

Designing Human-Centered AI Products

October 21, 2019

Jess Holbrook, PhD
Co-lead, People + AI Research at Google
Content credit to Kristie J. Fisher



Asynchronous Stochastic Optimization for Sequence Training of Deep Neural Networks

Venue

Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), IEEE, Firenze, Italy (2014)

Publication Year

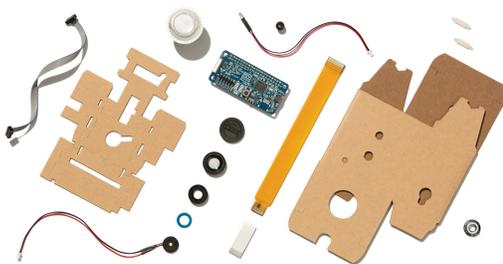
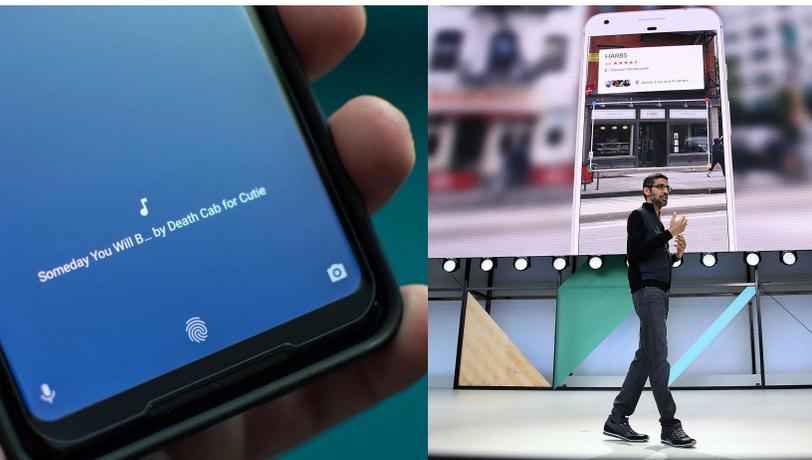
2014

Authors

Georg Heigold, Erik McDermott, Vincent Vanhoucke, Andrew Senior, Michiel Bacchiani

Abstract

This paper explores asynchronous stochastic optimization for sequence training of deep neural networks. Sequence training requires more computation than frame-level training using pre-computed frame data. This leads to several complications for stochastic optimization, arising from significant asynchrony in model updates under massive parallelization, and limited data shuffling due to utterance-chunked processing. We analyze the impact of these two issues on the efficiency and performance of sequence training. In particular, we suggest a framework to formalize the reasoning about the asynchrony and present experimental results on both small and large scale Voice Search tasks to validate the effectiveness and efficiency of asynchronous stochastic optimization.

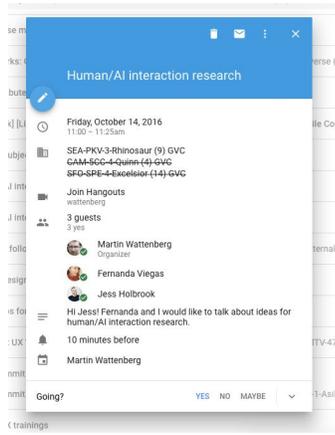


Sundar, 2018 Founders Letter

One of our clearest insights on AI so far is that its potential is greatest when paired with human intelligence. [...] We believe that we can develop AI in a way that complements human expertise, and we feel a deep responsibility to get this right.



One day Jess gets an invite from two Research Scientists on the Brain team who want to talk about "Human/AI interaction research."



