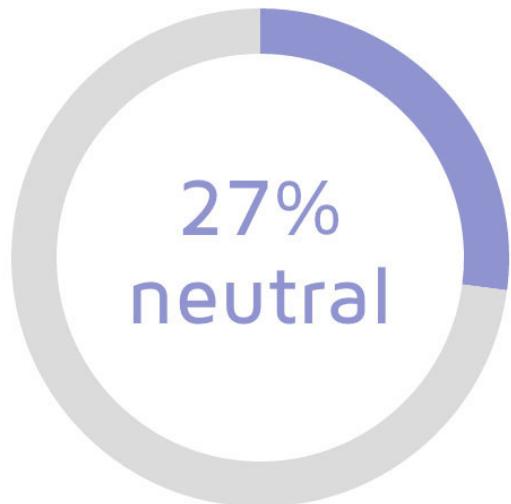
X



Skip Tagging Friends

Save Tags

Breakdown of all tweets



Most Positive Tweets



@SingInTheRine We're so happy you found us!
Feel free to send a message if you need any help,
and let us know what you end up building :)

99.4%

Most Negative Tweets



Why Overfitting is More Dangerous Than Just
Poor Accuracy: <http://t.co/ni6GdsAGXN>
#machinelearning <http://t.co/0A3wrvyLaO>

3.1%

A



B



C



D



E



F



Certain, closed systems:

- Well-defined inputs (e.g., bounded integers)
- Well-defined transformations and calculations

Certain, open systems:

- Open-ended inputs (e.g., user comments, speech, selfies, etc.)
- Well-defined transformations and calculations

Uncertain systems:

- **Open-ended inputs**
- **Uncertain transformations and calculations**





DO I NEED AI?

“[T]he technology led a lot of the things that happened there [...] we would like to make a game like this [...] let's see what technology we can do that's in this quadrant of game development space, and then figure out what game we can make with what we actually wind up with.”

John Carmack

<http://www.wired.com/2013/12/john-carmack-doom/>

Do you “need” AI because you want to have AI?

