Stanford CS224v Course

Conversational Virtual Assistants with Deep Learning

# Projects

Monica Lam

You are welcome
to propose your own projects.

Looking forward to a great quarter
inventing the future together!

# 3 Major Project Themes this Year

1. World Wide Knowledge (WWK)

2. Apprentice Training

3. Problem decomposition with a
   Well-Designed Intermediate Representation

# 1. World Wide Knowledge (WWK)

- Use AI to accelerate knowledge production & improve access
- Example projects:
  - **DataTalk**: Big Local News
    - Real-time data for election 2024
  - **HistoryChat**: Center of African Studies
    - Discover history from the ground up with African Times
  - **Multi-lingual news platform:**
    Armed Conflict Location and Event Data (ACLED)
    - Automatic multilingual qualitative coding of world news


# Cross-Lingual Multi-Perspective News Platform


**Tiffany Le**
Columbia Journalism School

**Jialiang Xu**
Stanford CS

**Monica Lam**
Stanford CS

**Ty Lawson**
Columbia Journalism School

**Jennifer Pan**
Stanford Political Science

Successful projects will be showcased at:

# A Workshop on a Public AI Assistant for WWK on Feb 13-14, 2025

- Sponsored by Sloan Foundation and HAI
- Invitation-only
  - Funding agencies:
    government, philanthropies, companies
  - Journalists, historians, social scientist, technologists

# 2. Apprentice Training

# Background: LLM-based Engineering

1. Static pipeline (Wikichat)
2. Agentic approach (e.g. Wikidata semantic parsing)
    - Define a set of actions, including evaluation (e.g. execution of query))
    - Agent iteratively: proposes solution, evaluates, backtracks, if necessary
    - New project: add tree search of solutions to the agentic pipeline.
3. Finetuning
    - When the instructions cannot fit in the prompts
    - To speed up or reduce cost of execution

