Wizarding Interaction

March 8, 2021

THIS SLIDE IS HERE

TO REMIND WENDY

TO START ZOOM RECORDING

AND TRANSCRIPTION

LAB STAR NOMINATIONS?

HISTORY OF WIZARDING

CHI'83 Proceedings December 1983

An empirical methodology for writing User-Friendly Natural Language computer applications

J. F. Kelley
IBM
Thomas J. Watson Research Center
Yorktown Heights, NY 10598

A six-step, iterative, empirical, human factors design methodology was used to develop *CAL*, a natural language computer application to help computer-naive business professionals manage their personal calendars. Language is processed by a simple, non-parsing algorithm having limited storage requirements and a quick response time. *CAL* allows unconstrained English inputs from users with no training (except for a 5 minute introduction to the keyboard and display) and no manual (except for a two-page overview of the system). In a controlled test of performance, *CAL* correctly responded to between 86% and 97% of the inputs it received, according to various criteria. This research demonstrates that the methodological tools of the engineering psychologist can help build user-friendly software that accommodates the unruly language of computer-naive, first-time users by eliciting the cooperation of such users as partners in an iterative, empirical development process.

The principal purpose of the research reported here was to design and test a systematic, empirical methodology for developing natural language computer applications. This paper describes that methodology and its successful use in the development of a natural language computer application: *CAL*, *C*alendar *A*ccess *L*anguage. The limited con-

6. Cross-validation. The final program was tested with six additional participants to see how well it performed. In this step, the program ran without any assistance from the experimenter. Various measures of program speed, "understanding" and efficiency were combined with the results of post-session interviews to evaluate

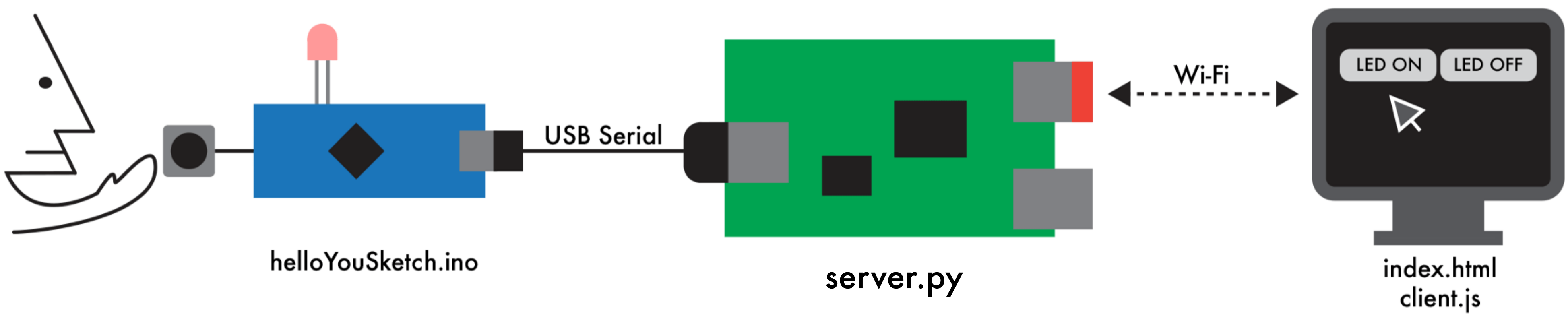
WIZARD OF OZ IN NATURAL LANGUAGE INTERFACES, J.F. KELLEY (1983)

A P P L E   K N O W L E D G E   N A V I G A T O R (1 9 8 7)

WIZARDING TO BOOTSTRAP INTERACTIONS

Interaction Engine

Microcontroller+Microproccesor!!!

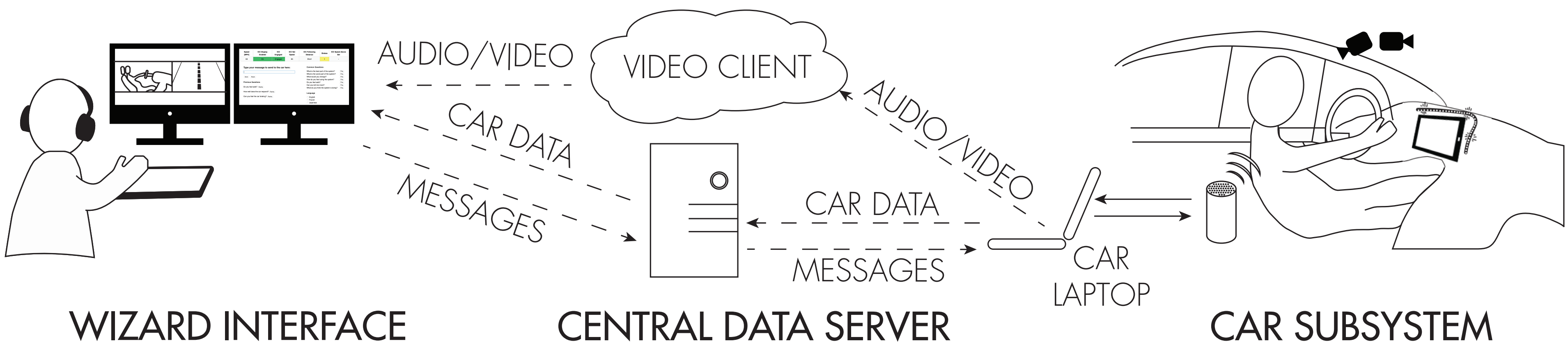


The Interaction Engine is a framework for prototyping web-connected hardware.

We use a set of widely supported tools to create a system to help interaction designers quickly realize new, multimodal interactive experiences.

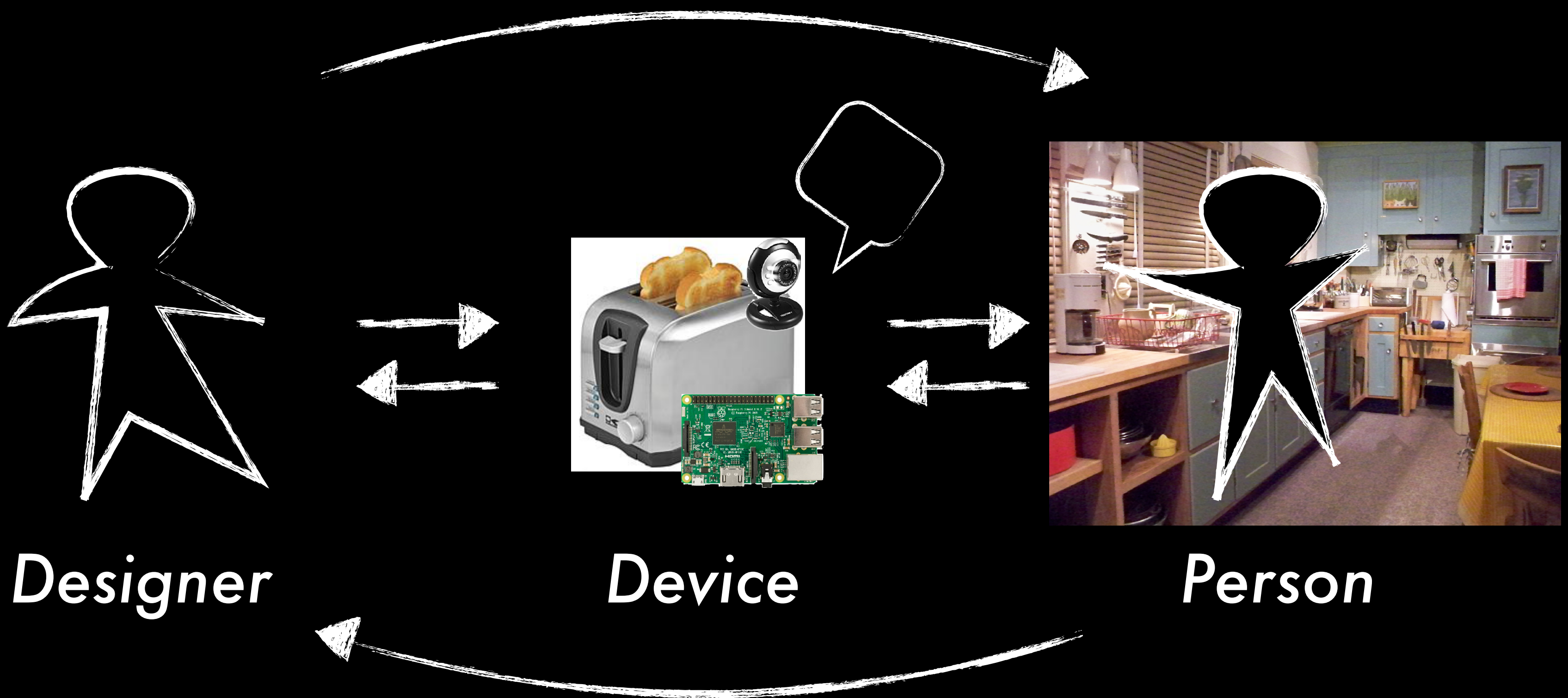
F R O M   N I K ' S   G U E S T   L E C T U R E :