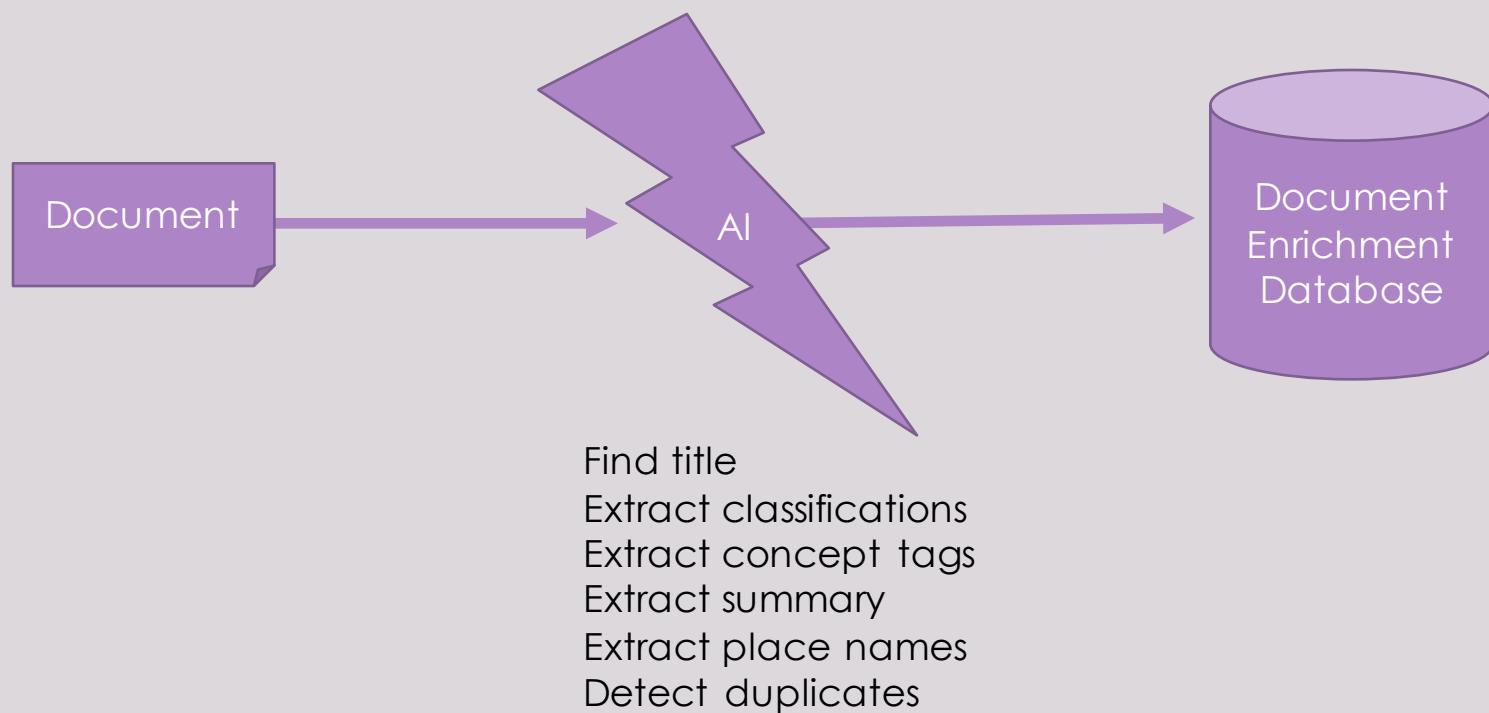




KEEP
CALM
**CLUTTER
GONE**



KEEP
CALM
**DATA
FOUND**



Do I need AI?

- **Online learning / planning**
- **Processing or generating forms of human communication**
- **Detecting or recognizing natural or human artifacts**
- **Encoding expert knowledge**
- **Fast, heuristic, “satisficing” search in massive search spaces**
- **Physical presence, operating among humans or the natural world**



AI = TECHNICAL DEBT

Machine Learning: The High-Interest Credit Card of Technical Debt

**D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov,
Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young**
{dsculley, gholt, dgg, edavydov}@google.com
{toddphillips, ebner, vchaudhary, mwyoung}@google.com
Google, Inc

Abstract

Machine learning offers a fantastically powerful toolkit for building complex systems quickly. This paper argues that it is dangerous to think of these quick wins as coming for free. Using the framework of *technical debt*, we note that it is remarkably easy to incur massive ongoing maintenance costs at the system level when applying machine learning. The goal of this paper is highlight several machine learning specific risk factors and design patterns to be avoided or refactored where possible. These include boundary erosion, entanglement, hidden feedback loops, undeclared consumers, data dependencies, changes in the external world, and a variety of system-level anti-patterns.

“[M]achine learning packages have all the basic code complexity issues as normal code, but also have a larger system-level complexity that can create hidden debt.

Thus, refactoring these libraries, adding better unit tests, and associated activity is time well spent but does not necessarily address debt at a systems level.”

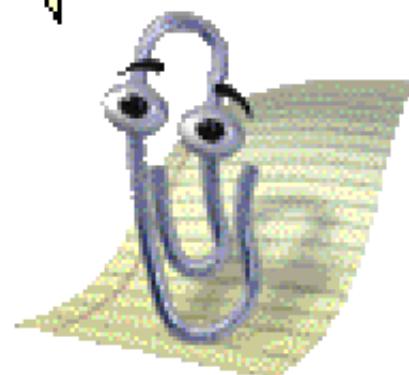
- Machine learning may “subtly erode abstract boundaries”
 - i.e., using outputs from a predictor in other parts of code
 - This creates a tight coupling
- ML is often treated as a black box
 - Results in lots of glue code and calibration code that locks in assumptions
- Changes in external world may make behavior change in unintended ways
 - Ratcheting up maintenance costs
- Monitoring that the system is performing as intended may be difficult
 - Need careful design to put bounds on the chaos a bad ML model can cause



Tay.ai

It looks like you're giving a presentation. Would you like some assistance?

- Yes, please find a way to embarrass me.
- No thanks, I'm handling it ok.



