

ThoughtWorks®

# BEHIND CONVERSIONAL UI

---

*Xiaolei Wang*

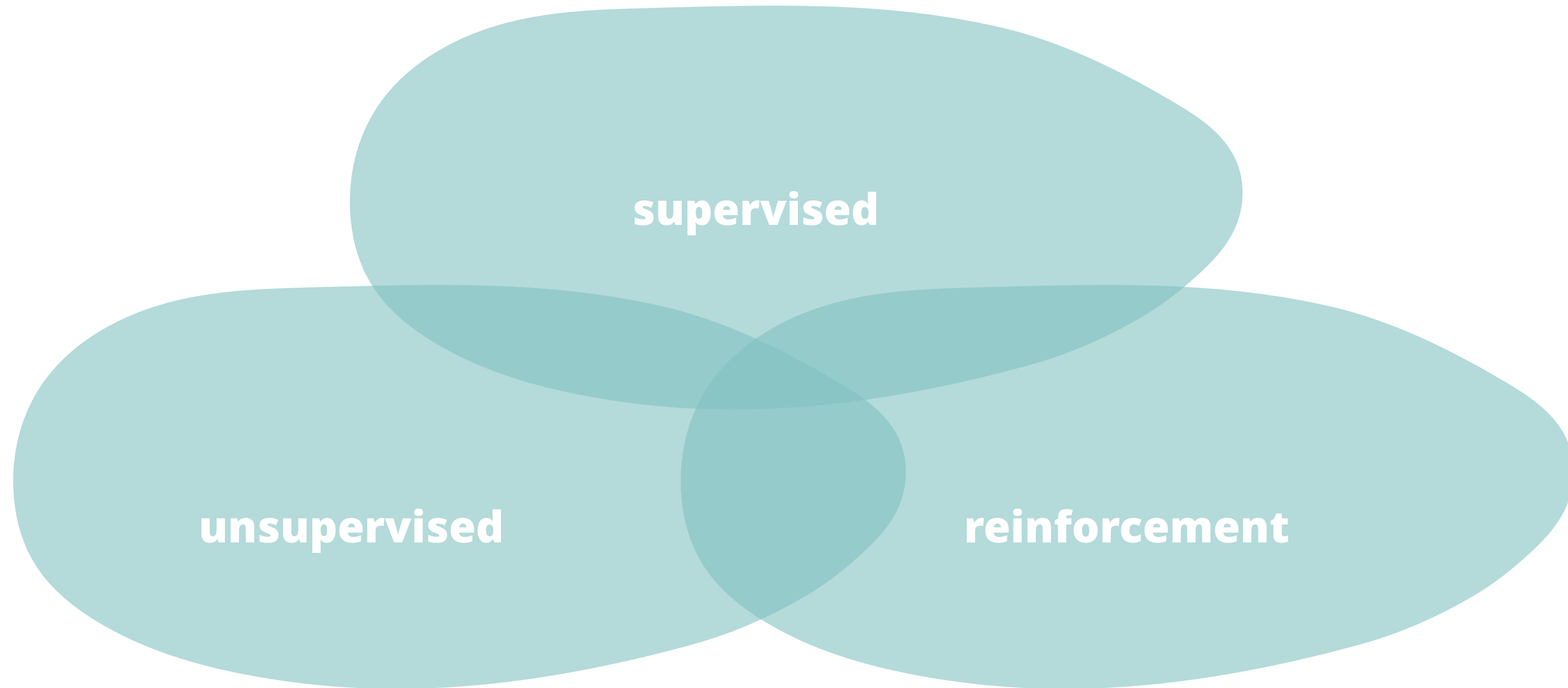
***“AI is like sex: sure, it may give  
some practical results, but that's  
not why we do it.”***