Putting designers in-the-loop

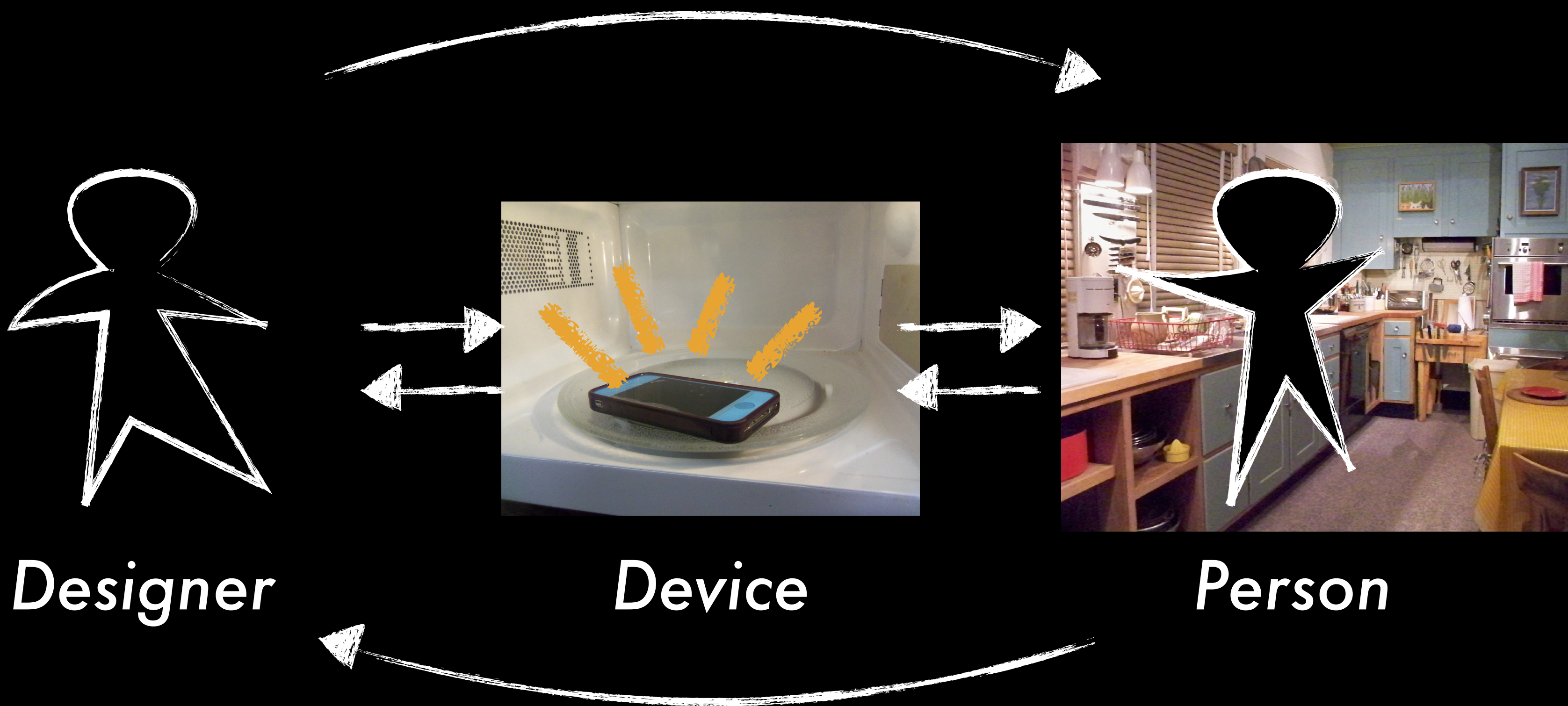


Nikolas Martelaro and Wendy Ju. 2017. WoZ Way: Enabling Real-time Remote Interaction Prototyping & Observation in On-road Vehicles. CSCW '17. ACM, New York, NY, USA, 169-182

F R O M   N I K ' S   G U E S T   L E C T U R E :



Tinkerbell



SPEECH SYNTHESIS

BELL LABS, HOMER DUDLEY, VODER (1939)

MonoThyratoron: https://www.youtube.com/watch?v=5hyI_dM5cGo

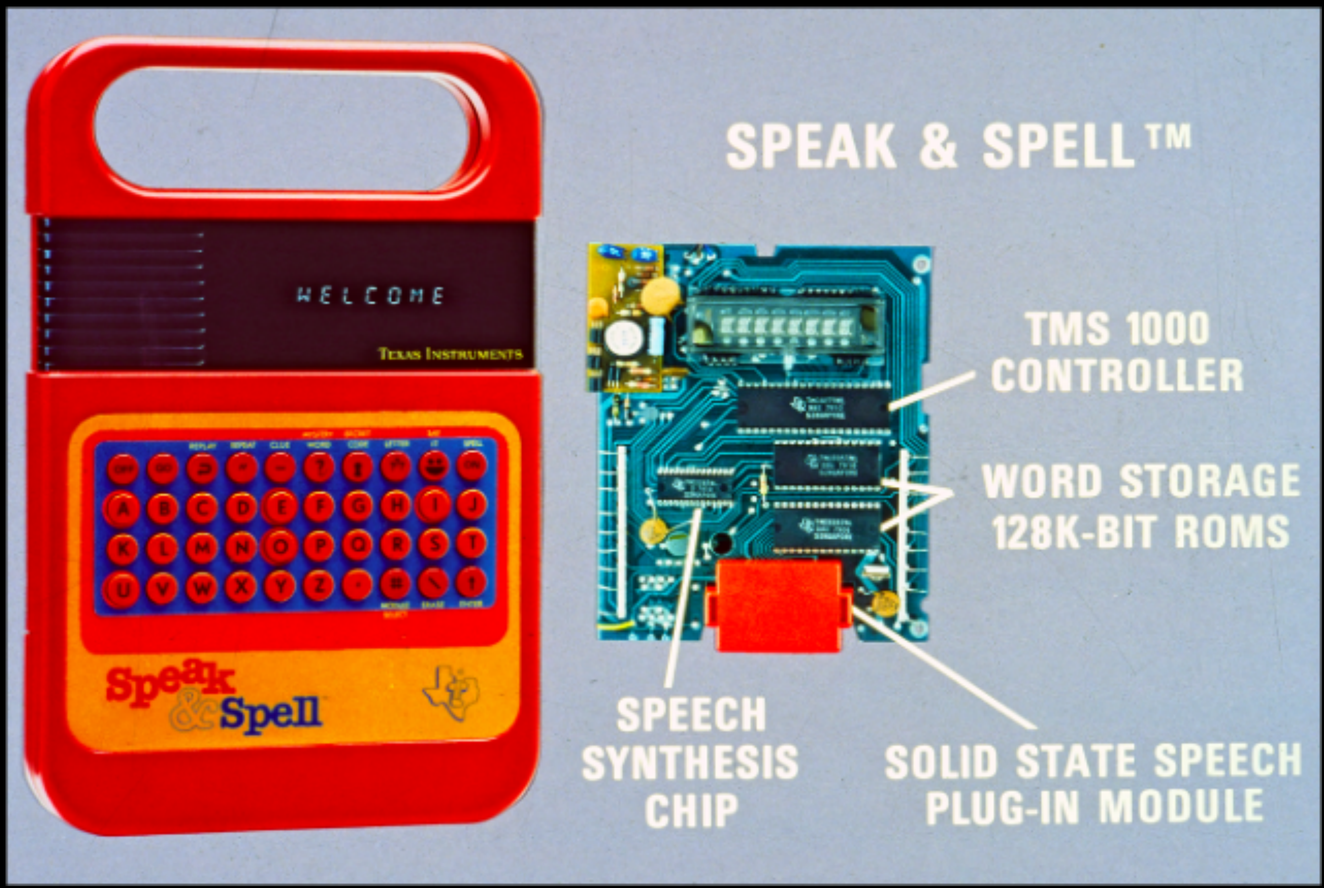
TEXAS INSTRUMENTS SPEAK & SPELL (1978)

Johnsebay: https://www.youtube.com/watch?v=UwWaeEyhPP0

TEXAS INSTRUMENTS SPEAK & SPELL (1978)

- Used the Texas Instruments TMC0280 linear predictive coding speech synthesizer IC (integrated circuit)
- First synthesized speech in educational toy (not taped or recorded)
- Phonemes for spoken words were stored on ROM (read only memory)

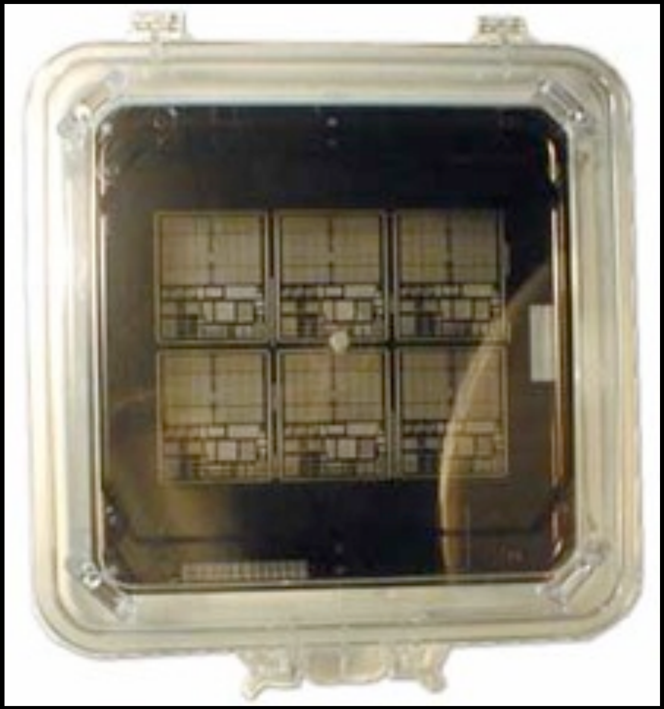
TEXAS INSTRUMENTS SPEAK & SPELL (1978)