*What if these techniques are not good enough?*

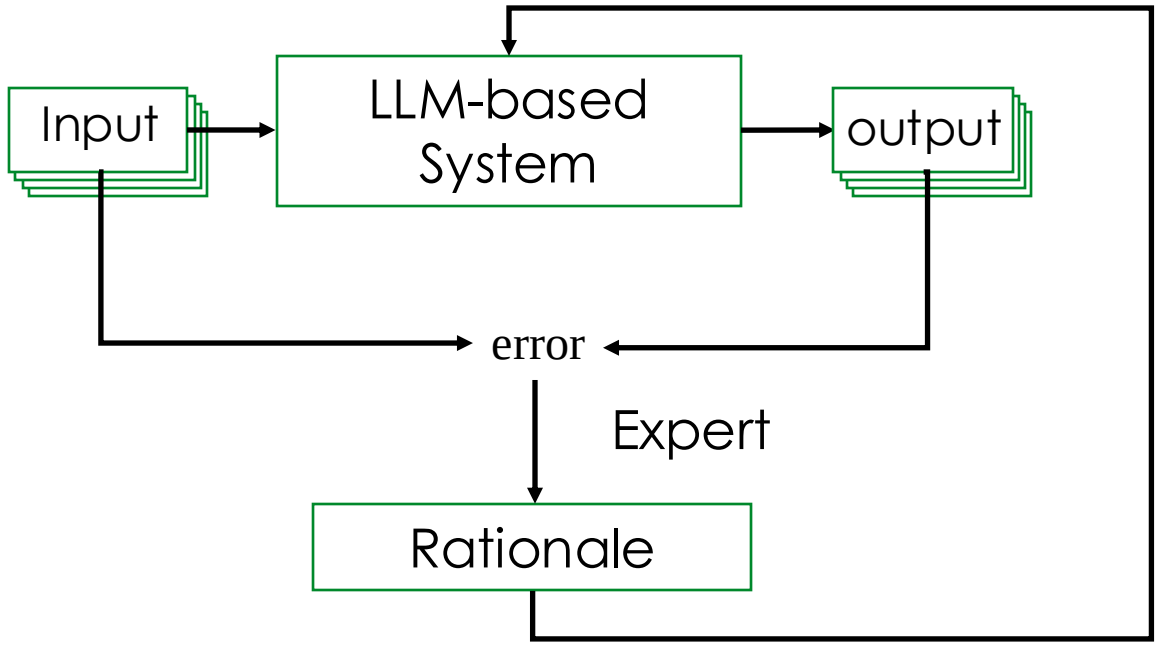
# Apprentice Training

- Key issue: *tacit knowledge*
  - Many tasks require info in experts' head
    - e.g. The donation column in the dataset does not include donations under $200.
      Therefore, adding the column up does not match the reported total.
  - We need to "distill the information" from experts
- Human solution: apprenticeships, internships – learn by practicing
  - Apply written knowledge, practice, ask, and get corected

# Data Driven → Rationale Driven

- Hard to get enough data points for high-level tasks
- Humans are data efficient: they learn from rationales:
  - E.g. when classifying events, an armed protest should be classified as an "armed event" and not a "protest event".
    - How many data points do you need to teach that principle?
- Humans can generalize rationales:
  - Choose the event type that is most escalated

# Architecture of Apprentice Training




e.g. Rules to apply in agentic approach

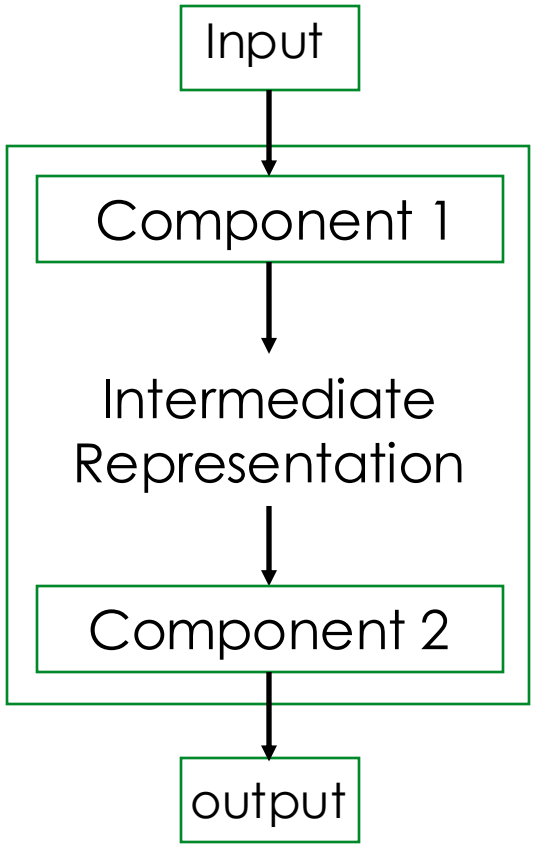
- check the different rules in different prompts?

# Research Approach

- Apply this technique to many more domains:
  - Currently used in 2 real-world scenarios
    Election Datatalk & ACLED event classification
- Create a general apprentice framework

# 3. Intermediate Representations

# Improves Data Efficiency
# Well-Designed Intermediate Representation




- **Massive data analysis**: qualitative coding
  - Coding according to a given codebook (events, resumes, medical transcripts)
  - Auto-designed schema by inferring them with questions from experts – they can simulated
- **Compliance**: SMT (Satisfiable Modulo Theory)
  - Course requirements, invoice auditing

AI

# DEPARTMENT OF COMPUTER SCIENCE
## MSCS Program Sheet (2023-24)
## Artificial Intelligence Specialization

Name: _____ Advisor: _____ Date: _____

Student ID#: _____ Email: _____ Proposed date for degree conferral: _____ HCP? _____ Coterm? _____

**GENERAL INSTRUCTIONS:** Before the end of your first quarter, complete a program sheet by filling in the number, name and units of each course you intend to take for your degree. Program sheets, as well as foundation course waiver forms, may be submitted electronically on the Gates Information Network at http://gin.stanford.edu/ using your CS ID and password (click on the 'Dashboard' link). See http://cs.stanford.edu/degrees/mscs/programsheets/psguide2324.pdf for detailed instructions.