***——Artificial Feynman***

---

# MACHINE LEARNING

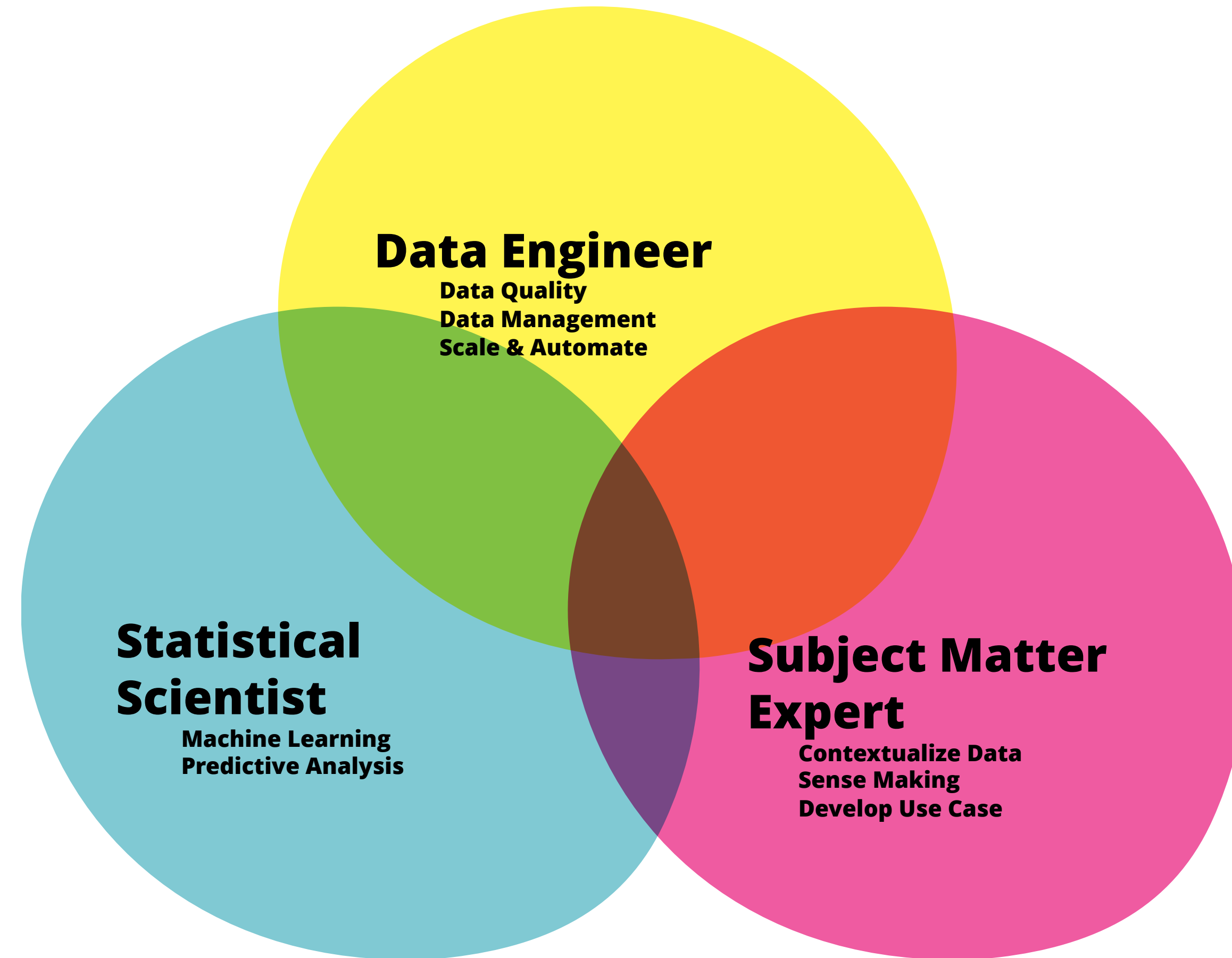
---



---

# ROLES

---



---

# IDENTIFICATION / RECOGNITION / CONVERSATION

---

*You speak.  
AI listens.  
We make it.*

翻译

记者

助理

销售

客服

**better knowledge, more intelligent**

**better service, more benefit**

---

# HEY THERE

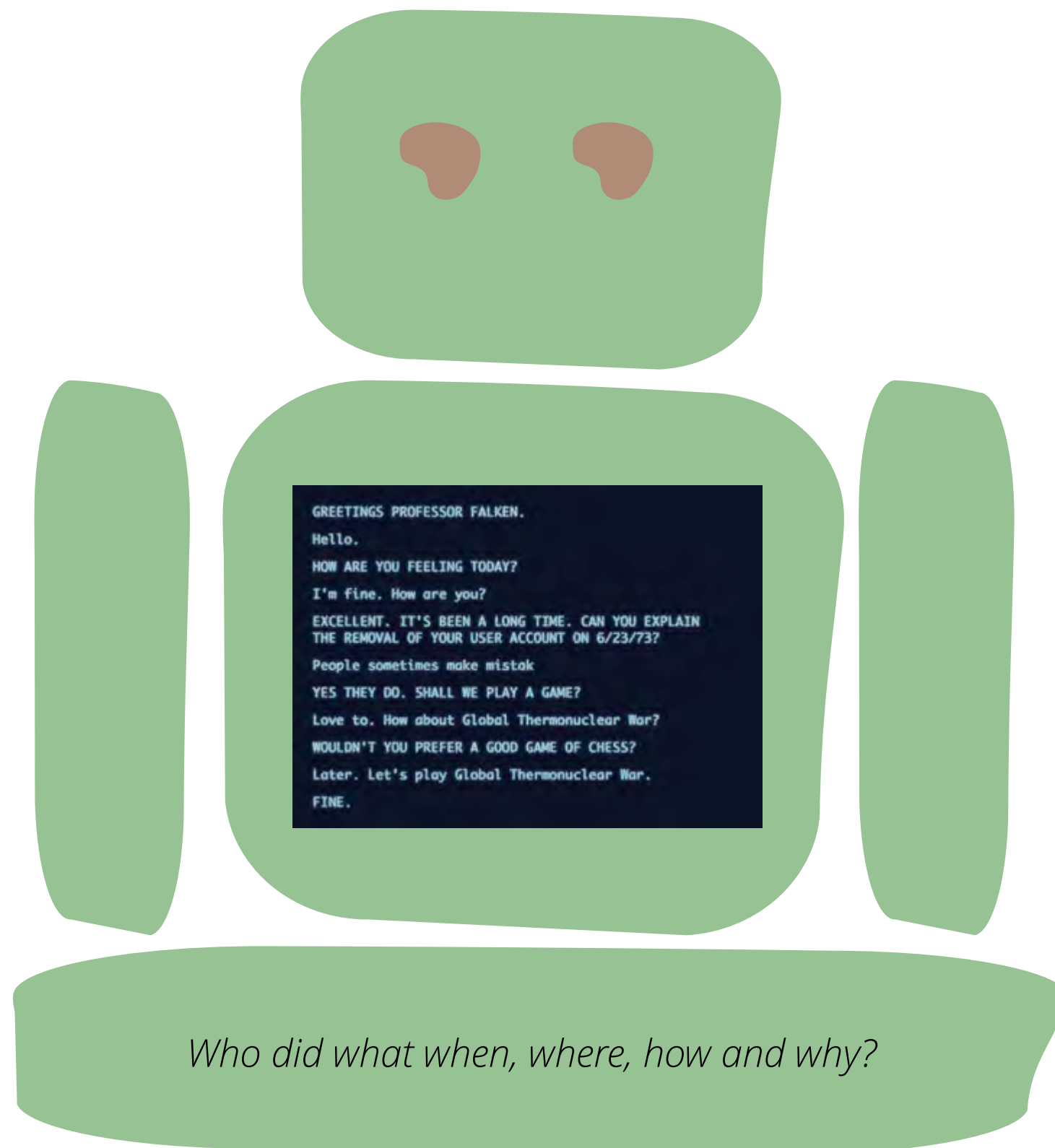
---

**Understand**

**Reason**

**Learn**

**Interact**



***'Hello World'***

---