http://hackeducation.com/2015/01/13/speak-and-spell

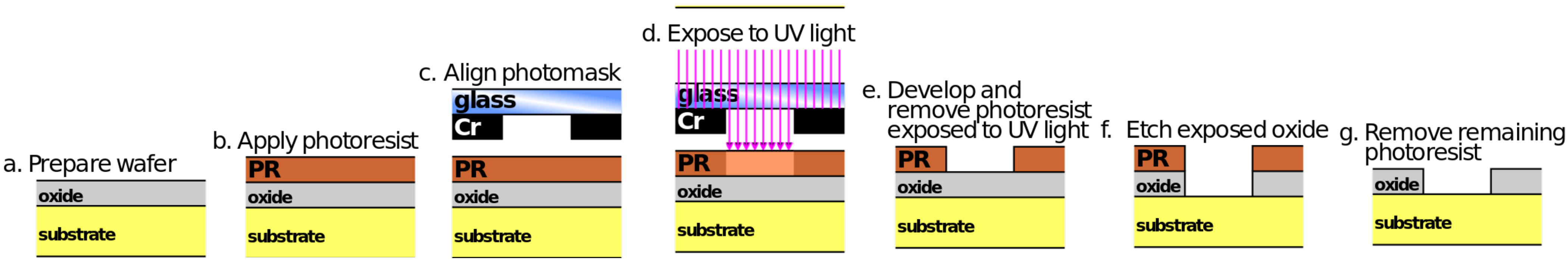
Integrated Circuits Manufacturing

Similar to Silkscreening



Integrated Circuits Manufacturing

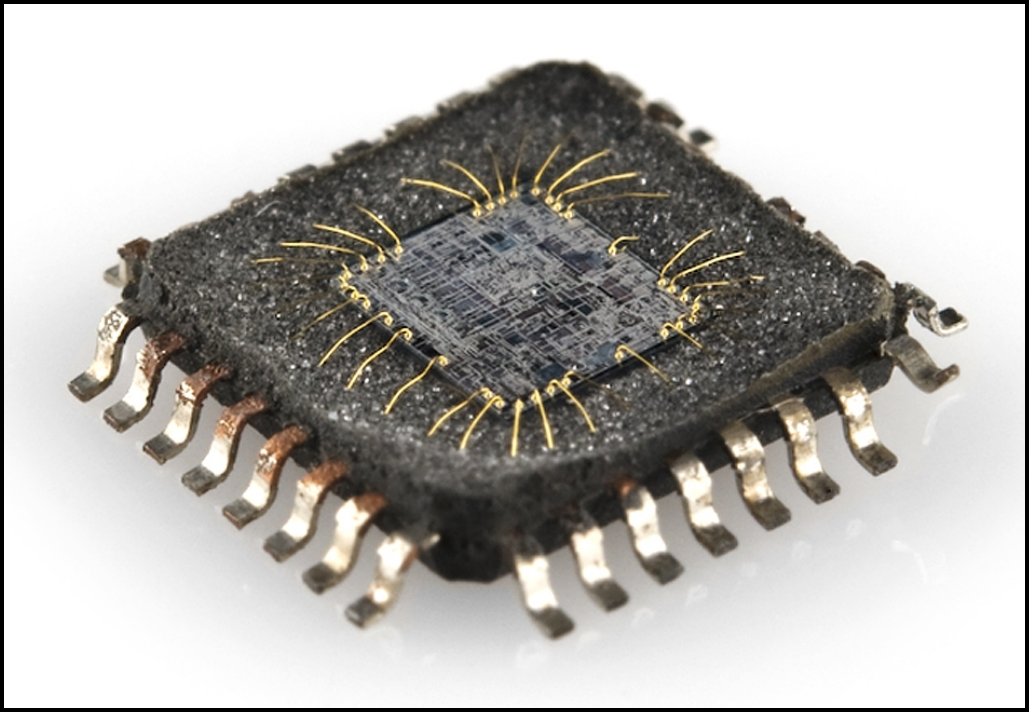
Photolithography



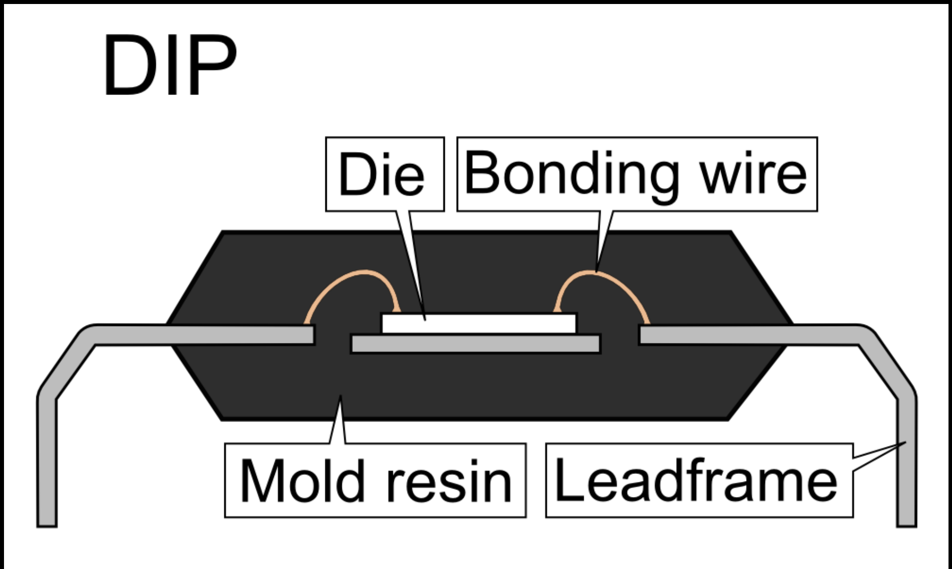
Photolithography
Wikipedia

Integrated Circuit Packaging

Where does the silicon go?



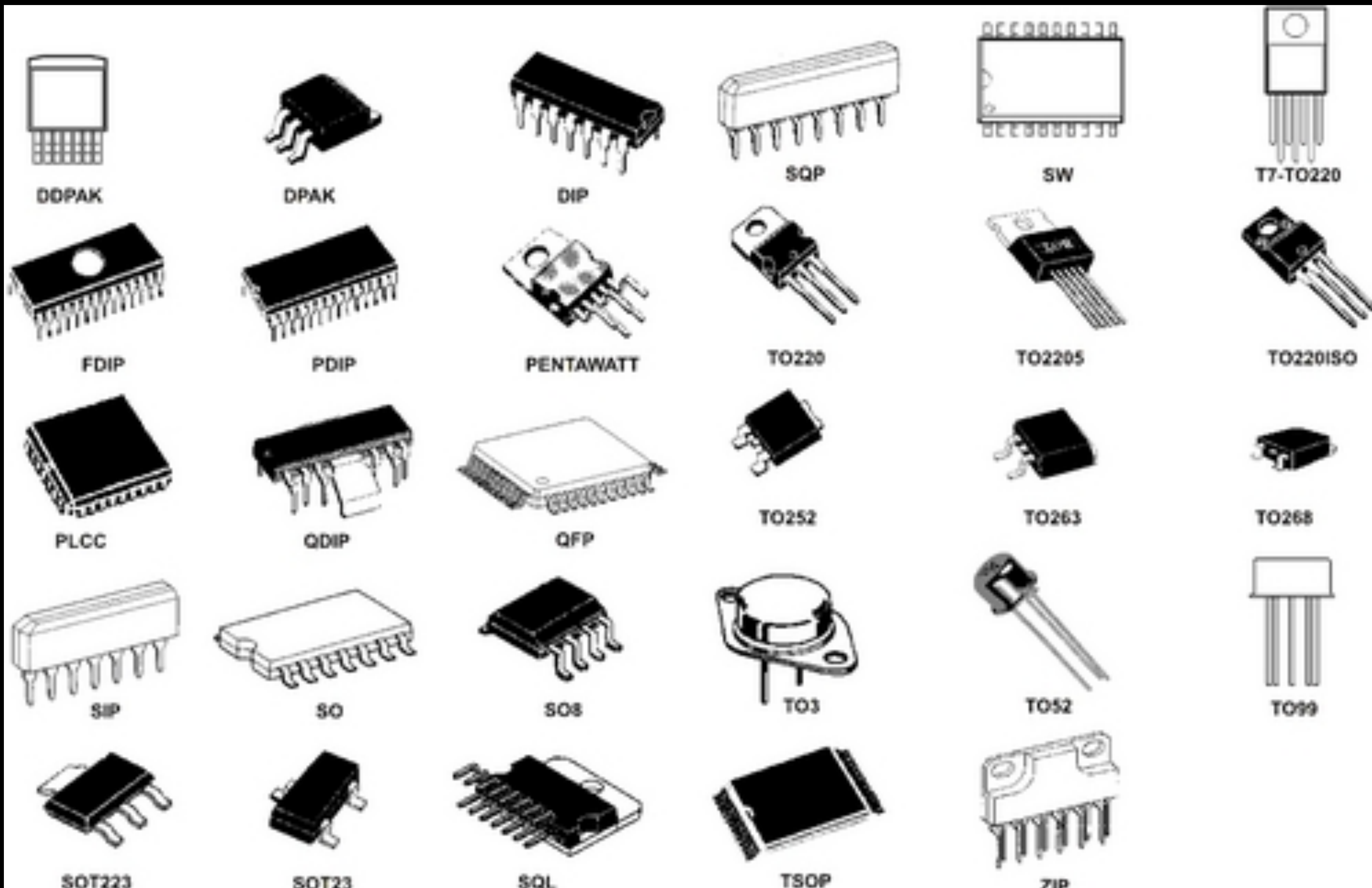
The guts of an integrated circuit
https://learn.sparkfun.com/tutorials/integrated-circuits/all