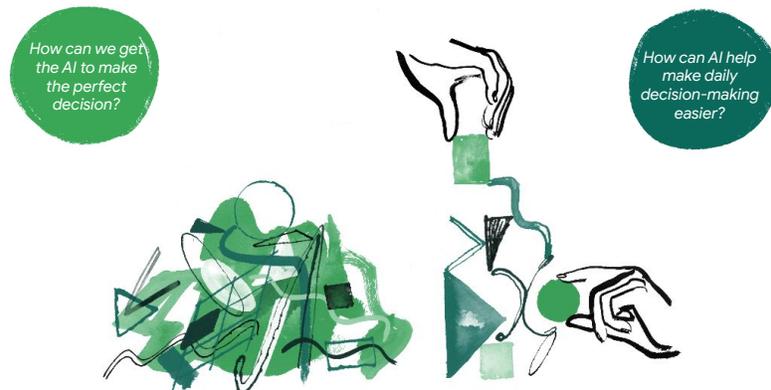
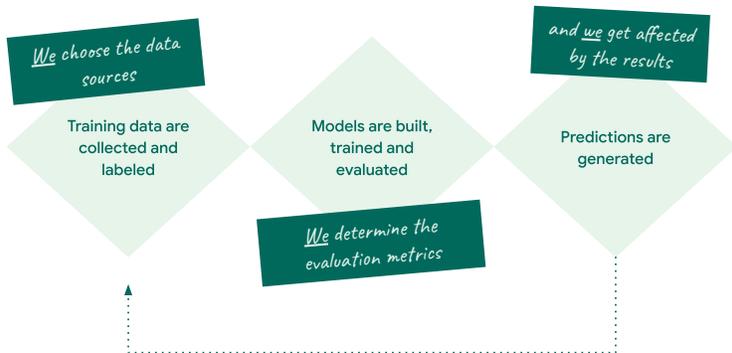
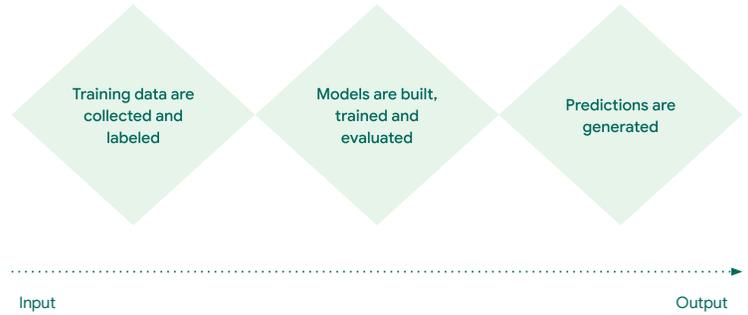
People + AI Research

UX & HCI

Seattle, LA, NYC, SFO, Boston, Accra, Bangalore

ML Research & Engineering

Cambridge, Mountain View, NYC, DC, London



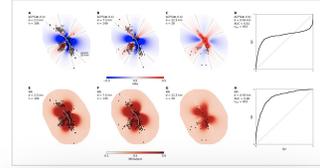
PAIR does 3 things

1. We conduct and publish human-AI interaction research.
2. We create and launch open-source tools and platforms to build AI responsibly.
3. We widen the circle of who can participate in the development and application of AI.

PAIR does 3 things

1. We conduct and publish human-AI interaction research.

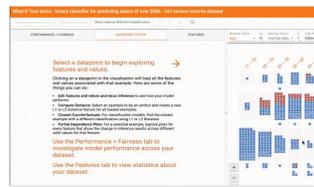
...to push the fields of ML and AI, HCI, and human-centered design forward. We research AI explainability, interpretability, fairness, data visualization, human-centered AI, AI-assisted healthcare, and ML for science discovery. We conduct research independently and in partnership with leading academics in multiple fields.



PAIR does 3 things

2. We create and launch open-source tools and platforms to build AI responsibly.

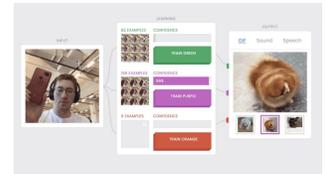
...to enable the builders of these systems to understand the consequences of their decisions, simulate the effects of different decisions, and make building AI responsibly the easiest option.



PAIR does 3 things

3. We widen the circle of who can participate in the development and application of AI.

...to increase participation and agency in how AI will affect our lives. We engage through multiple channels including building development tools, writing AI explainers and demos, inviting external presentations, and art installations.