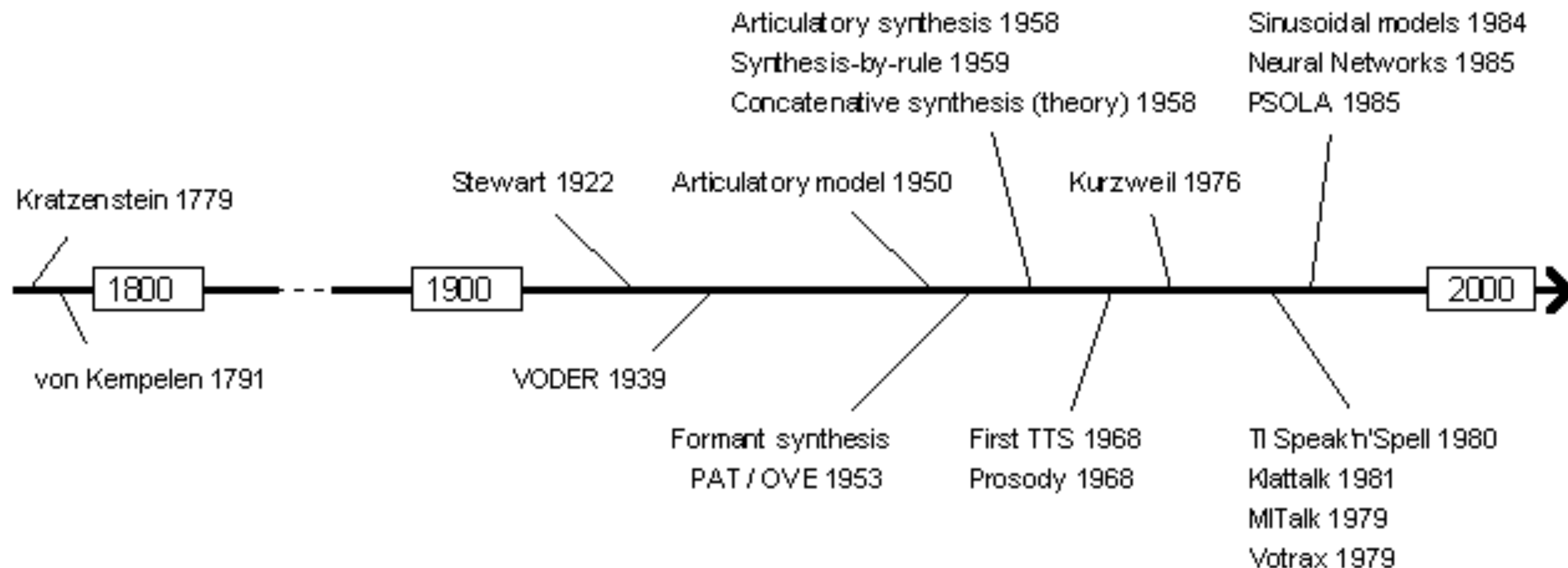
# DEMO: VON KEMPELEN'S SPEAKING MACHINE

---



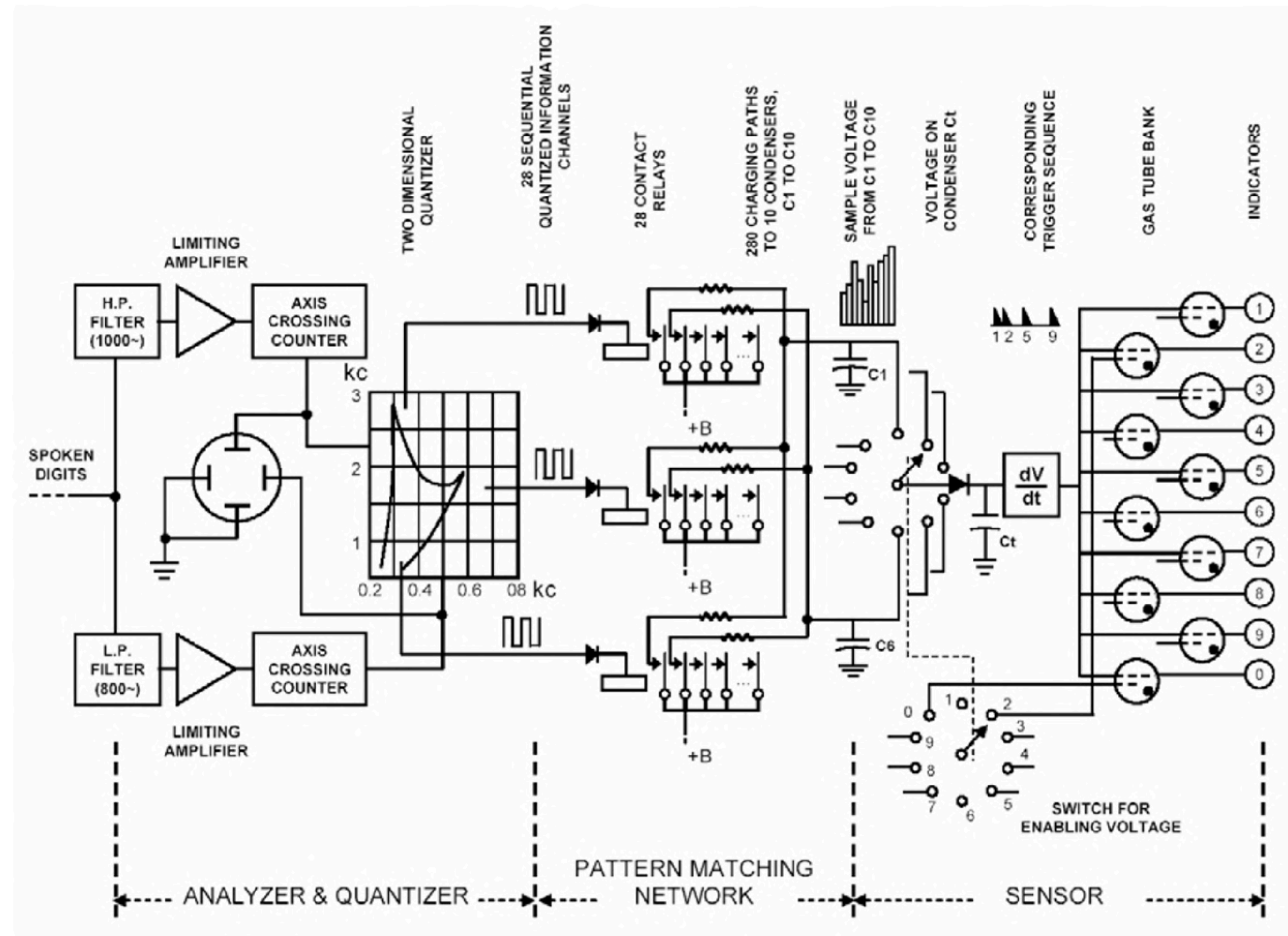


# MILESTONES



***To understand ...***

# EARLY AUTOMATIC SPEECH RECOGNIZERS

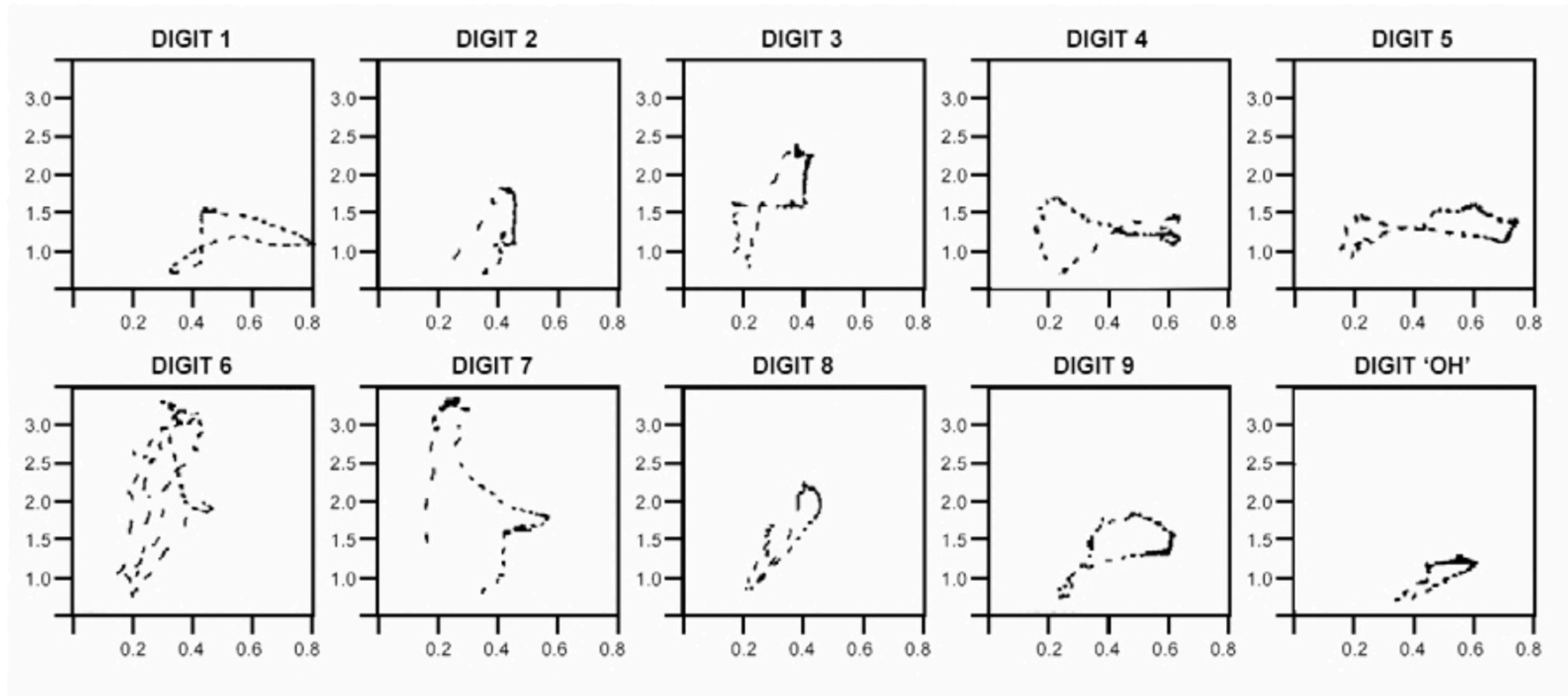


K. H. Davis, R. Biddulph, and S. Balashek, Automatic Recognition of Spoken Digits, J. Acoust. Soc. Am., Vol 24, No. 6, pp. 627-642, 1952.