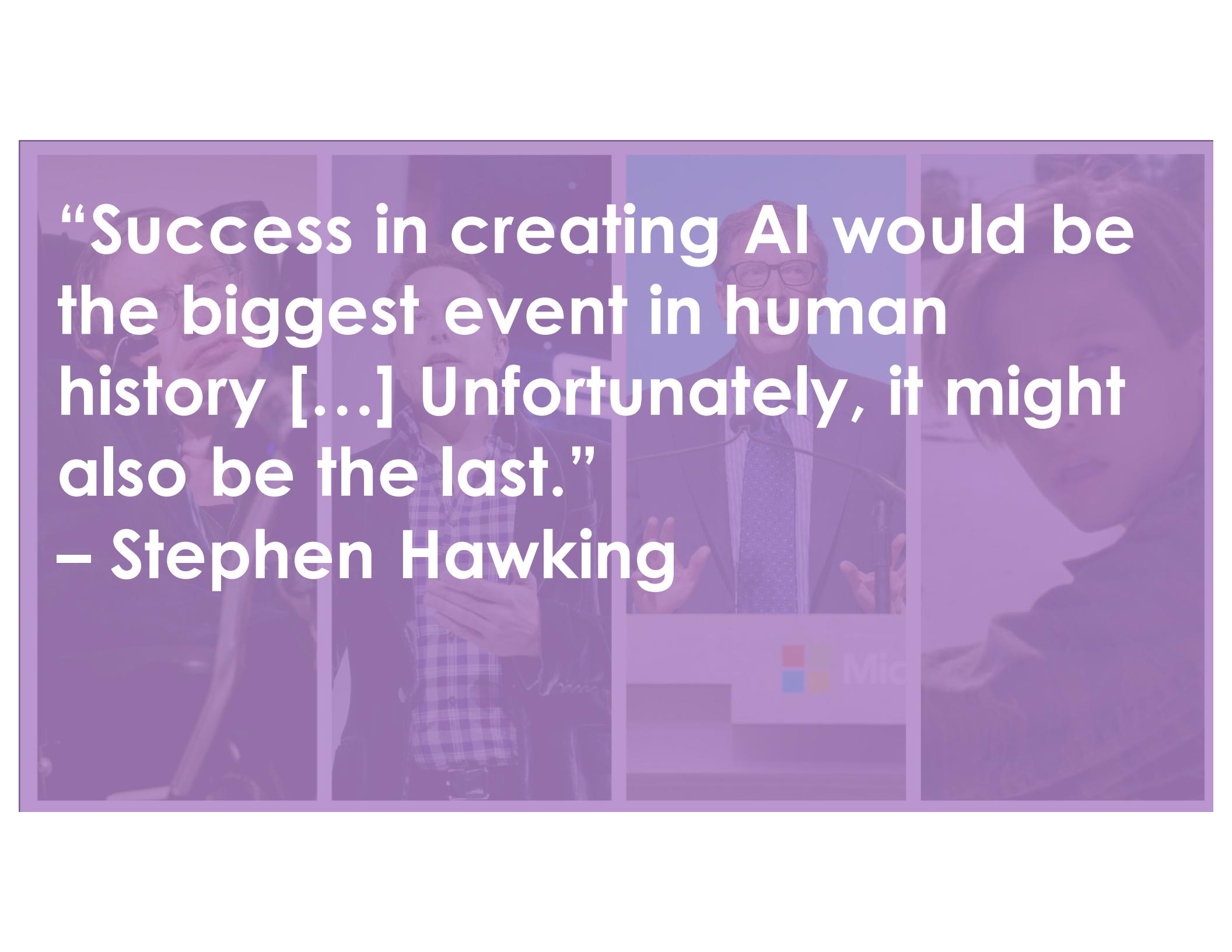
Email Management – Microsoft Clutter

Posted on March 4, 2016 by Lucy Piva

In our efforts to improve service and simplify our systems, we will be disabling the “Clutter” feature in your mailboxes.

Microsoft recently added the “Clutter” feature which attempts to learn our email behaviors and assist us by moving what it deems “not important” into a folder titled “Clutter.” Due to reports of concerns from many users about this new feature, we will be disabling this feature on all accounts as of March 9.



A collage of four images showing people interacting with Microsoft Surface devices. The images include a person in a dark suit, a man in a plaid shirt, a woman in a patterned dress, and a close-up of hands using a tablet. The background is a light purple.