**FOUNDATIONS REQUIREMENT:** You must satisfy the requirements listed in each of the following areas; all courses taken elsewhere must be approved by your advisor via the foundation course waiver form. Do not enter anything in the "Units" column for courses taken elsewhere or for Stanford courses counted towards your undergraduate degree.

| **Required:** | **Equivalent elsewhere** (course numbers/titles/institution) | **Grade** | **Units** |
|---|---|---|---|
| Logic, Automata & Complexity (CS103) | | | |
| Probability (CS109, Stat116, CME106, MS&E220, or EE178) | | | |
| Algorithmic Analysis (CS161) | | | |
| Computer Organ & Sys (CS107 or 107E) | | | |
| Principles of Computer Systems (CS111) | | | |

**TOTAL UNITS USED TO SATISFY FOUNDATIONS REQUIREMENT** (May not exceed **10** units.) ____

**SIGNIFICANT IMPLEMENTATION REQUIREMENT:** At least one course in your MS program should be designated as satisfying the Significant Implementation Requirement. Note that this course will also be used to satisfy some other requirement (i.e., Depth, or Elective). Must be taken for a letter grade,* must be taken at Stanford. Coterm students who took two of these courses at Stanford as undergraduates may waive this requirement. Deviations must be approved by the faculty director of the MS program, Cynthia Lee.

| **Course Number:** | **Title:** | **Grade** |
|---|---|---|
| | | |

***Courses that may be used to satisfy the significant implementation requirement include:*** CS 140, 140E, 143, 144, 145, 148, 151, 190, 210B, 212, 224R, 227B, 231N, 239, 243, 248A, 248B, 330, and 341.

**BREADTH REQUIREMENT:** Three courses, with each course chosen from a different Breadth area A, B, C or D. Breadth courses may not be waived, must be taken for at least 3 units each, and must be completed for a letter grade.* Each of the three Breadth courses must be from different Areas. Note that these courses will also be used to satisfy some other requirement (i.e., Depth or Elective). Letter grade only.*

| **Course Number:** | **Title:** | **Grade** |
|---|---|---|
| | | |
| | | |
| | | |

**Area A**. Formal Foundations: CS 154, 157, 166, 168, 205L, 229T, 250, 254, 254B, 255, 256, 257, 258, 259Q, 261, 263, 265, 269I, 328, 331, 334A, 354, 355, 359D, 369O; EE 364A, 364B; Phil 251

**Area B**. Learning and Modeling: CS 173A, 205L, 221, 223A, 224N, 224R, 224S, 224U, 224V, 224W, 227B, 228, 229, 229B, 229T, 230, 231A, 231N, 233, 234, 236, 237A, 237B, 248A, 248B, 271, 272, 273A, 273B, 274, 279, 281, 326, 329D, 329X, 330, 332, 333, 348A, 348B, 348C, 348E, 348I, 348N, 371, 373

**Area C.** Systems: CS 112 (with CS111 as prereq), 140E, 143, 144, 145, 149, 155, 212, 240, 240LX, 242, 243, 244, 244B, 245, 246, 249I, 316, 348K, 356, 358; EE 180, 282, 382E

**Area D**. People and Society: CS 147, 148, 152, 181/W, 182/W, 194H, 206, 247 (any suffix), 256, 269I, 278, 281, 329T, 329X, 347, 377 (any suffix), 384, 448B; AMSTUD 145; ANTHRO 132C, 132D; COMM 254, 266, 286, 324; DESIGN 255; EARTHSYS 213; EDUC 315A; ENGLISH 184D, HISTORY 244F; INTLPOL 268; LAW 4039; ME 177; MS&E 193, 231, 234, 254; POLISCI 150A; PUBLPOL 103F for 3 units, 353B