IC packaging, Wikipedia

Integrated Circuit Packaging

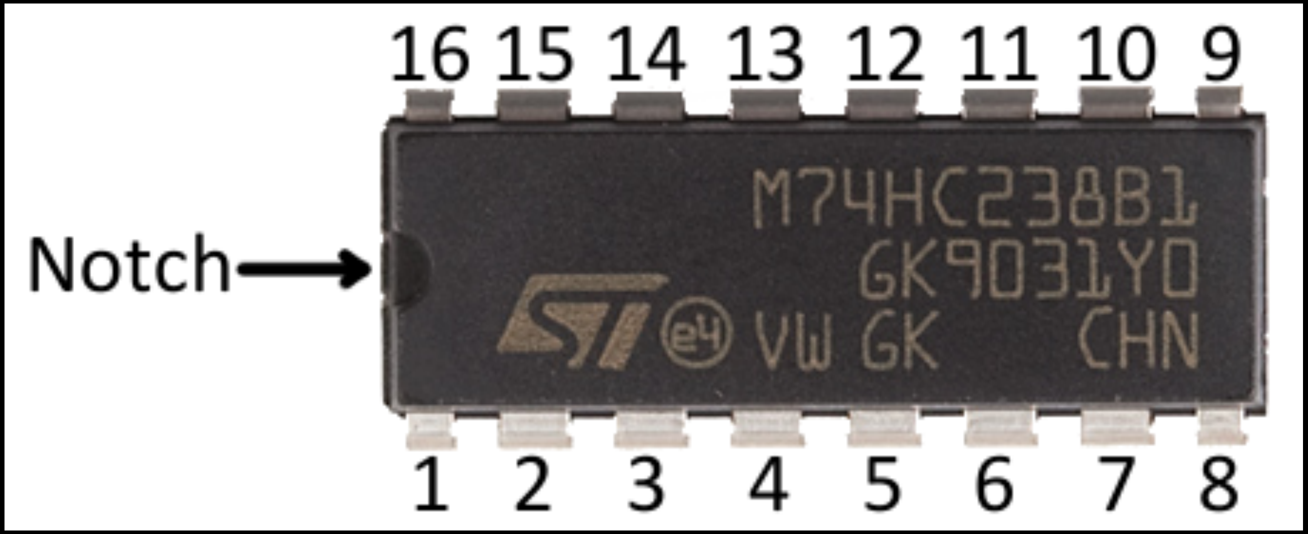
Where does the silicon go?



https://learn.sparkfun.com/tutorials/integrated-circuits/all

Integrated Circuit Packaging

Polarity, Pin numbering



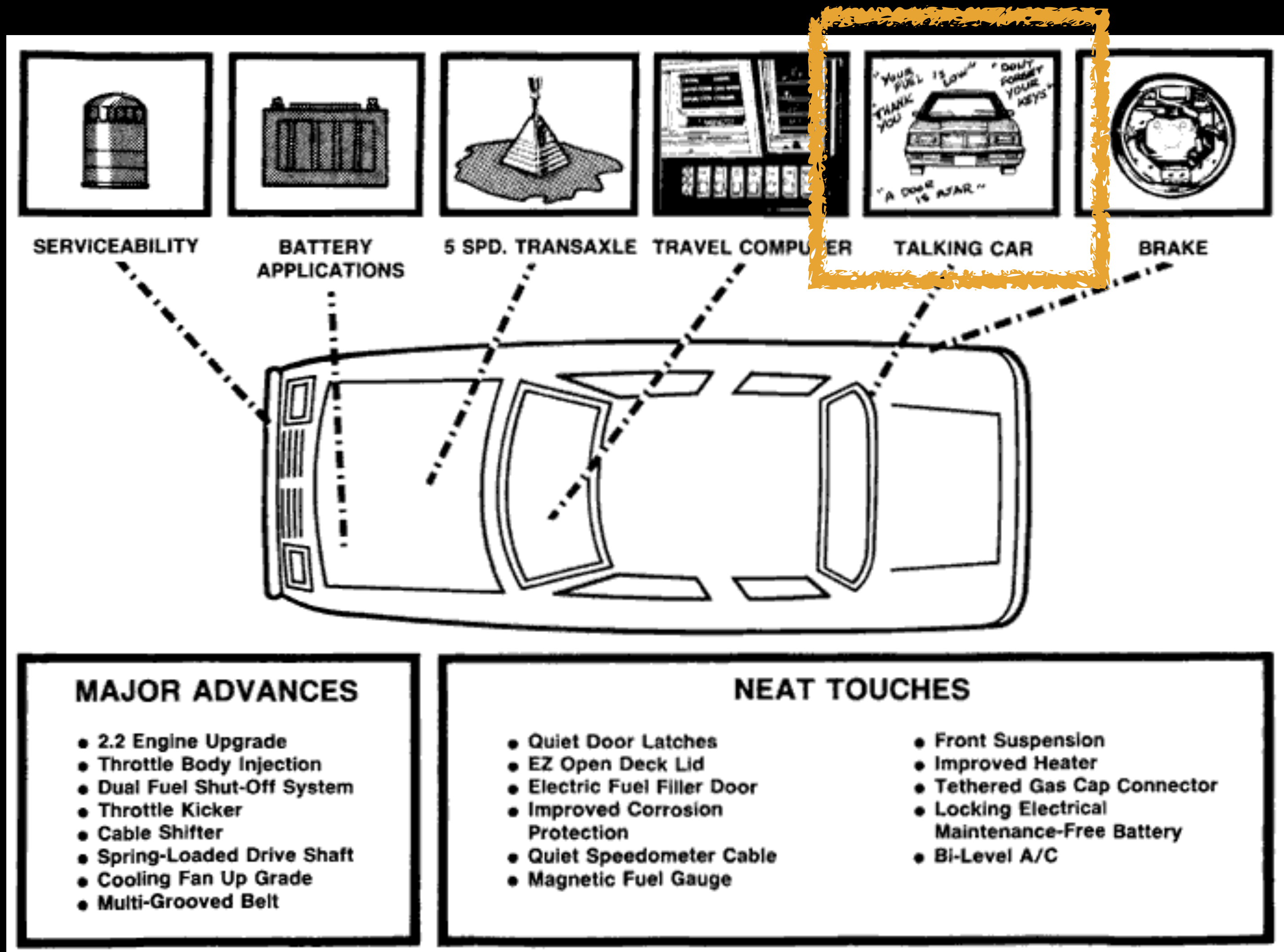
https://learn.sparkfun.com/tutorials/
integrated-circuits/all