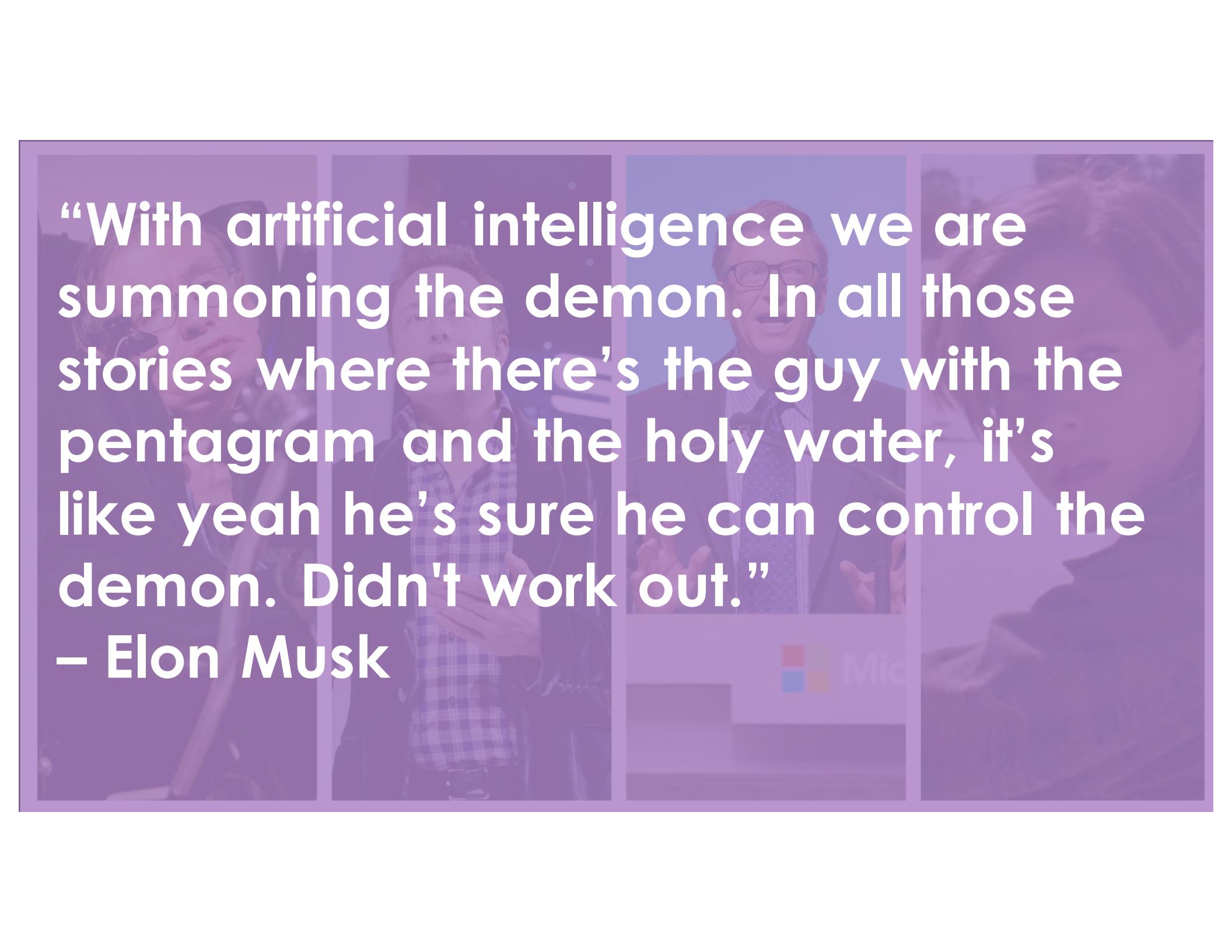
**“Success in creating AI would be
the biggest event in human
history [...] Unfortunately, it might
also be the last.”**