### ARTIFICIAL INTELLIGENCE DEPTH

All depth courses must be taken for a letter grade* for 3 or more units. A maximum of 6 units of CS 399 Independent Study may be counted towards the depth. Any deviations from the stated requirements must be noted and approved by your advisor in the Advisor Notes box provided (maximum of one advisor-approved deviation allowed). Courses taken for your Stanford undergraduate degree do not need to be repeated. Enter course selection in the table below:

a) CS 221 (students with equivalent course work may waive with advisor approval)

b) At least four of: CS 223A, 224N, 224R, 224S, 224U, 224V, 224W, 228, 229, 231A, 231N, 234, 237A, 237B, 238

c) A total of at least 21 units from categories (a), (b) and the following: CS 205L, 224C, 224R, 225A, 227B, 229B, 229S, 229T, 230, 233, 235, 236, 239, 246, 257, 270, 271, 173A/273A, 273B, 274, 275, 279, 281, 293, 322, 324, 324H, 325B, 326, 327A, 328, 329 (any suffix), 330, 331, 332, 333, 336, 348I, 348N, 353, 361, 368, 371, 375, 377† (any suffix), 379† (any suffix), 398, 399†, 428A, 428B, 432; EE 263, 276, 278, 364A, 364B, 377; ENGR 205, 209A; MS&E 226, 252; PSYCH 209; STATS 202, 315A, 315B (courses with † notation require approval of MS advisor)

**Advisor Notes**

| **Course number** | **Title (Depth courses must be taken for at least 3 units.)** | **Grade** | **Units** |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

**TOTAL DEPTH UNITS APPLIED TO MSCS (must total at least 21 units)** ***Letter grades only.**** ____

### ELECTIVES

List here any additional courses used to complete the 45-unit requirement for the MSCS degree. You may count up to a maximum of **3 units** of 1-2 unit seminars offered in the School of Engineering. All other electives must be taken for 3 or more units. CS courses numbered above 111, excluding CS 161A, 196, CS198, and CS390A/B/C, can be used as MS electives. Non-CS courses must be technical courses numbered above 100, related to the degree program, and approved by the advisor and the MS program administrator. *Note that CS 129 may not be counted towards the MS if CS 229 is being counted towards any BS or MS requirement.*

| **Course number** | **Title** | **Grade** | **Units** |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

**TOTAL ELECTIVE UNITS APPLIED TO MSCS** ____

**TOTAL UNITS APPLIED TO MSCS** ____

### ADDITIONAL REQUIREMENTS

- All courses submitted for the MSCS degree must be numbered 100 or greater.
- At most 10 units of Foundations requirement courses may be counted toward your 45 units.
- At most 3 units of 1-2 unit seminars may be counted toward your 45 units.
- At least 36 units submitted for the MSCS degree, including all courses taken for breadth and depth, must be taken for a letter grade.*
- The average grade in the courses submitted for the MSCS must be at least a B (3.0 in Stanford's GPA scale).
- Units previously applied toward BS requirements may not also be counted toward the MSCS.
- You must complete at least 45 graduate units at Stanford before receiving the MSCS degree.

*All classes taken Spring 19-20 and Fall through Summer 20-21 for a CR or S grade will satisfy MSCS requirements as if taken for a letter grade.