CHRYSLER ELECTRONIC VOICE ALERT 1983-88

TBL: https://www.youtube.com/watch?v=0BFdvJwl8T8

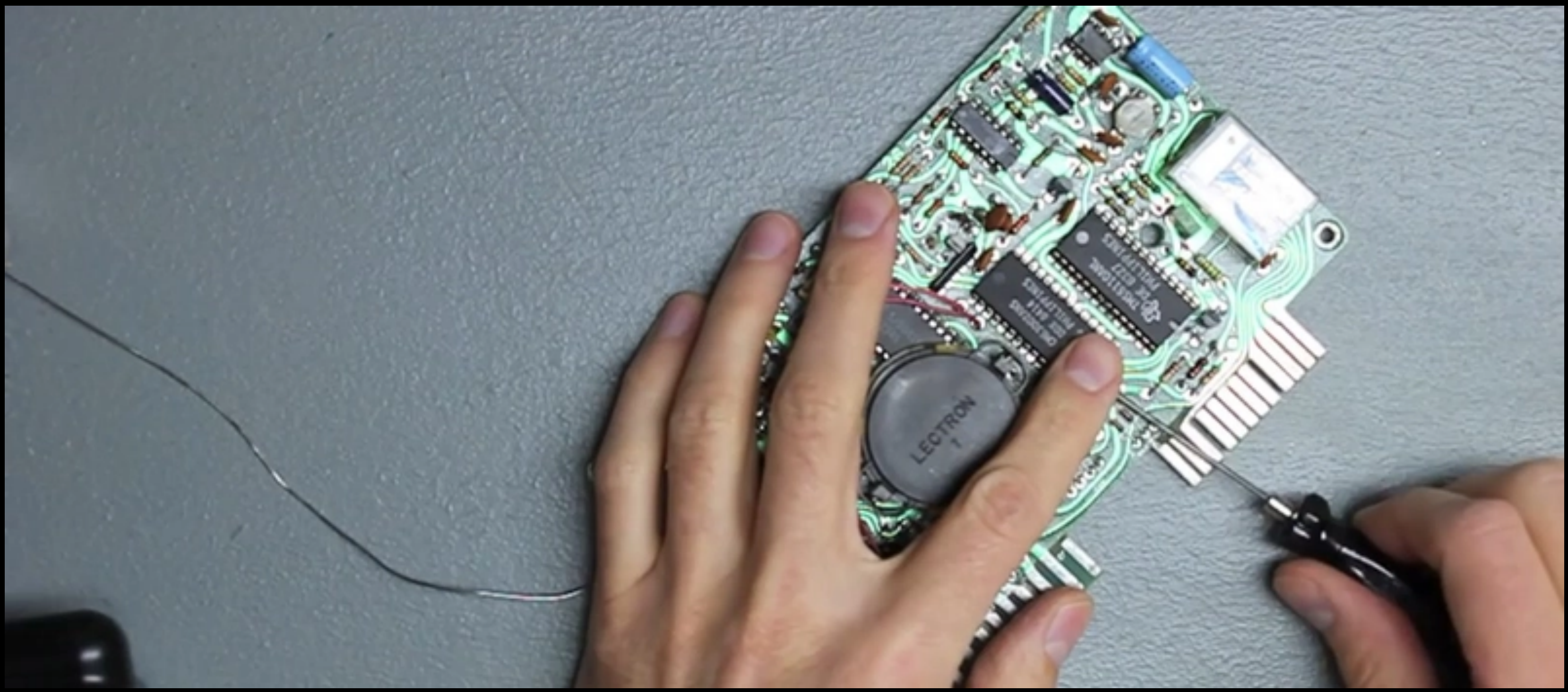
CHRYSLER ELECTRONIC VOICE ALERT 1983-88

- Also used the Texas Instruments LPC speech chips



CHRYSLER ELECTRONIC VOICE ALERT 1983-88

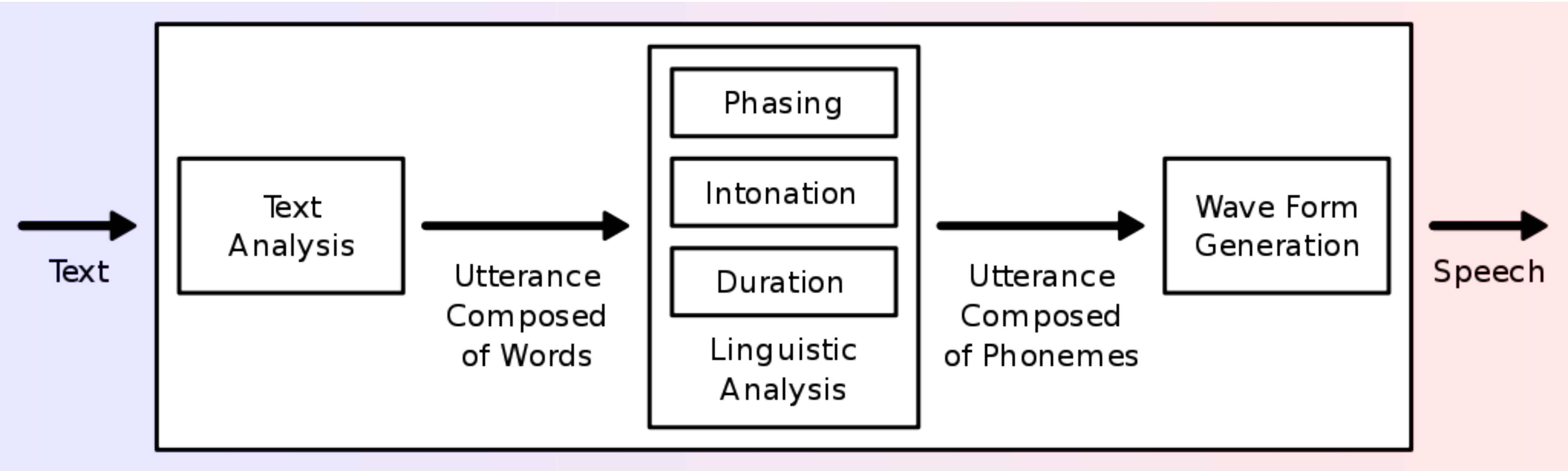
- Also used the Texas Instruments LPC speech chips



https://hackaday.com/2016/01/20/saving-old-voices-by-dumping-roms/

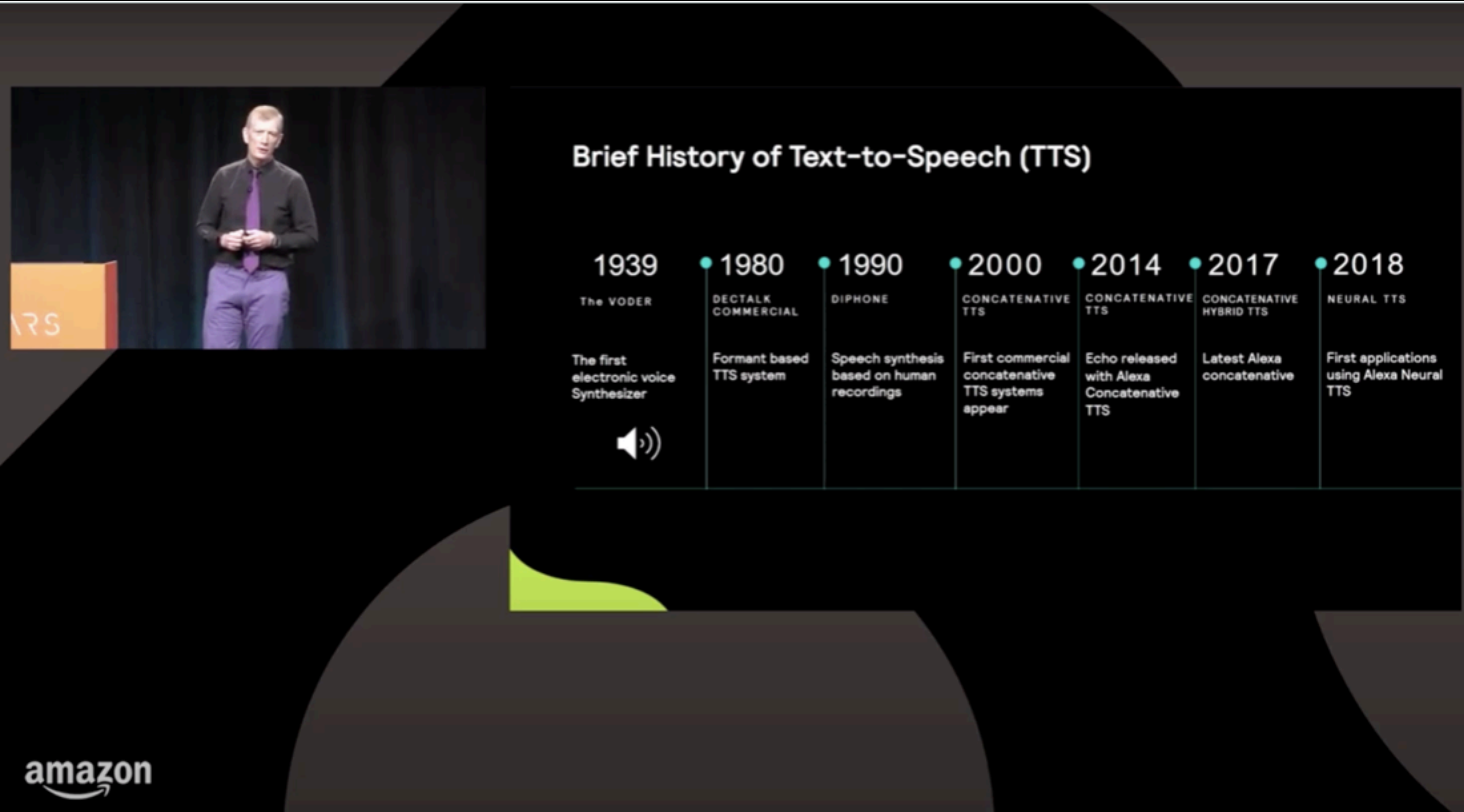
TEXT TO SPEECH

TYPICAL TTS DATA FLOW



https://en.wikipedia.org/wiki/Speech_synthesis

T T S   T I M E L I N E



https://www.amazon.science/latest-news/advances-in-text-to-speech-technologies-help-computers-find-their-voice

TTS ON RASPBERRY PI

- Cepstral TTS- works offline, voice quality higher quality than open-source solutions, they charge licensing fee for commercial applications https://www.cepstral.com/en/raspberrypi
- Festival TTS- works offline, Good robot-y voice
- Espeak- works offline, Alien-y voice
- Google TTS- sounds awesome, depends on Google servers
- Pico TTS: Android TTS Engine

SPEECH RECOGNITION

TIGER ELECTRONICS FURBY (1998)

https://www.instructables.com/Hack-Your-Furby-into-Zombie-Furby/

S P E E C H R E C O G N I T I O N

- Speech recognition is harder than speech synthesis because the speech is human initiated rather than machine initiated—the problem space can be un-bounded.

- Modern recognition on consumer-grade computers dates to the 1990s (Dragon Dictate, Voice Recognition Call Processing)

- Until recently, systems could recognize one person very well over a large vocabulary, or everyone well over a small vocabulary (speaker independence)

- Specialized speech corpuses can improve general untrained recognition in specific contexts

- Current technology works as well as human transcribers on high powered computers

SPEECH RECOGNITION ON PI

- Speech recognition on Pi works well over fixed vocabularies
- (Ilan to demo Vosk)

DATA IS IMPORTANT TO

INTERACTION DESIGN

Rich Sutton's The Bitter Lesson

Made into a slide deck,
and with post-reflections by Wendy Ju

David Goedicke and Wendy Ju

Based on

- http://www.incompleteideas.net/IncIdeas/BitterLesson.html

The Bitter Lesson

Rich Sutton

March 13, 2019

The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin. The ultimate reason for this is Moore's law, or rather its generalization of continued exponentially falling cost per unit of computation. Most AI research has been conducted as if the computation available to the agent were constant (in which case leveraging human knowledge would be one of the only ways to improve performance) but, over a slightly longer time than a typical research project, massively more computation inevitably becomes available. Seeking an improvement that makes a difference in the shorter term, researchers seek to leverage their human knowledge of the domain, but the only thing that matters in the long run is the leveraging of computation. These two need not run counter to each other, but in practice they tend to. Time spent on one is time not spent on the other. There are psychological commitments to investment in one approach or the other. And the human-knowledge approach tends to complicate methods in ways that make them less suited to taking advantage of general methods leveraging computation. There were many examples of AI researchers' belated learning of this bitter lesson, and it is instructive to review some of the most prominent.