– Stephen Hawking



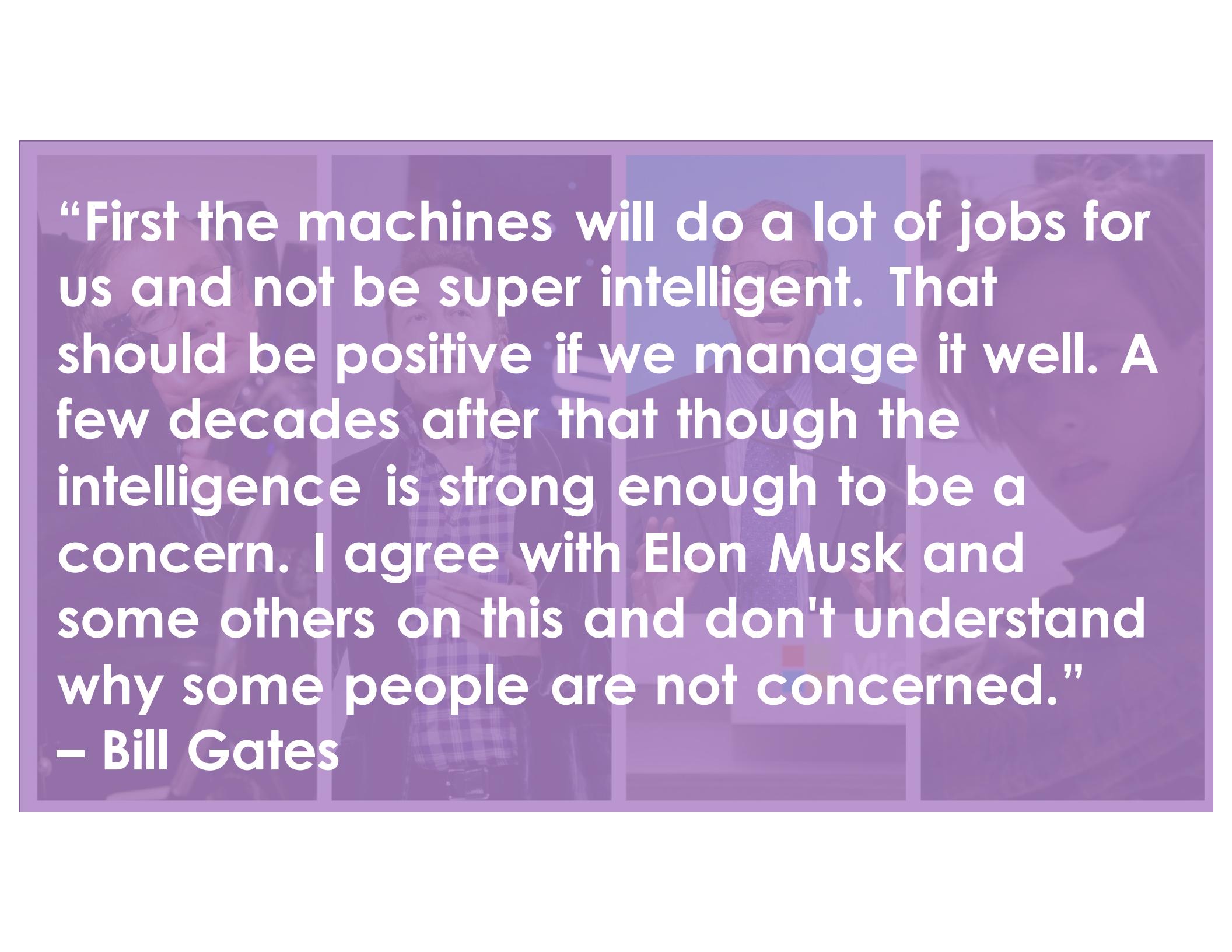
“I think we should be very careful about artificial intelligence. If I were to guess like what our biggest existential threat is, it's probably that.”

– Elon Musk

A collage of four images featuring Elon Musk. The top-left image shows him from the chest up, wearing glasses and a dark jacket. The top-right image shows him from the waist up, wearing a light-colored