Introducing the

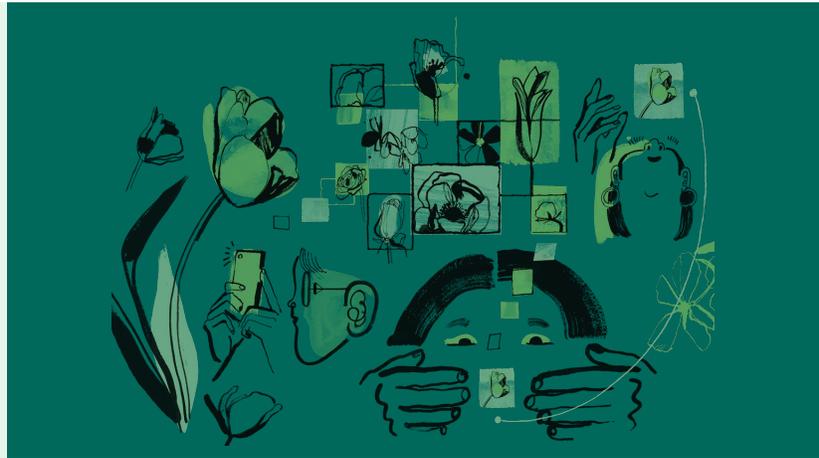
People + AI Guidebook

Designing human-centered AI products



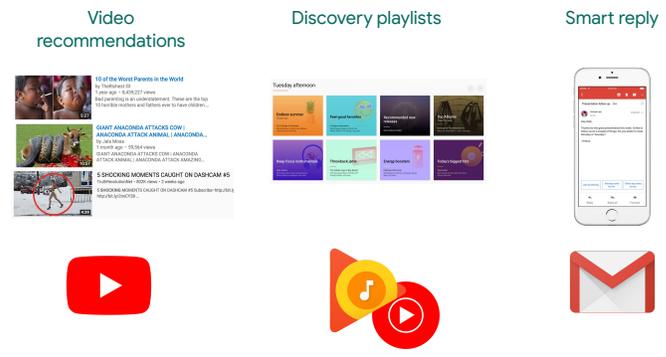


design.google/ai-guidebook



Myth 1

If the AI system needs to ask for explicit user feedback to succeed, then it has already failed.



Reframed

To properly measure AI success and improve the system, explicit feedback is typically required.

Feedback + control

Feedback best practices

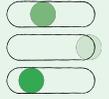
Dedicate sufficient time for testing & iterating on feedback systems - must work for users and for the AI

Align as a team: if we get feedback that ____, then we will change ____

Communicate value & time to impact

design.google/ai-guidebook

#paiguidebook



Feedback + control

Implicit + explicit feedback

What user insights or AI model improvements might you be missing out on due to lack of or under-utilized feedback systems?

Is feedback system development properly resourced?

design.google/ai-guidebook

#paiguidebook

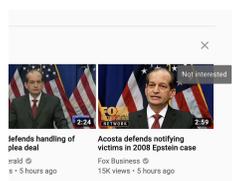
Myth 2

*AI prediction **errors are the worst**. Minimize or ignore them and move on.*

People auto-tagging



Dismissing "incorrect" video suggestions



Reframed

*"Failure" is an inherent part of a probabilistic system like AI. **Errors are an opportunity** to gather feedback and calibrate user trust. They show the **boundaries** of the AI.*

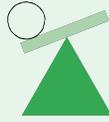
Errors + graceful failure

Errors must be defined, with user help

“Context errors” are unique to AI

Incorporate explanations into errors

Paths forward: feedback, manual override



Recovering from errors

Are there common errors in our products that we're not using as opportunities?

How confident are we that the team and the users are defining “error” in the same way?

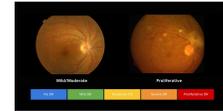
Myth 3

“Helpful” AI always means doing things **for** the user.