In computer chess, the methods that defeated the world champion, Kasparov, in 1997, were based on massive, deep search. At the time, this was looked upon with dismay by the majority of computer-chess researchers who had pursued methods that leveraged human understanding of the special structure of chess. When a simpler, search-based approach with special hardware and software proved vastly more effective, these human-knowledge-based chess researchers were not good losers. They said that ``brute force" search may have won this time, but it was not a general strategy, and anyway it was not how people played chess. These researchers wanted methods based on human input to win and were disappointed when they did not.

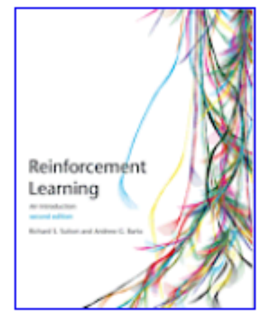
A similar pattern of research progress was seen in computer Go, only delayed by a further 20 years. Enormous initial efforts went into avoiding search by taking advantage of human knowledge, or of the special features of the game, but all those efforts proved irrelevant, or worse, once search was applied effectively at scale. Also important was the use of learning by self play to learn a value function (as it was in many other games and even in chess, although learning did not play a big role in the 1997 program that first beat a world champion). Learning by self play, and learning in general, is like search in that it enables massive computation to be brought to bear. Search and learning are the two most important classes of techniques for utilizing massive amounts of computation in AI research. In computer Go, as in computer chess, researchers' initial effort was directed towards utilizing human understanding (so that less search was needed) and only much later was much greater success had by embracing search and learning.

Richard S. Sutton

Distinguished Research Scientist
DeepMind

Professor and iCORE chair
Department of Computing Science
University of Alberta
2-21 Athabasca Hall
Edmonton, Alberta
Canada T6G 2E8

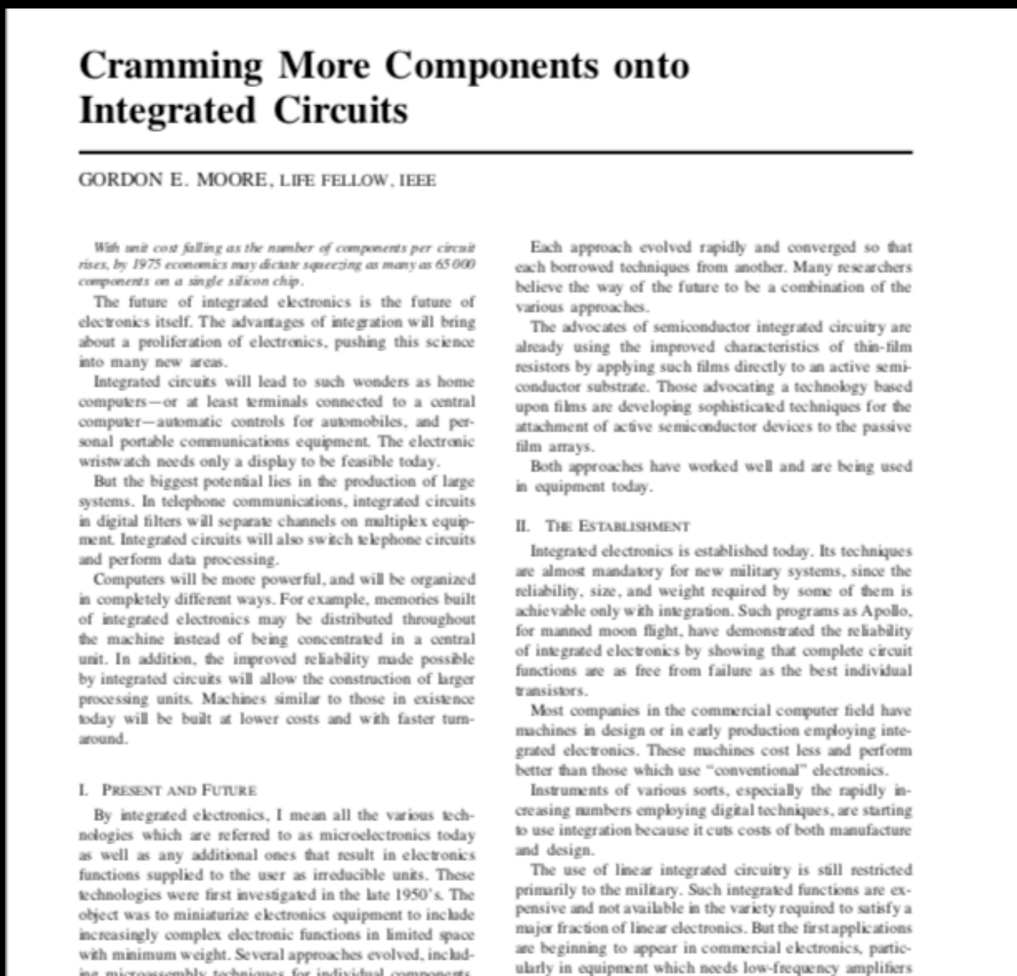
office 3-13 Athabasca Hall
email rsutton@ualberta.ca or rich@richsutton.com
fax 1-780-492-1111
web http://richsutton.com



- Research
  - Worldwide RLAI research at rlai.net
  - 2013 NSERC technical proposal
  - 2012 iCORE research proposal
- Information for prospective students
- Brief biography, CV
- Publications (Google Scholar)
- Talks, video of interview on AI and society, video of academic talk on *The Future of AI*
- Reinforcement Learning: An Introduction (textbook)
- RL FAQ - Frequently asked questions about reinforcement learning (from 2004)
- CMPUT 397 - Reinforcement Learning MOOC
- new CMPUT 609 - Reinforcement Learning II
- old CMPUT 609 - Reinforcement Learning I (Reinforcement Learning for Artificial Intelligence)
- CMPUT 325 - Non-procedural Programming Languages (Fall 2006)

The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin.

Moore’s Law, predicted



Cramming More Components onto Integrated Circuits

GORDON E. MOORE, LIFE FELLOW, IEEE

With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65 000 components on a single silicon chip.

The future of integrated electronics is the future of electronics itself. The advantages of integration will bring about a proliferation of electronics, pushing this science into many new areas.

Integrated circuits will lead to such wonders as home computers—or at least terminals connected to a central computer—automatic controls for automobiles, and personal portable communications equipment. The electronic wristwatch needs only a display to be feasible today.

But the biggest potential lies in the production of large systems. In telephone communications, integrated circuits in digital filters will separate channels on multiplex equipment. Integrated circuits will also switch telephone circuits and perform data processing.

Computers will be more powerful, and will be organized in completely different ways. For example, memories built of integrated electronics may be distributed throughout the machine instead of being concentrated in a central unit. In addition, the improved reliability made possible by integrated circuits will allow the construction of larger processing units. Machines similar to those in existence today will be built at lower costs and with faster turnaround.

I. Present and Future

By integrated electronics, I mean all the various technologies which are referred to as microelectronics today as well as any additional ones that result in electronics functions supplied to the user as irreducible units. These technologies were first investigated in the late 1950's. The object was to miniaturize electronics equipment to include increasingly complex electronic functions in limited space with minimum weight. Several approaches evolved, including microassembly techniques for individual components,

Each approach evolved rapidly and converged so that each borrowed techniques from another. Many researchers believe the way of the future to be a combination of the various approaches.

The advocates of semiconductor integrated circuitry are already using the improved characteristics of thin-film resistors by applying such films directly to an active semiconductor substrate. Those advocating a technology based upon films are developing sophisticated techniques for the attachment of active semiconductor devices to the passive film arrays.

Both approaches have worked well and are being used in equipment today.

II. The Establishment

Integrated electronics is established today. Its techniques are almost mandatory for new military systems, since the reliability, size, and weight required by some of them is achievable only with integration. Such programs as Apollo, for manned moon flight, have demonstrated the reliability of integrated electronics by showing that complete circuit functions are as free from failure as the best individual transistors.

Most companies in the commercial computer field have machines in design or in early production employing integrated electronics. These machines cost less and perform better than those which use "conventional" electronics.

Instruments of various sorts, especially the rapidly increasing numbers employing digital techniques, are starting to use integration because it cuts costs of both manufacture and design.

The use of linear integrated circuitry is still restricted primarily to the military. Such integrated functions are expensive and not available in the variety required to satisfy a major fraction of linear electronics. But the first applications are beginning to appear in commercial electronics, particularly in equipment which needs low-frequency amplifiers

Fig. 2.