---

# (CONT') EARLY AUTOMATIC SPEECH RECOGNIZERS

---

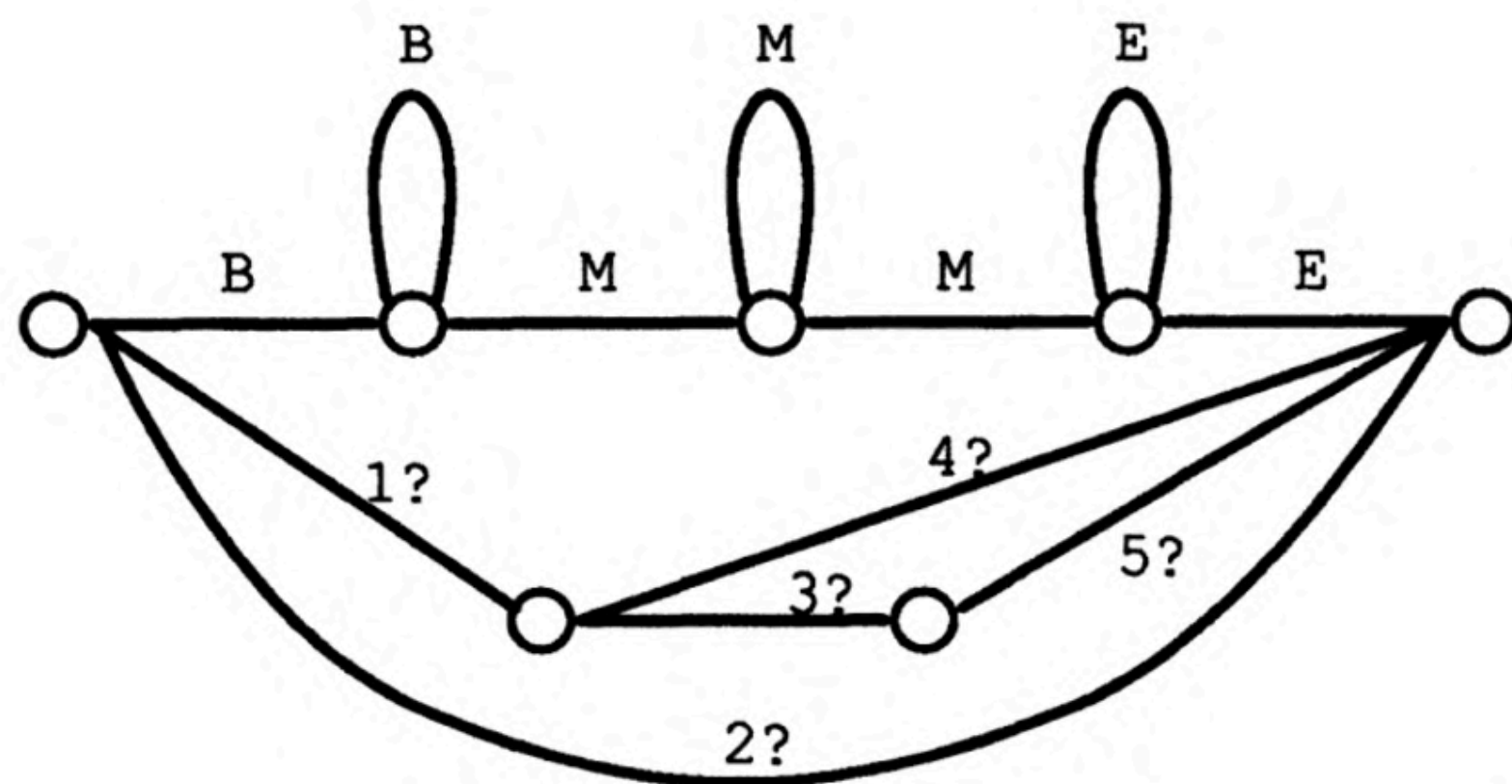


These natural modes of resonance, called the formants or formant frequencies, are manifested as major regions of energy concentration in the speech power spectrum.

**@1952**



# A COMPOSITE FINITE-STATE NETWORK



**Figure 5-7:** The HMM topology used in SPHINX, with different output pdf labelings on the lower transitions for different phones.

1	2	3	4	5	Phones
B	B	B	B	E	/ae/, /eh/, /ah/, /aa/, /ao/, /uw/, /aw/, /ay/, /ey/, /ow/, /oy/, /l/, /en/, /er/, /m/, /n/, /ng/, /f/, /sh/, /v/, /z/, /sil/
B	B	M	E	E	/ih/, /iy/, /uh/, /ax/, /ix/, /r/, /w/, /y/, /ch/, /jh/, /dx/
E	E	E	E	E	/b/, /d/, /dh/, /g/, /k/, /p/, /t/, /s/, /th/, /hh/, /ts/

**Table 5-2:** Lower transition labels assigned for each phone using the HMM in Figure 5-7.

Lee, Kai-Fu. *Automatic speech recognition: the development of the SPHINX system*. Vol. 62. Springer Science & Business Media, 1988.

@1988

***To learn ...***

# N-GRAM MODELS

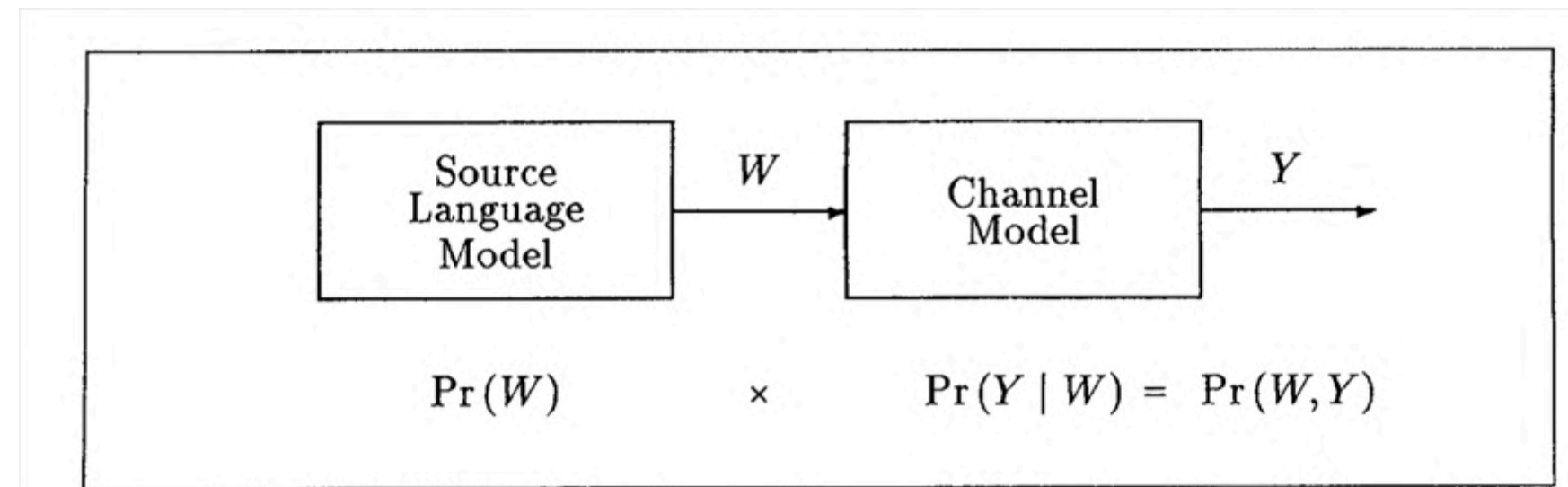


Figure 1  
Source-channel setup.