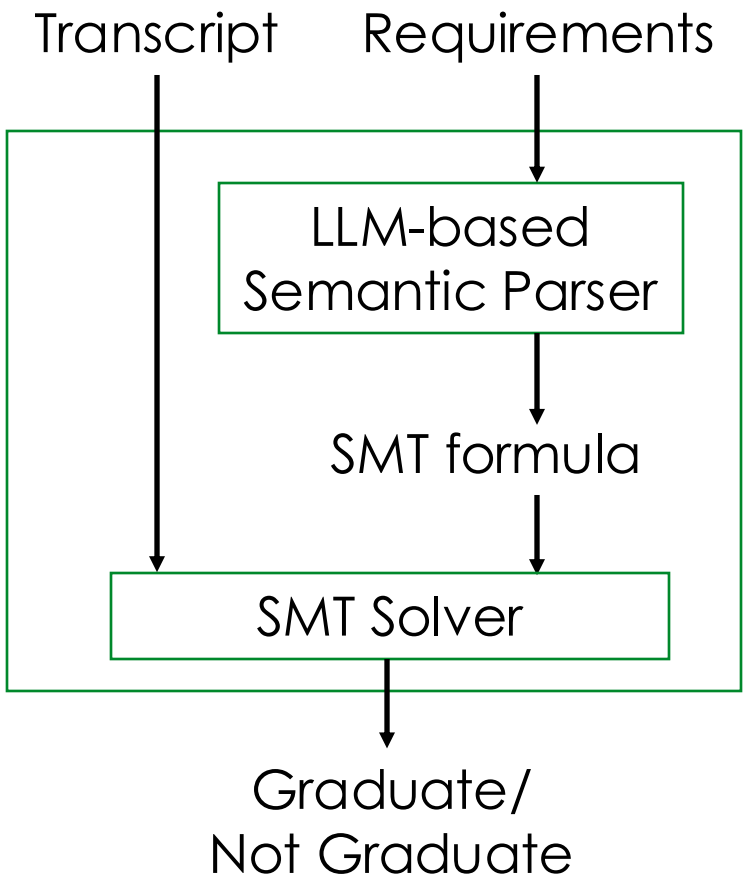
Advisor's signature: _____ Date: _____

LAM

STANFORD

# Intermediate Representation: SMT

(Satisfiability Modulo Theory)




- LLMs are reasonably good for simple conditions
  - But what if it screws up?
- Reduce the LLM complexity
  - Translate to just the SMT constraints
    - A simpler and necessary skill
    - Interpretability: constraints can be reviewed once and for all
  - Use a theorem prover
    - Guaranteed to be correct

# An Easy Example

FOUNDATIONS REQUIREMENT:

You must satisfy the requirements listed in each of the following:

1. Logic, Automata & Complexity (CS103)*
2. Probability (CS109, Stat116, CME106, MS&E220, or EE178)
3. Algorithmic Analysis (CS161)
4. Computer Organ & Sys (CS107 or 107E)
5. Principles of Computer Systems (CS110 or CS111)

# LLM-Generated SMT Formula

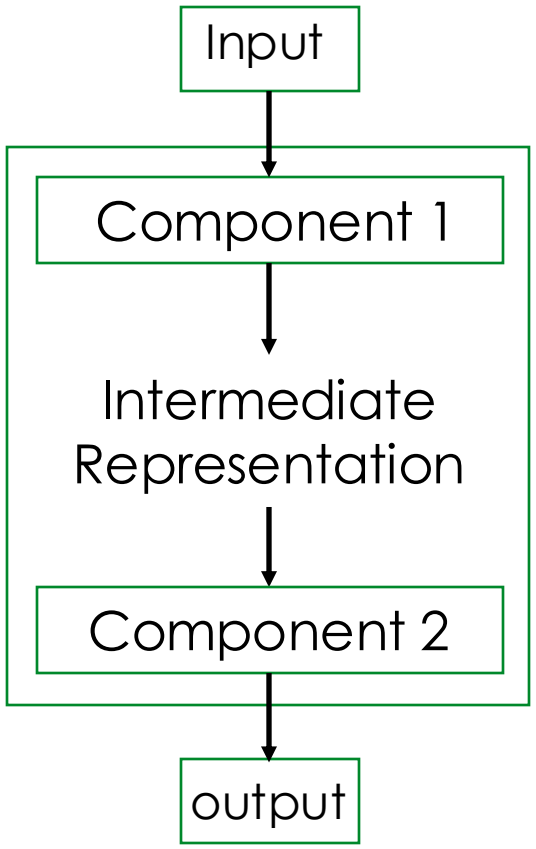
1. Logic, Automata & Complexity (CS103)*
2. Probability (CS109, Stat116, CME106, MS&E220, or EE178)
3. Algorithmic Analysis (CS161)
4. Computer Organ & Sys (CS107 or 107E)
5. Principles of Computer Systems (CS110 or CS111)

```
(set-logic ALL)
;; declaring variables
(declare-const course1 String) (declare-const course2 String) (declare-const course3 String) (declare-const course4 String) (declare-const course5 String)

;; check required course(s) of each group
assert(and(and(and(and((= course1 "CS103")
(or (= course2 "CS109") (= course2 "Stat116") (= course2 "CME106") (=course2 "MS&E220")  (=course2  "EE178")))
(= course3 "CS161")) (or (= course4 "CS107") (= course4 "CS107E")))
(or (= course5  "CS110") (= course5 "CS111")))

;; check required course(s) of each group are in aa given transcript
assert(and(and(and(and((= course1 course.get("Course_Id") for course in Transcript.get("Courses_Taken")
(= course2 course.get("Course_Id") for course in Transcript.get("Courses_Taken"))
(= course3 course.get("Course_Id") for course in Transcript.get("Courses_Taken"))
(= course4 course.get("Course_Id") for course in Transcript.get("Courses_Taken"))
(= course5 course.get("Course_Id") for course in Transcript.get("Courses_Taken"))
(check-sat)
```