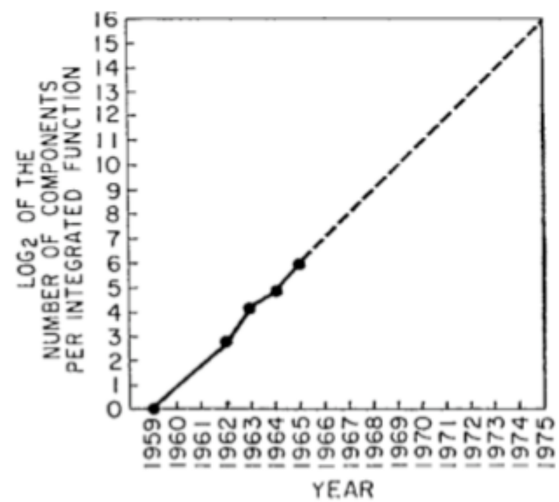
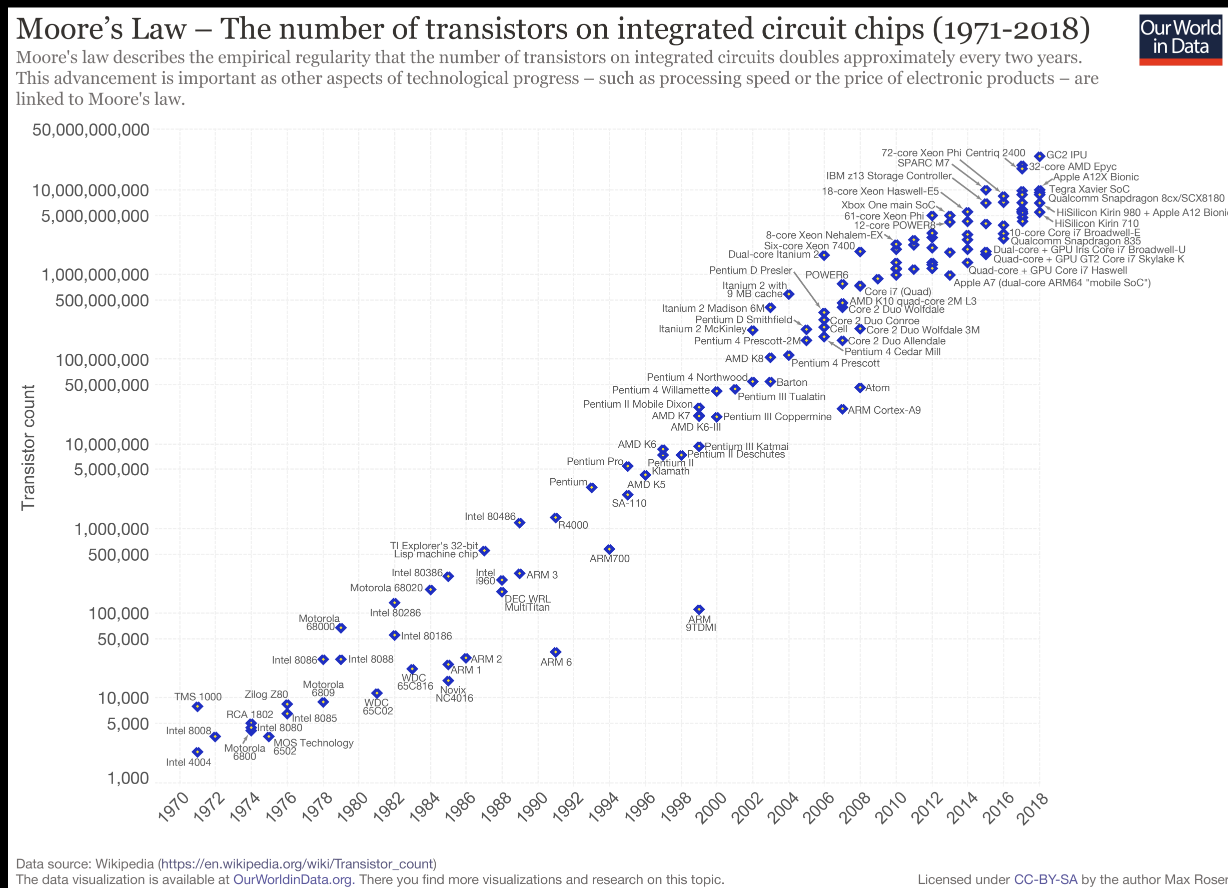


Fig. 3.

Moore, Gordon E. (1965). "Cramming more components onto integrated circuits" (PDF). Electronics Magazine. p. 4. Retrieved 2006-11-11.

Moore’s Law, Actualized



https://en.wikipedia.org/wiki/Moore%27s_

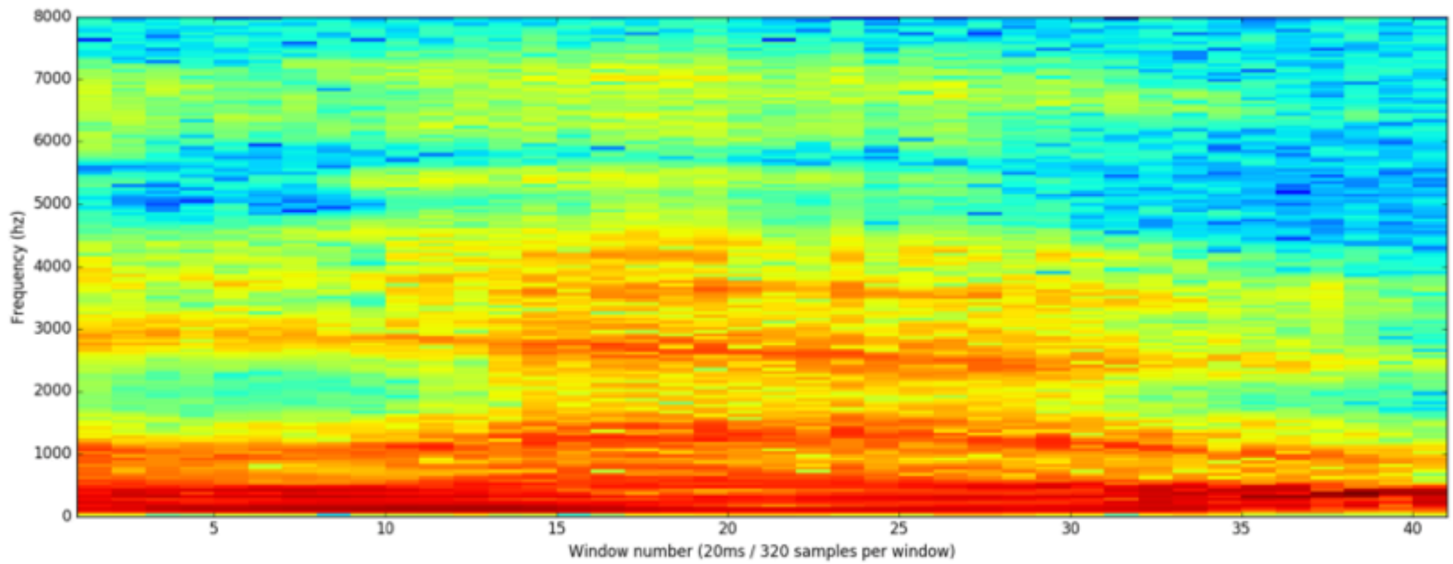
Computer Chess

https://rarehistoricalphotos.com/kasparov-deep-blue-

Alpha Go

https://www.netflix.com/title/8019084

Speech Recognition



The full spectrogram of the "hello" sound clip

https://medium.com/@ageitgey/machine-learning-is-fun-part-6-how-to-do-speech-recognition-with-deep-learning-28293c162f7a

Computer Vision

Cats vs Dogs

https://towardsdatascience.com/image-classifier-cats-vs-dogs-with-convolutional-neural-networks-cnns-and-google-colabs-4e9af21ae7a8

The bitter lesson is that:

1. AI researchers have often tried to build knowledge into their agents,
2. this always helps in the short term, and is personally satisfying to the researcher, but
3. in the long run it plateaus and even inhibits further progress, and
4. breakthrough progress eventually arrives by an opposing approach based on scaling computation by search and learning.

One thing that should be learned from the bitter lesson is the **great power of general purpose methods**, of methods that continue to scale with increased computation even as the available computation becomes very great. **The two methods that seem to scale arbitrarily in this way are *search* and *learning***.

Wendy's thoughts

One thing that is overlooked by Sutton's essay is that the key knowledge is captured by the data, and the people evaluating the success of the algorithms.

The work that goes into labelling data, or forming clusters, is in fact a human-intelligence task.

The evaluation of people is another and very important binary classification that determines the stopping function.

Stephen Yang, Brian Mok, David Sirkin, Hillary Page Ive, Rohan Maheshwari, Kerstin Fischer, Wendy Ju. Experiences Developing Socially Acceptable Interactions for a Robotic Trash Barrel. In RO-MAN 2015. August 31-September 3, 2015. Kobe, Japan.

Is Now a Good Time?