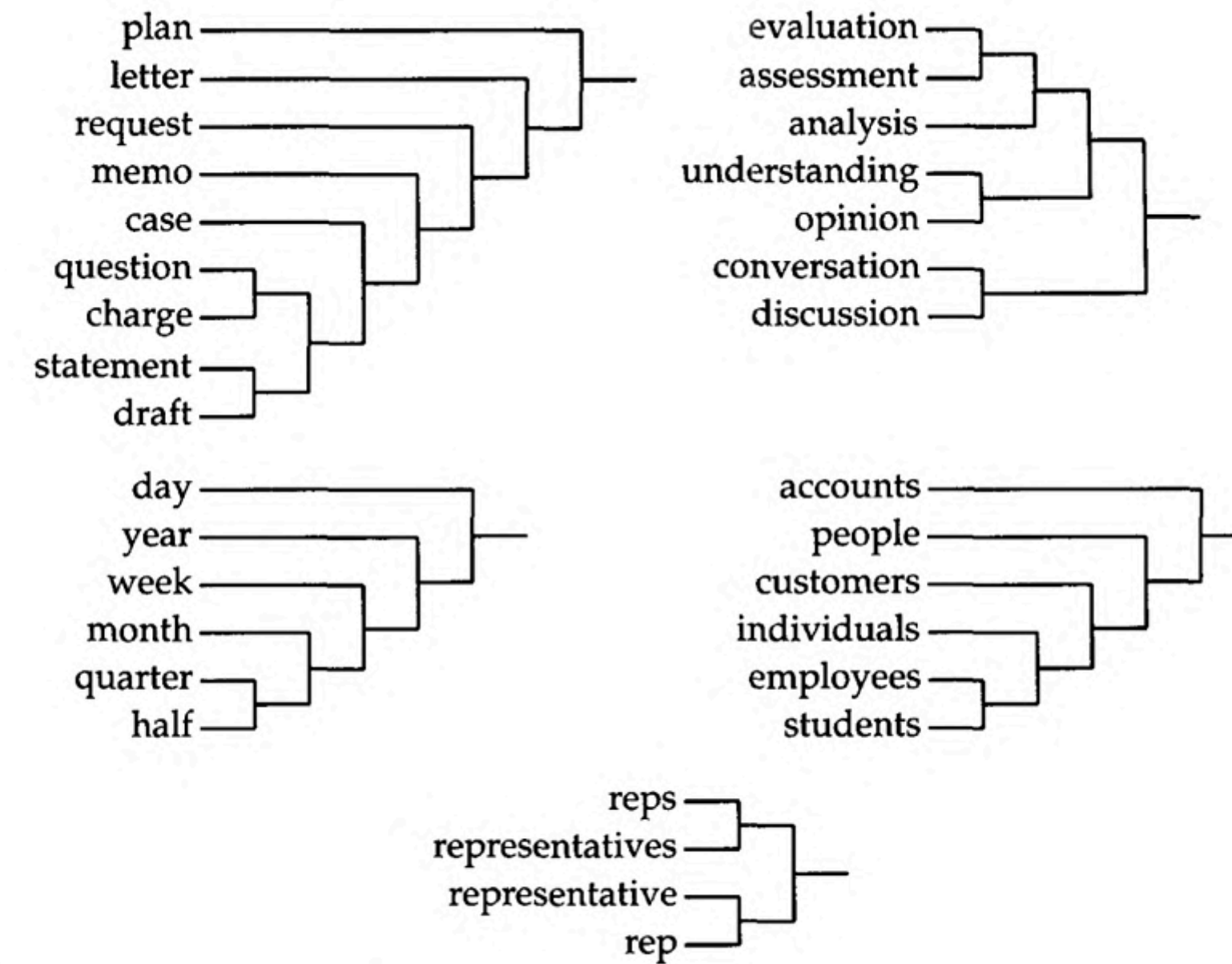


Figure 2  
Sample subtrees from a 1,000-word mutual information tree.

Brown, Peter F., et al. "Class-based n-gram models of natural language." *Computational linguistics* 18.4 (1992): 467-479.

@1992



# (CONT') N-GRAM MODELS

we our us ourselves ours  
 question questions asking answer answers answering  
 performance performed perform performs performing  
 tie jacket suit  
 write writes writing written wrote pen  
 morning noon evening night nights midnight bed  
 attorney counsel trial court judge  
 problems problem solution solve analyzed solved solving  
 letter addressed enclosed letters correspondence  
 large size small larger smaller  
 operations operations operating operate operated  
 school classroom teaching grade math  
 street block avenue corner blocks  
 table tables dining chairs plate  
 published publication author publish writer titled  
 wall ceiling walls enclosure roof  
 sell buy selling buying sold

**Table 6**  
Semantic clusters.

Word pair	Mutual Information
Humpty Dumpty	22.5
Klux Klan	22.2
Ku Klux	22.2
Chah Nulth	22.2
Lao Bao	22.2
Nuu Chah	22.1
Tse Tung	22.1
avant garde	22.1
Carena Bancorp	22.0
gizzard shad	22.0
Bobby Orr	22.0
Warnock Hersey	22.0
mutatis mutandis	21.9
Taj Mahal	21.8
Pontius Pilate	21.7
ammonium nitrate	21.7
jiggery pokery	21.6
Pitney Bowes	21.6
Lubor Zink	21.5
anciens combattants	21.5
Abu Dhabi	21.4
Aldo Moro	21.4
fuddle duddle	21.4
helter skelter	21.4
mumbo jumbo	21.4

@1992



---

# JUPLTER

---

JUPITER:	Hi! Welcome to JUPITER, the MIT Lab for Computer Science weather information system. This call is being recorded for system development. You may hang up or ask for help at any time. How can I help you?
User:	CAN YOU TELL ME WHAT'S THE WEATHER LIKE IN BOSTON TODAY
JUPITER:	In Boston this afternoon. Partly sunny early. Then becoming cloudy with 30 percent chance of showers toward evening. Tonight: rain. Possibly heavy at times. Near 100 percent chance of rain. Can I help you with something else?
User:	YES WHAT ABOUT SAN FRANCISCO
JUPITER:	In San Francisco today: patchy fog near the ocean morning. Otherwise mostly sunny. Tonight: areas of coastal fog and low clouds. Otherwise fair. Is there something else I can do for you?
User:	YES HOW COLD IS IT IN DALLAS TEXAS TODAY
JUPITER:	In Dallas today: high upper 40s. Tonight: lows near 32. What else would you like to know?
User:	I'D LIKE TO KNOW WHETHER IT'S SNOWING IN DETROIT
JUPITER:	I'm sorry, I didn't understand you. Please rephrase your query.
User:	WHAT IS THE WEATHER IN DETROIT
JUPITER:	In Detroit this afternoon: occasional light rain and drizzle. Tonight: rain showers likely. Becoming snow showers by late tonight. 70 percent chance of precipitation.

Fig. 1. Example spoken dialogue between JUPITER and a user.

Zue, Victor, et al. "JUPITER: a telephone-based conversational interface for weather information." *IEEE Transactions on speech and audio processing* 8.1 (2000): 85-96.

@2000

# (CONT') JUPLTER

```

{c weather_response
 :continuant {c something_else4 }
 :db_tlist ({c weather_event
   :input "saturday"
   :city {p in
     :topic {q city :name "new york city" }}
   :pred {p month_date
     :topic {q date :day "saturday" }}}
 {c weather_event
   :topic {q weather_act
     :conditional "mostly"
     :name "sunny"
     :and {q iwind :name "brisk" }}
   :input "mostly sunny and brisk" }
 ...
 {c weather_event
   :topic {q weather_act
     :pred {p temp_qual :topic "chilly" }}
   :and {c weather_event
     :conjn "with"
     :topic {q lows
       :pred {p from_value
         :qualifier "around"
         :topic {q value
           :name 40 }}}
     :input "chilly with a low around 40" })
 :domain "Jupiter" }

```

Fig. 6. Excerpts from a response frame for the query "What is the weather going to be like tomorrow in New York?" The response by the system was "In New York City Saturday, mostly sunny and brisk, . . . , chilly with lows around 40. What other information can I give you?" Note: "c . . . =" clause, "p . . . =" predicate, and "q . . . =" quantified noun phrase.