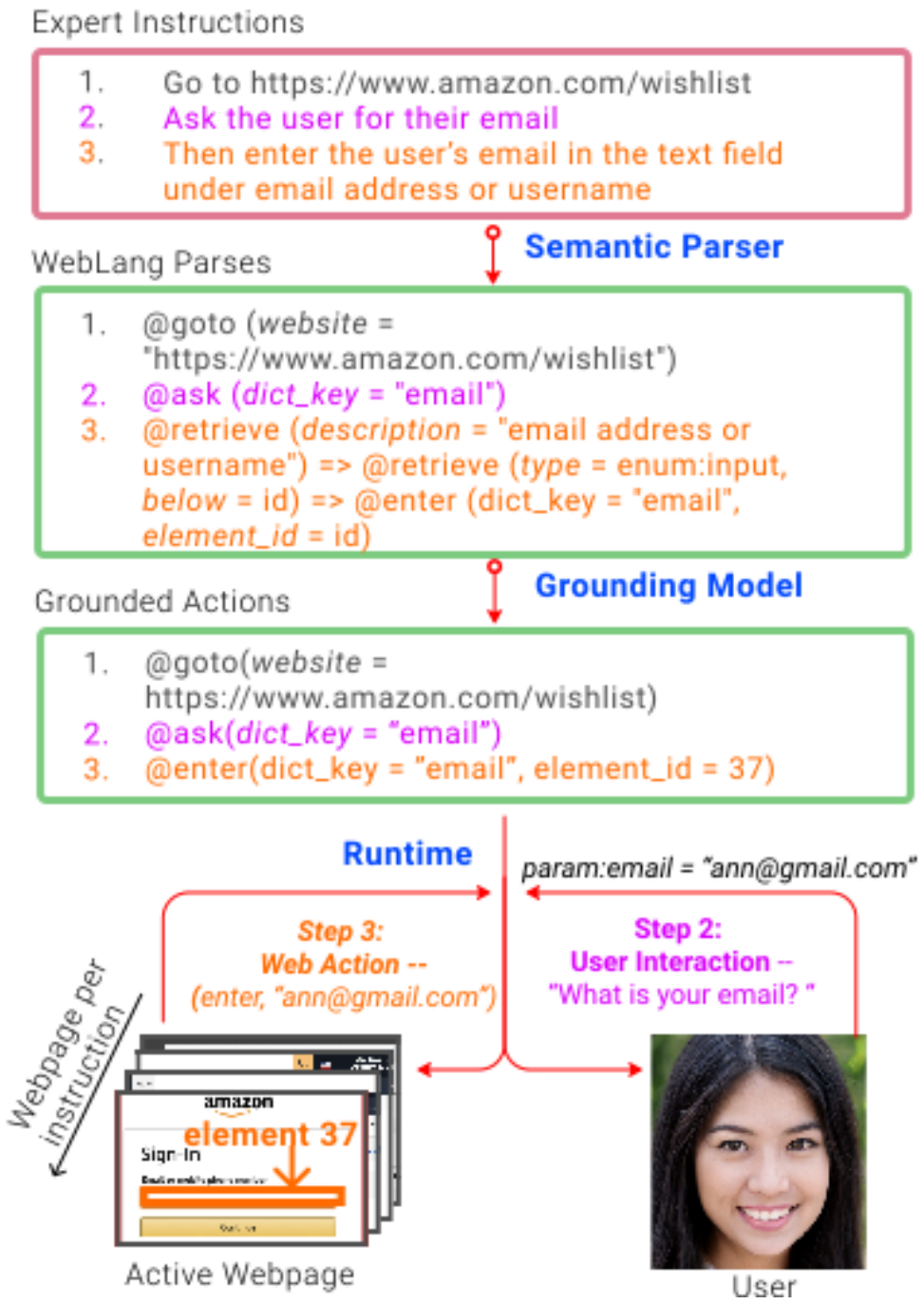
# Improves Data Efficiency
# Well-Designed Intermediate Representation