shirt and a tie. The bottom-left image shows him from the waist up, wearing a plaid shirt and jeans. The bottom-right image shows him from the chest up, wearing a dark suit and tie.

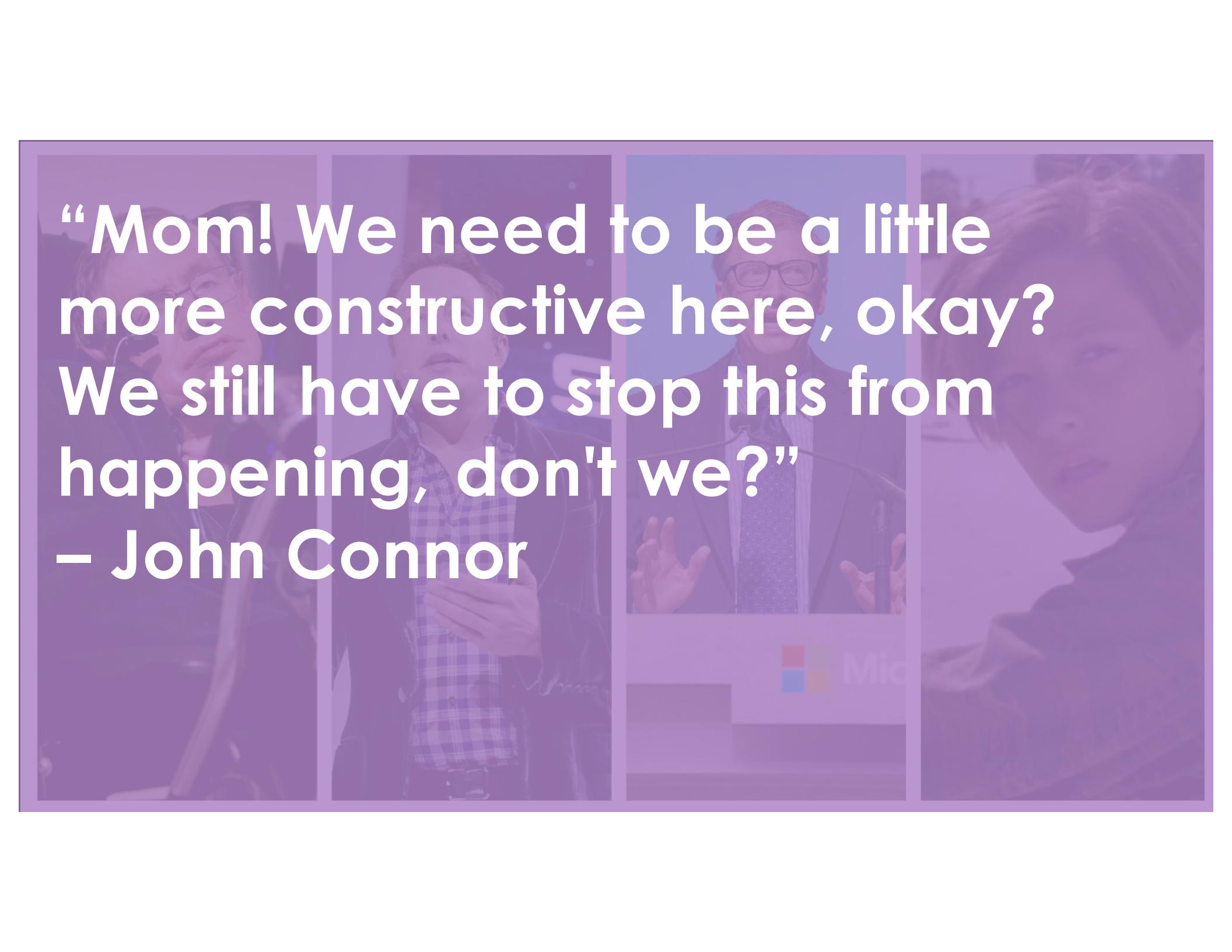
“With artificial intelligence we are summoning the demon. In all those stories where there’s the guy with the pentagram and the holy water, it’s like yeah he’s sure he can control the demon. Didn’t work out.”