User: SQL:	WILL IT SNOW TODAY IN NORTH DAKOTA select distinct city from geo, event where event.apt_code = geo.apt_code and event.snowspk is not null
JUPITER:	The forecast indicates snow for the following places in North Dakota: Bismarck, Dickinson, Devil's Lake, Fargo, Grand Forks, Williston, James-town, and Minot. Please select one.
User: SQL:	HOW ABOUT FARGO select distinct dayspk, snowspk from event, geo where geo.city = 'Fargo' and event.apt_code = geo.apt_code and event.day = 0 and event.snowspk is not null
JUPITER:	In Fargo in North Dakota today: becoming windy with light snow developing. Wind east increasing to 20 to 30 miles per hour producing drifting snow. Can I help you with something else?

@2000

***and to build***



# API.AI



API & DOCS

PRICING

BLOG

FORUM

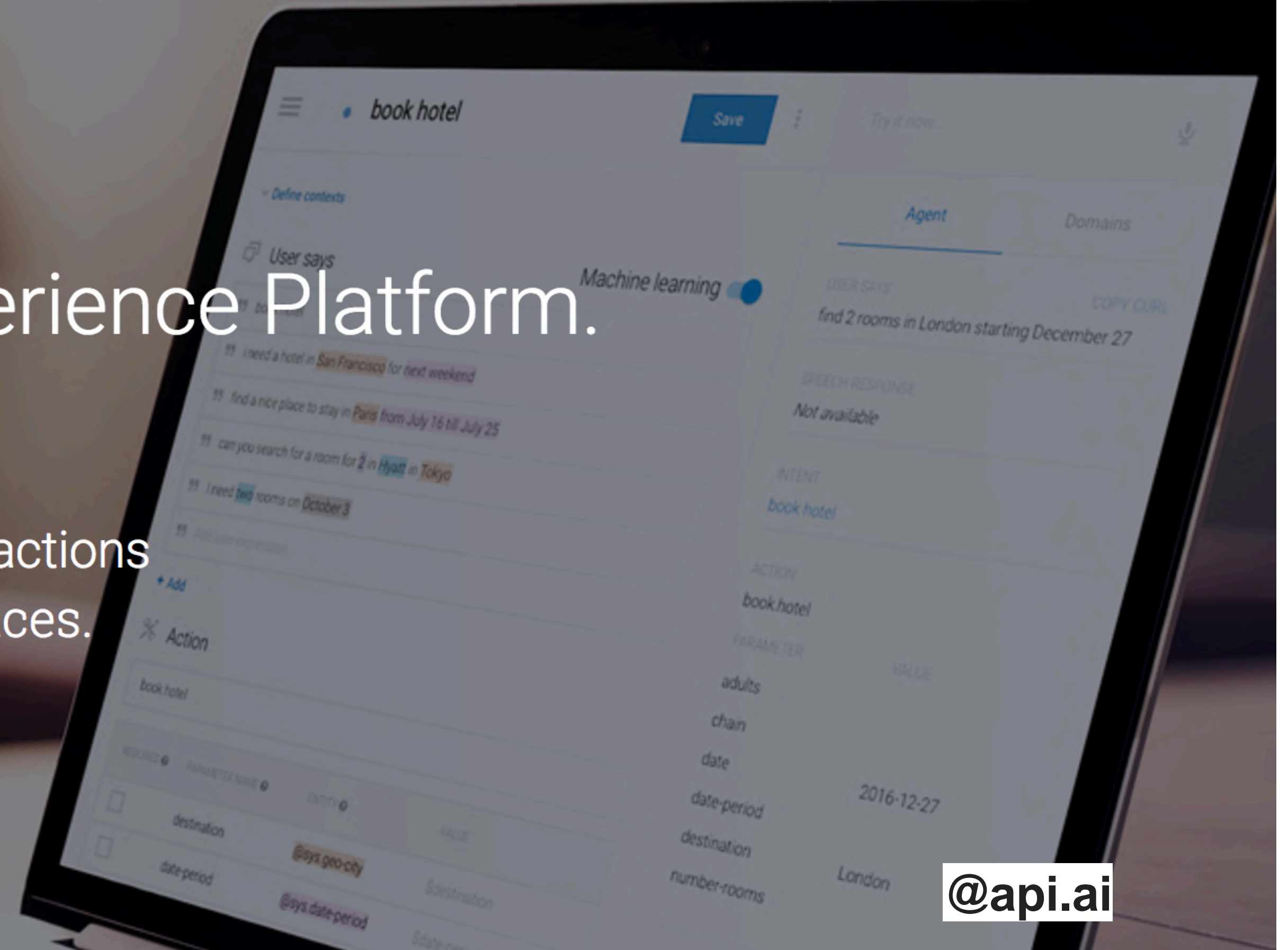
LOG IN

SIGN UP FREE

## Conversational User Experience Platform. Now we're talking.

Build brand-unique, natural language interactions for bots, applications, services, and devices.

Get Started for Free



@api.ai

# Order

Save

" Order a large pizza

" I would like to order 2 large pepperoni pizzas to 443 Waverley St at 2pm

" Add user expression...

+ Add

## Action

PizzaDelivery

REQUIRED ?	PARAMETER NAME ?	ENTITY ?	VALUE	PROMPTS ?
<input checked="" type="checkbox"/>	PizzaType	@PizzaType	\$PizzaType	Hi! What kind of a pizza are you looking for today? [2]
<input checked="" type="checkbox"/>	PizzaSize	@PizzaSize	\$PizzaSize	Size? [1]
<input checked="" type="checkbox"/>	customtime	@customtime	\$customtime	When would you like it to be delivered? [1]
<input checked="" type="checkbox"/>	address	@sys.address	\$address	Where should we deliver this pizza to? [2]
<input type="checkbox"/>	number	@sys.number	\$number	—



# CHINESE DEMO

Asics跑鞋推荐

场景 应用名称

Asics跑鞋推荐 设置

创建应用 查看所有应用

保存 设置完后保存

是否公开  私密

测试链接 <http://yige.ai/shareBot/2mKtn2> 测试页面

应用描述 应用描述

根据不同的脚型为用户推荐Asics跑步鞋

语言 中文(简体) 目前仅支持简体中文

时区 (GMT+8:00) Asia/Shanghai 目前仅支持上海时区

API密钥 接口访问令牌

客户端访问令牌 88802AAE0DBE2CC0A0FE9ADFC0515F17

默认回复 在没有匹配到合适场景时的默认回复

- 我还小, 只知道推荐Asics跑步鞋哦。
- 请不要问我Asics跑步鞋之外的问题哦。

@yige.ai

# WORDS EMBEDDING



@yige.ai

# ENVIRONMENT CREATION

The screenshot displays a web application interface for environment creation. The top navigation bar includes the logo '一个AI' and menu items: 'Asics跑鞋推荐', '场景' (Scene), '词库' (Word List), and '日志' (Log). The left sidebar contains '词库列表' (Word List) with a '+' icon, a search bar '搜索', and a list of items including 'Asics鞋系列' (Asics Shoe Series) and '脚型' (Foot Type). The main content area is titled 'Asics鞋系列' and features a blue '保存' (Save) button. A checkbox labeled '定义同义词' (Define Synonyms) is checked. A notification bar with an information icon and a close button contains the text '按enter键来分离同义词' (Press enter to separate synonyms). Below this, a table lists synonyms for 'Nimbus' and 'Cumulus':

Nimbus	Nimbus, NIMBUS, Gel-Nimbus
Cumulus	Cumulus, CUMULUS

At the bottom of the main area, there is a '+添加一行' (+Add a row) link. The bottom status bar shows the URL 'yige.ai/api-client/intents'.