In-cabin + road video

Elicited Response

Biosignals

Automotive Data

Driving Route

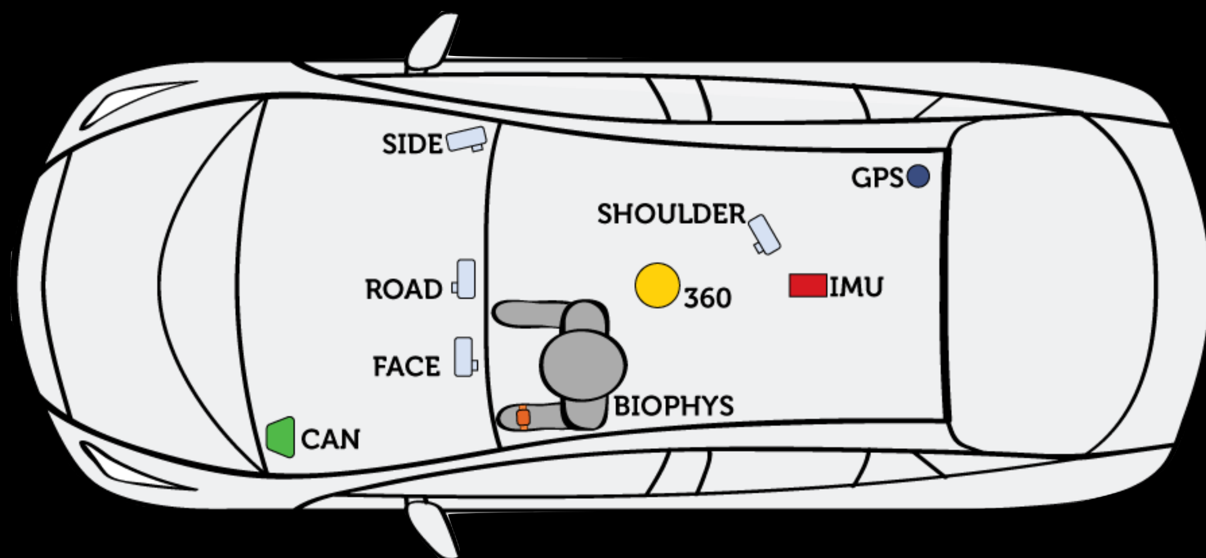
Dataset development with TRI

63 drivers

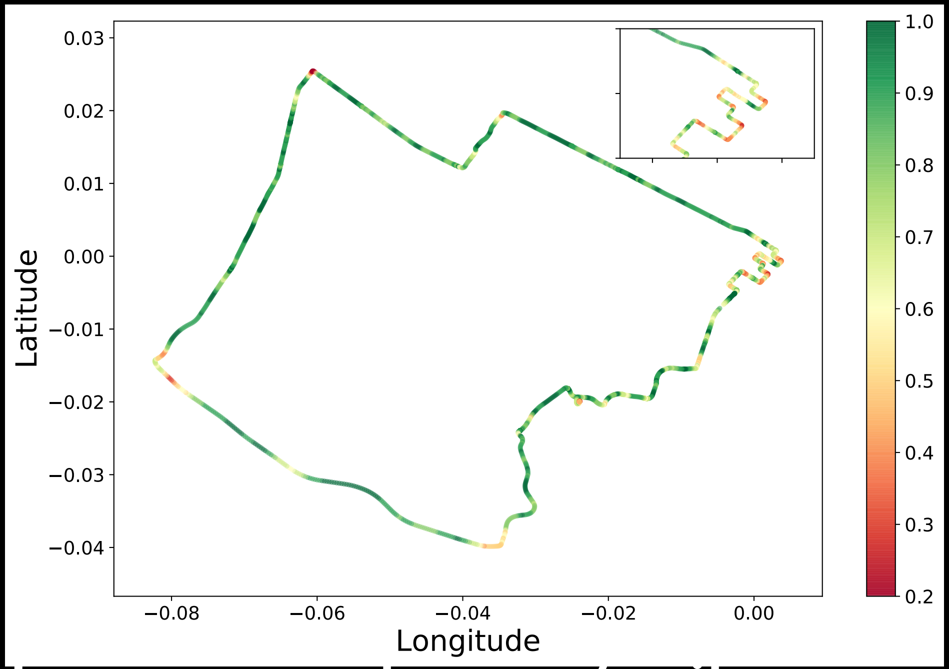
59 hours of video

2734 responses

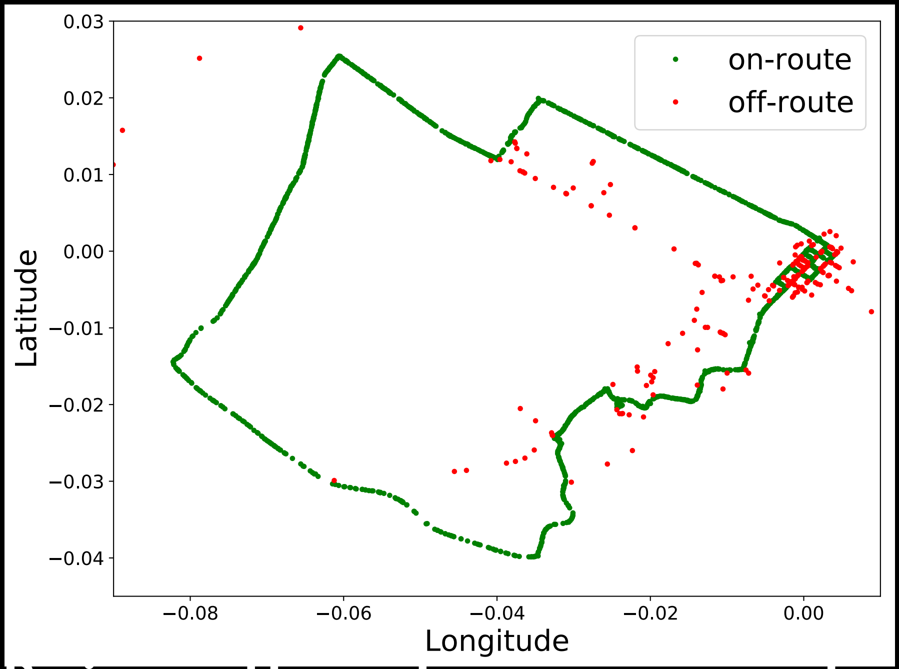
Overall response rate:
78% yes
20% no
2% no answer



Challenging Moments



Freeway entrance/exit, challenging turns, urban driving



Being off route was one of the stronger predictors of NO responses

No, Now is Not a Good Time

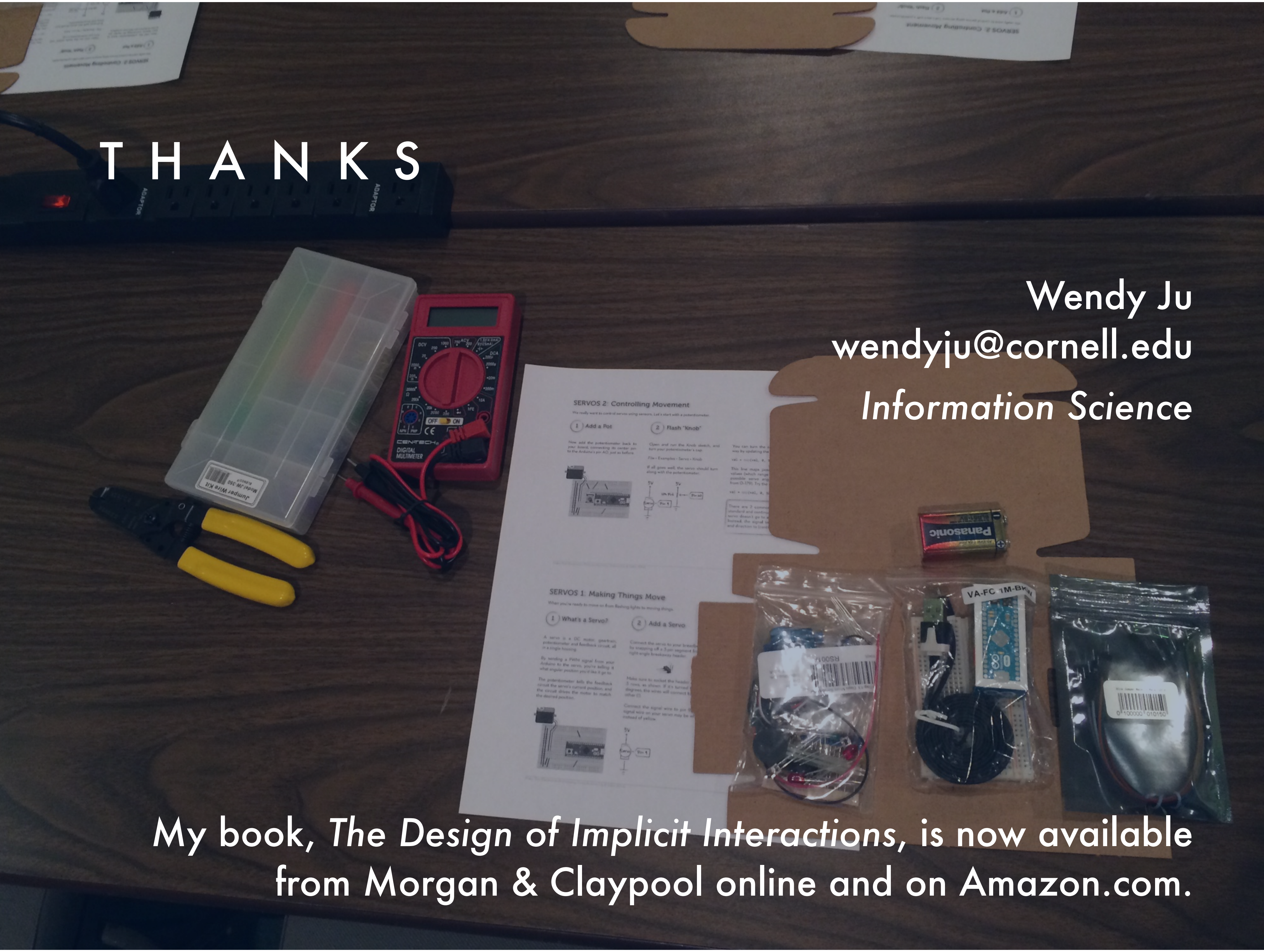
DATASETS FOR INTERACTION

LAB 3: CHATTERBOXES

Part 1 will be up on Tuesday afternoon: You will be trying out TTS and STT code.

No deliverables for 3/15 aside from setting up Github for the lab.

You will make some talking device, people’s responses, and think about what kind of data you would gather by Wizarding.



T H A N K S

Wendy Ju
wendyju@cornell.edu
Information Science

My book, *The Design of Implicit Interactions,* is now available from Morgan & Claypool online and on Amazon.com.