– Elon Musk



“First the machines will do a lot of jobs for us and not be super intelligent. That should be positive if we manage it well. A few decades after that though the intelligence is strong enough to be a concern. I agree with Elon Musk and some others on this and don't understand why some people are not concerned.”

– Bill Gates



**“Mom! We need to be a little
more constructive here, okay?
We still have to stop this from
happening, don't we?”**

– John Connor

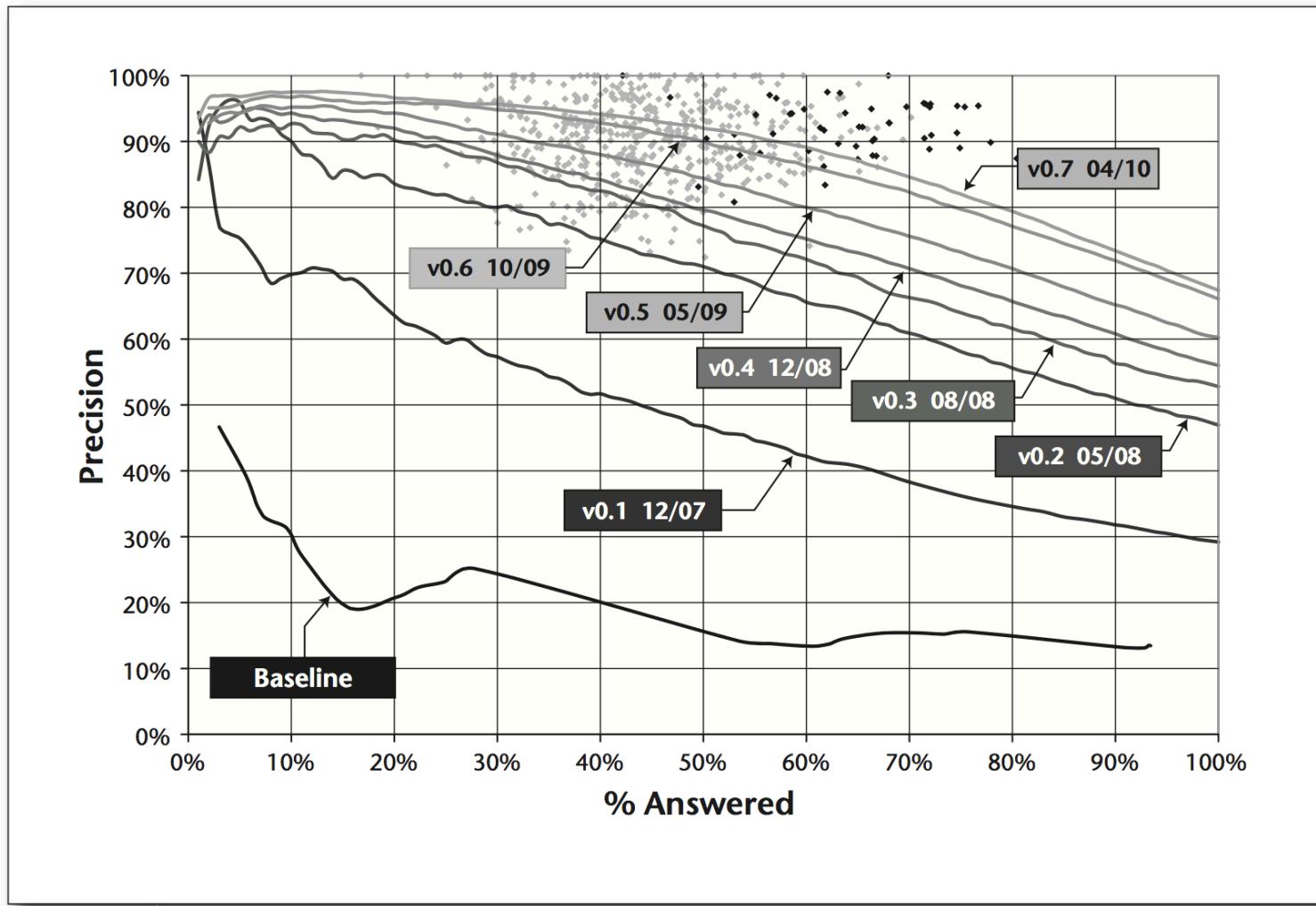
“How can AI be dangerous?