Email actions



Diagnosis assistance



Search ad settings, Universal campaigns



Reframed

Helpful is helpful. Only automate when you're extremely confident people don't want a choice.

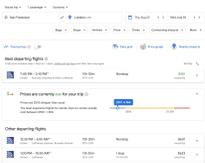
Automate	Augment
People lack the knowledge or ability to do the task	People enjoy the task
Tasks are boring, repetitive, awkward, or dangerous	Personal responsibility for the outcome is required or important
	The stakes of the situation are high
	Specific preferences are hard to communicate

Automation vs. augmentation

Are there processes we're trying to automate for our users where we should really be augmenting their unique human capabilities?

Do we have sufficient insights about our users regarding their automation needs and concerns?

Price predictions



Directions



Performance forecasts



Myth 4

People need to trust us more.
We should always be trying to increase trust.

Reframed

People need to develop **calibrated** trust because AI won't always be right, and this takes time.

Calibrated trust

Do we have a good understanding of users' current level of trust in our Smart Home systems? Any signs of avoidance or of over-trust?

How are we currently helping users calibrate their trust appropriately?

Myths		Reframed
If the AI system needs to ask for explicit user feedback to succeed, then it has already failed.	>>	To properly measure AI success and improve the system, explicit feedback is usually required.
AI prediction errors are the worst. Minimize or ignore them and move on.	>>	Failure is an inherent part of a probabilistic system like AI. Errors are an opportunity to gather feedback and calibrate user trust.
"Helpful" AI always means doing things for the user.	>>	Helpful is helpful. Only automate when you're extremely confident people don't want a choice.
People need to trust us. We should always be trying to increase trust.	>>	People need to develop calibrated trust because AI won't always be right, and this takes time.

Thank you!

jessh@google.com



Designing Human-Centered AI Products

Guidebook Resources

People + AI Guidebook

Designing human-centered AI products

User Needs + Defining Success

Identify user needs, find AI opportunities, and design your reward function.

Data Collection + Evaluation

Decide what data are required to meet your user needs, source data, and tune your AI.

Mental Models

Introduce users to the AI system and set expectations for system-change over time.

Explainability + Trust

Explain the AI system and determine if, when, and how to show model confidence.

Feedback + Control

Design feedback and control mechanisms to improve your AI and the user experience.

Errors + Graceful Failure

Identify and diagnose AI and context errors and communicate the way forward.

Data Collection + Evaluation Chapter worksheet

Instructions

Block out time to get as many cross-functional leads as possible together in a room to work through these exercises & checklists.

Exercises

1. Get to know your data

Decide what kind of data you need, whether or not it already exists, and understand the sources.

2. Dataset essential checklist

Walk through the checklist to prepare your data for use in your AI product.

3. Design for your raters

Data collection + evaluation

1. Get to know your data

The first task your team has to complete is to identify the type and scope of data needed to train an ML model that can meet your users' needs.

User needs & data needs template

Use this template for each unique user need your ML model will impact.

Example: building a recipe recommendation service that suggests new dishes to cook.

User needs & data needs	
Users	Home chefs
User need	Try new recipes or cuisines
User action	Click a new dish using the recipe-based on recommendation
ML system output	Recommendations for new recipes
ML system learning	Patterns of behavior around choosing recipe recommendations
Training dataset needed	Get of recipes user has previously found, used, and liked

pair.withgoogle.com/resources

People + AI Guidebook

Introduction

User Needs + Defining Success

Data Collection + Evaluation

Mental Models

Explainability + Trust

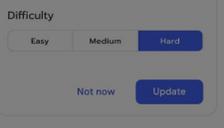
Feedback + Control

What's new when working with AI

Align feedback with model improvement

Communicate value & time to impact

Balance control & automation



✓ Aim for

Allow users to adjust their prior feedback and reset the system. [Learn more](#)

Key concept

Take time to think about your users' expectations for control over certain tasks or processes. Use this quick checklist for you and your team to run through before setting up feedback and control.

Feedback