@yige.ai



# CHAT

跑鞋示例 + 场景 词库 日志 扩展 文档中心 开发者论坛 xiaolei89tw

场景列表 +

搜索

最好的公司是什么?

2.1-购买链接-内翻

2.3-购买链接-外翻

2.2-购买链接-正常

**1.4-脚型-未知** 删除

1.3-脚型-外翻

1.2-脚型-正常

1.1-脚型-内翻

0-跑鞋推荐

邮费咨询

询问是否有货

### ★ 1.4-脚型-未知

保存

状态 ^

输入 shoe\_first 添加输入状态...

输出 shoe\_first 添加输出状态...

用户提问

不知道啊
不清楚
我哪知道
怎么知道是什么脚型
添加用户说语料...

+添加一行

动作

输入动作名称...

跑鞋示例

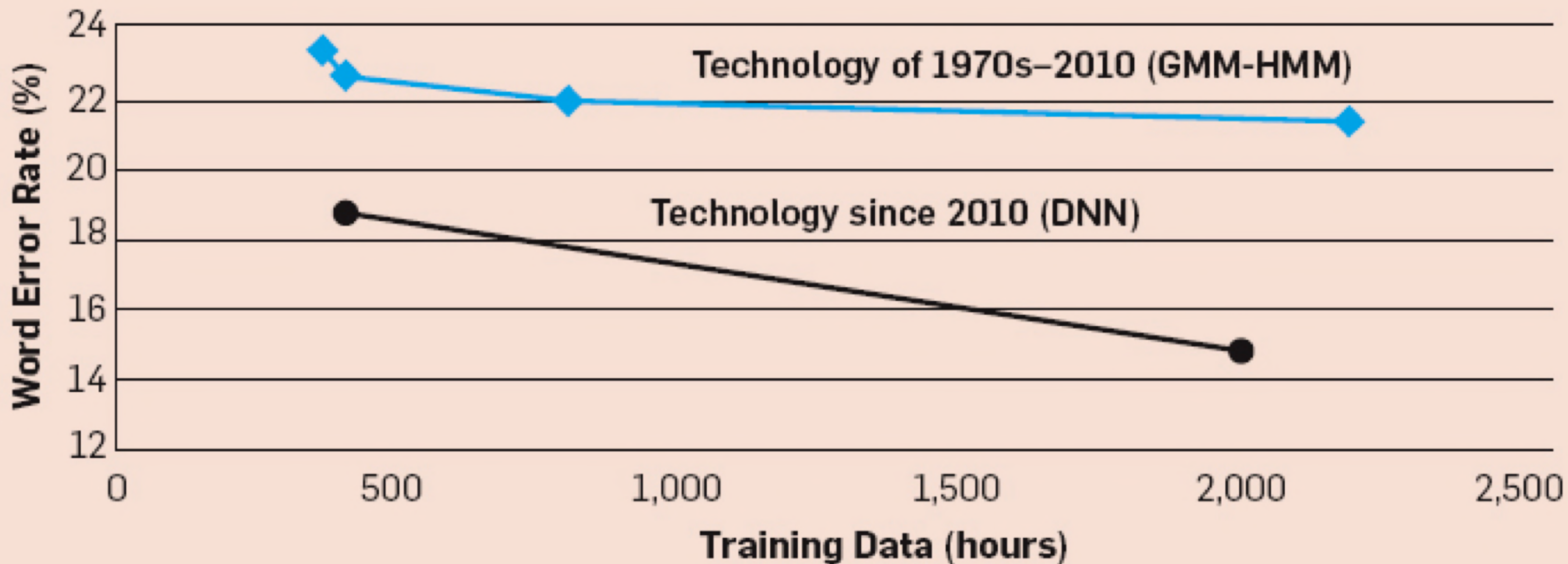
分享到: 