- **Massive data analysis**: qualitative coding
  - Coding according to a given codebook (events, resumes, medical transcripts)
  - Auto-designed schema by inferring them with questions from experts – they can simulated
- **Compliance**: SMT (Satisfiable Modulo Theory)
  - Course requirements, invoice auditing
- **Web agents**: Verbal description of visual elements on web pages

# Universal WebAgent

- Build a conversational agent to replace a web interface, given instructions.
- Large dataset for webtasks: The Mind2Web dataset

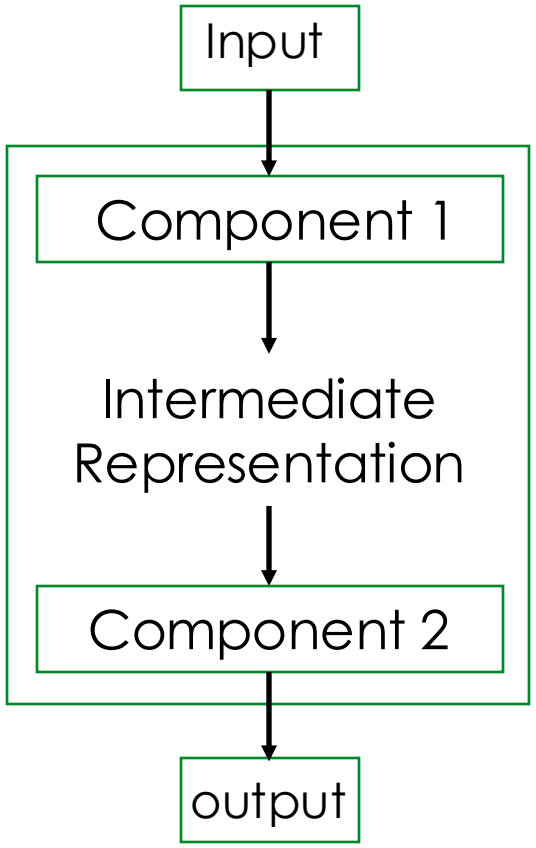




*Grounding Open-Domain Instructions to Automate Web Support Tasks*
Nancy Xu, Sam Masling, Michael Du, Giovanni Campagna, Larry Heck, James Landay, Monica S Lam, NAACL '21.

Mind2Web: Towards a Generalist Agent for the Web. Xiang Deng, Yu Gu, Boyuan Zheng, Shijie Chen, Samuel Stevens, Boshi Wang, Huan Sun, Yu Su. In NeurIPS, 2023

# Improves Data Efficiency with a Well-Designed Intermediate Representation




- **Massive data analysis**: qualitative coding
  - Coding according to a given codebook (events, resumes, medical transcripts)
  - Auto-designed schema by inferring them with questions from experts – they can simulated
- **Compliance**: SMT (Satisfiable Modulo Theory)
  - Course requirements, invoice auditing
- **Web agents**: Verbal description of visual elements on web pages
- **Papers to decision making**: Derive a decision tree

# 3 Major Project Themes this Year

1. World Wide Knowledge (WWK)

2. Apprentice Training

3. Problem decomposition with a
   Well-Designed Intermediate Representation