- The AI is programmed to do something devastating.
 - The AI is programmed to do something beneficial, but it develops a destructive method for achieving its goal.”
- Future of Life Institute

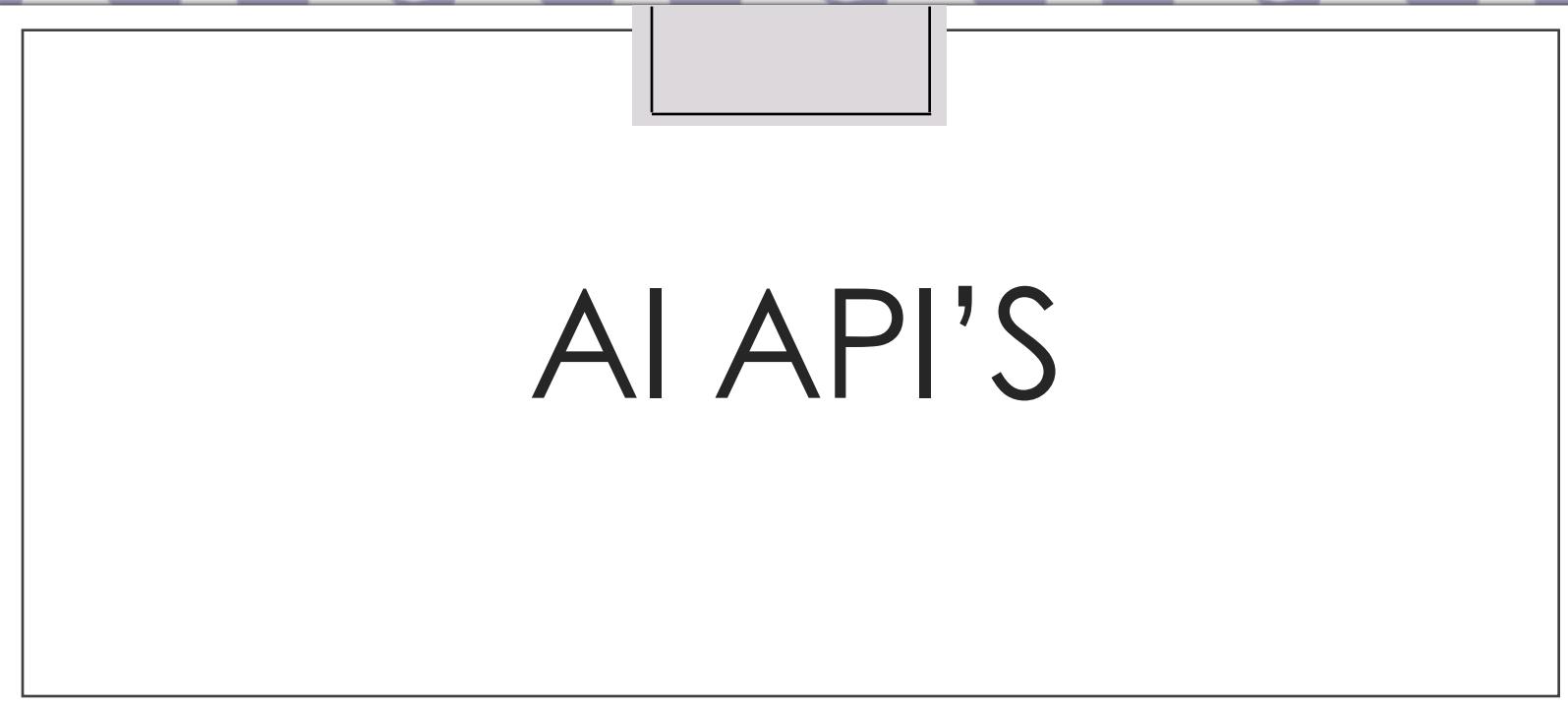


ADVICE

- 1. Develop a set of “ground truth” examples**
 - These can be unit tests!
 - Can also serve as a performance metric (% accuracy, etc.)
- 2. Add safeguards**
 - Check AI's output for dangerous conditions (bad words, extreme values, etc.)
- 3. Limit the impact of AI on the system**
 - AI output can be unpredictable; don't depend on a small range of outputs
 - Other code may break if it makes wrong assumptions about AI output
- 4. Provide an alternative UI pathway**
 - Support disabling Clippy
 - Show data that was deemed “irrelevant” in some less-visible menu
- 5. Be very cautious about making assumptions for people**
 - You could be wrong
 - People hate when your system ignores their intentions
- 6. Be very very cautious about online learning**
 - At least pay attention to the system – it could be evolving in the wrong direction
- 7. Don't put your learning bot on Twitter**
 - i.e., don't let the larger world influence your AI system in a feedback loop
 - Just don't.



Ferrucci, D.; Brown, E.; Chu-Carroll, J.; Fan, J.; Gondek, D.; Kalyanpur, A. A.; Lally, A.; Murdock, J. W.; Nyberg, E.; Prager, J. & others. Building Watson: An overview of the DeepQA project. *AI magazine*, **2010**, 31, 59-79



AI API'S

Kitchen sink APIs

- **Microsoft Project Oxford**
 - vision: image categories, thumbnail generation (with proper crop), OCR, face detection and recognition, emotion recognition; video: face detection and tracking, auto stabilization, motion detection; speech: voice commands, speaker recognition; language: parse natural language queries
- **IBM Watson**
 - text: keyword extraction, sentiment analysis, named entity recognition, translation; speech: tts, recognition; vision: find objects, people, text in images

Vision APIs

- **Google Vision**
 - concept extraction, face/object detection and recognition, detect offensive content, image sentiment analysis
- **Clarifai**
 - concept extraction (tagging) from images and video

Voice / natural language APIs

- **Api.ai**
 - voice interface; text-to-speech (tts), natural language queries (text or sound), supports user-defined synonyms
- **Wit.ai**
 - voice interface; very similar to Api.ai
- **MindMeld**
 - voice interface; integrates with existing speech recognition (e.g., in ios); supports defining custom “knowledge graph” that matches commands specific to your app
- **Chatbots.io**
 - chatbots, customer care agent bots



AI LIBRARIES

Javascript AI libraries

- **Good luck**
- There are some weak candidates:
 - www.npmjs.com/search?q=artificial-intelligence
- Most AI work requires a model training phase
 - This is computationally expensive
 - A mobile device or website won't do model training
 - Why do this in Javascript?
- Better: create a backend web-service to host the AI
 - Use a language with a solid AI library: Python, Java, C++, etc.

Python AI libraries

- **scikit-learn** (scikit-learn.org)
 - Classification: identify the category of text, images, ...
 - Regression: predict a value, such as a stock price, weather, user ratings, ...
 - Clustering: find natural groupings to predict user preferences, document similarity, ...
- **Theano** (deeplearning.net/software/theano)
 - Deep learning: tag photos, identify faces, extract meaning from text, voice search, speech transcription, ...

Java AI libraries

- **Weka** (www.cs.waikato.ac.nz/ml/weka)
 - Classification
 - Regression
 - Clustering
 - Spiffy GUI for trying out various techniques before writing code
- **OpenNLP** (opennlp.apache.org)
 - Text processing: finding names, organizations, dates, etc.
- **Deep Learning for Java** (deeplearning4j.org)



SUMMARY