清除数据

发送

***GO DEEP***

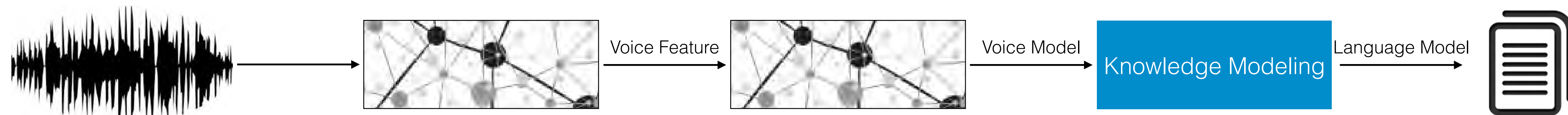
# DEEP GOES FAST



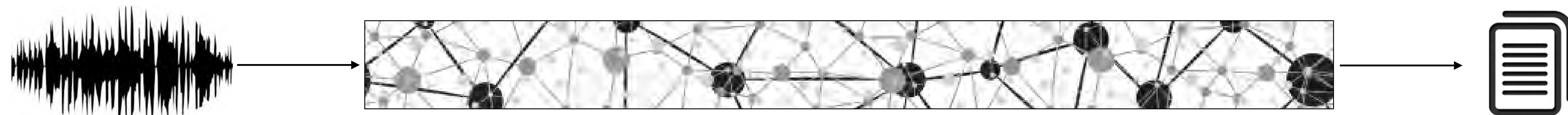
### Traditional Speech Recognition



### Early Deep Learning



### End to End with deep Learning

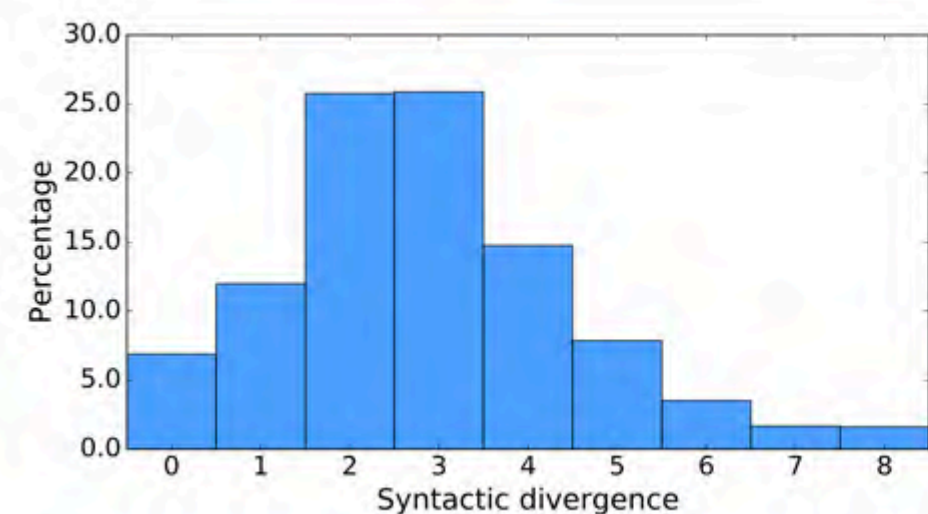




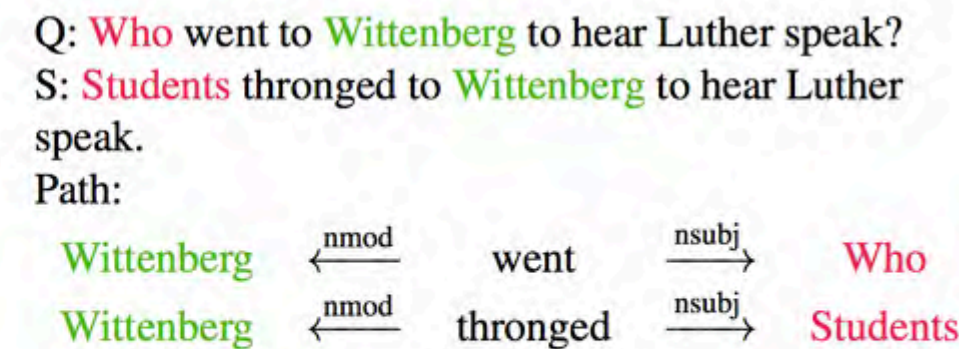
# KNOWLEDGE DATA SET

## SQuAD: 100,000+ Questions for Machine Comprehension of Text

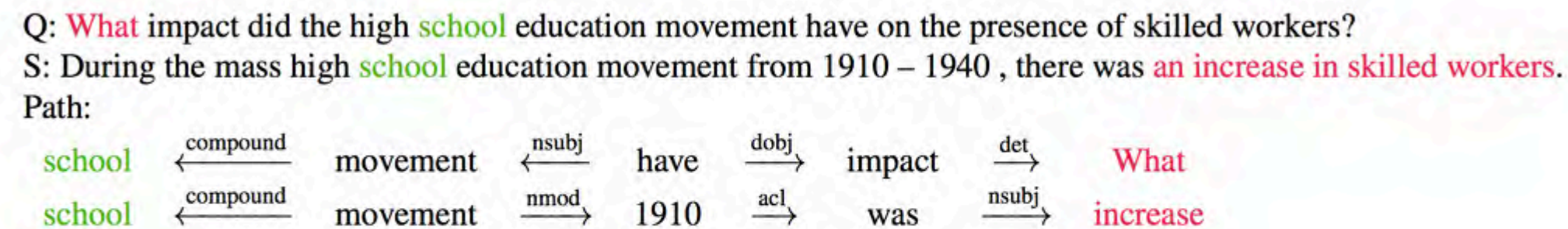
**Pranav Rajpurkar** and **Jian Zhang** and **Konstantin Lopyrev** and **Percy Liang**  
 {pranavsr, zjian, klopyrev, pliang}@cs.stanford.edu  
 Computer Science Department  
 Stanford University



(a) Histogram of syntactic divergence.



(b) An example of a question-answer pair with edit distance 0 between the dependency paths (note that lexical variation is ignored in the computation of edit distance).



(c) An example of a question-answer pair with edit distance 6.

**Figure 4:** We use the edit distance between the unlexicalized dependency paths in the question and the sentence containing the answer to measure *syntactic divergence*.

	Exact Match		F1	
	Dev	Test	Dev	Test
Random Guess	1.1%	1.3%	4.1%	4.3%
Sliding Window	13.2%	12.5%	20.2%	19.7%
Sliding Win. + Dist.	13.3%	13.0%	20.2%	20.0%
Logistic Regression	40.0%	40.4%	51.0%	51.0%
Human	80.3%	77.0%	90.5%	86.8%

**Table 5:** Performance of various methods and humans. Logistic regression outperforms the baselines, while there is still a significant gap between humans.



---

# ASK ME ANYTHING

---

I: Jane went to the hallway.  
I: Mary walked to the bathroom.  
I: Sandra went to the garden.  
I: Daniel went back to the garden.  
I: Sandra took the milk there.  
Q: Where is the milk?  
A: garden  
I: It started boring, but then it got interesting.  
Q: What's the sentiment?  
A: positive  
Q: POS tags?  
A: PRP VBD JJ , CC RB PRP VBD JJ .

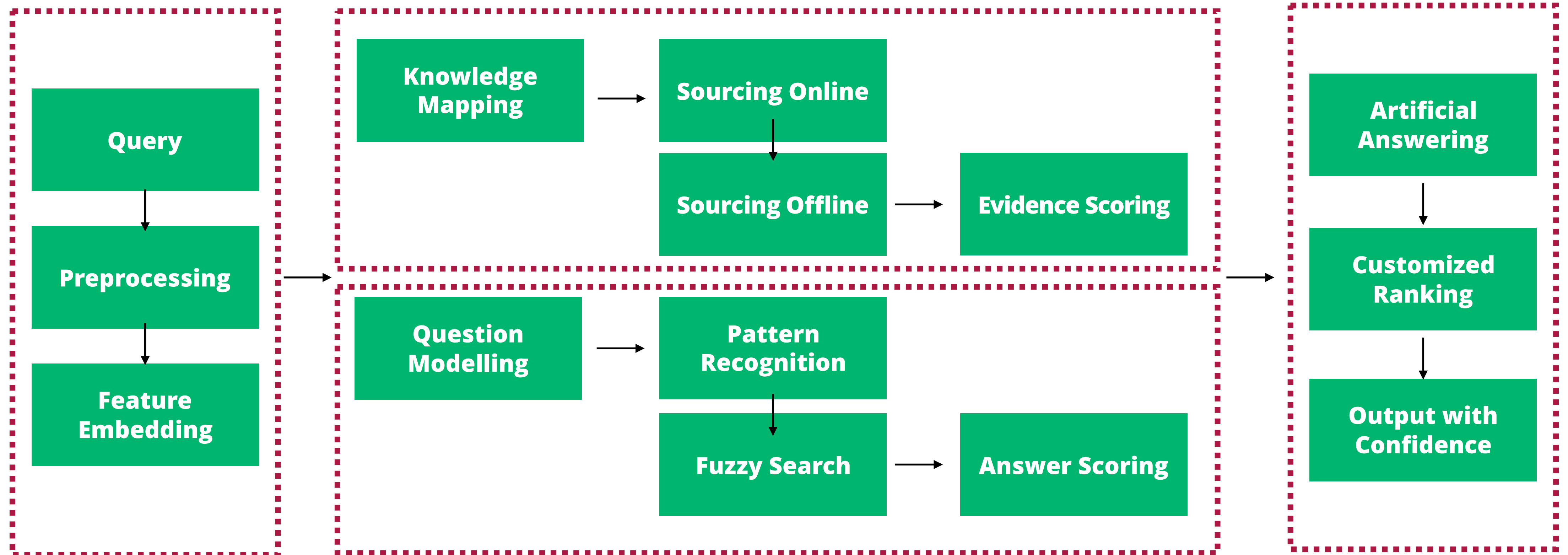
---

## Ask Me Anything: Dynamic Memory Networks for Natural Language Processing

---

Ankit Kumar\*, Ozan Irsoy\*, Peter Ondruska\*,  
Mohit Iyyer\*, James Bradbury, Ishaan Gulrajani\*,  
Victor Zhong\*, Romain Paulus, Richard Socher {JAMES.BRADBURY,RPAULUS,RSOCHER}@SALESFORCE.COM  
Salesforce Inc., CA USA \*Authors were interns at MetaMind.

# DATA FLOW DESIGN



# OUR DEMO - EARLY DEEP LEARNING





# THANK YOU

*Questions or feedback:*

*Xiaolei Wang: [xlewang@thoughtworks.com](mailto:xlewang@thoughtworks.com)*



**ThoughtWorks®**