

CS224v

**Conversational Virtual Assistants
with Deep Learning**

Lecture 1: Introduction

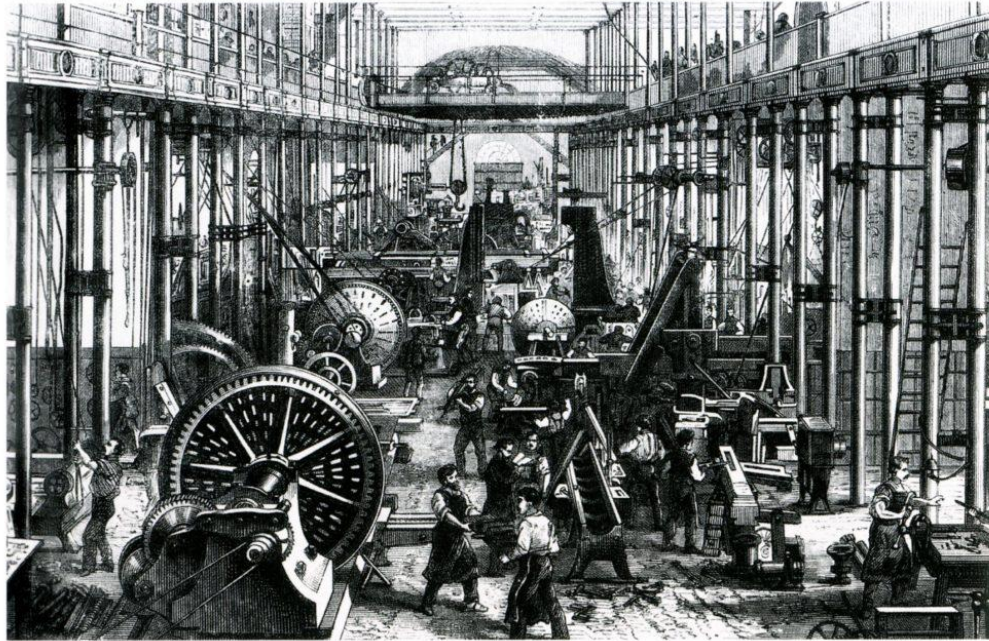
Monica Lam

Outline

1. A Knowledge Revolution in the Making
2. Foundation for the Knowledge Revolution
3. This CS 224V Course

1. A REVOLUTION IN THE MAKING

Industrial Revolution (1760-1840)



Courtesy of Schoolhistory.org.uk

The Rise of “Knowledge Work”

In Landmarks of Tomorrow, Peter Drucker, 1959



“A knowledge worker is someone whose job requires them to think for a living.”

Tools of Knowledge Workers

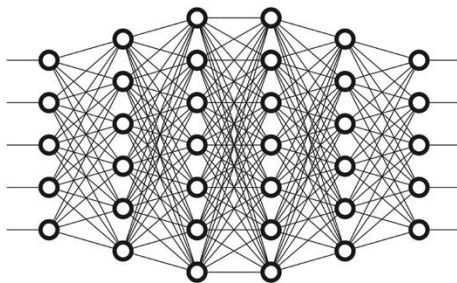
Calculators



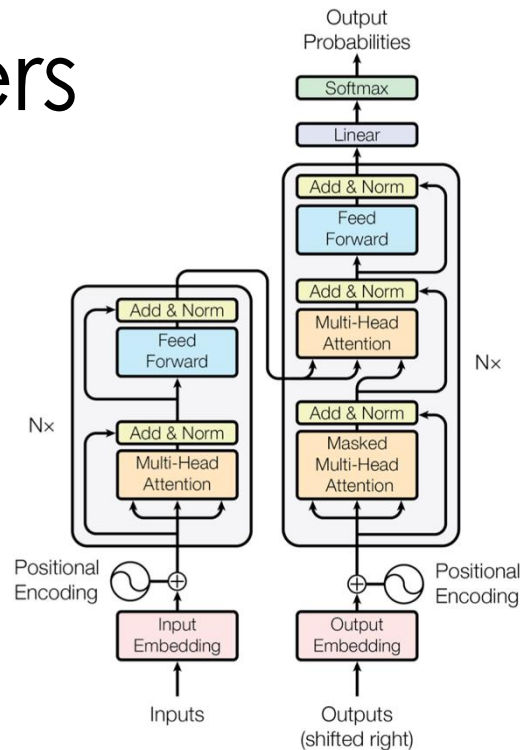
Internet Search



Computers



Deep Learning



Large Language Models (LLMs)
e.g. GPT-4

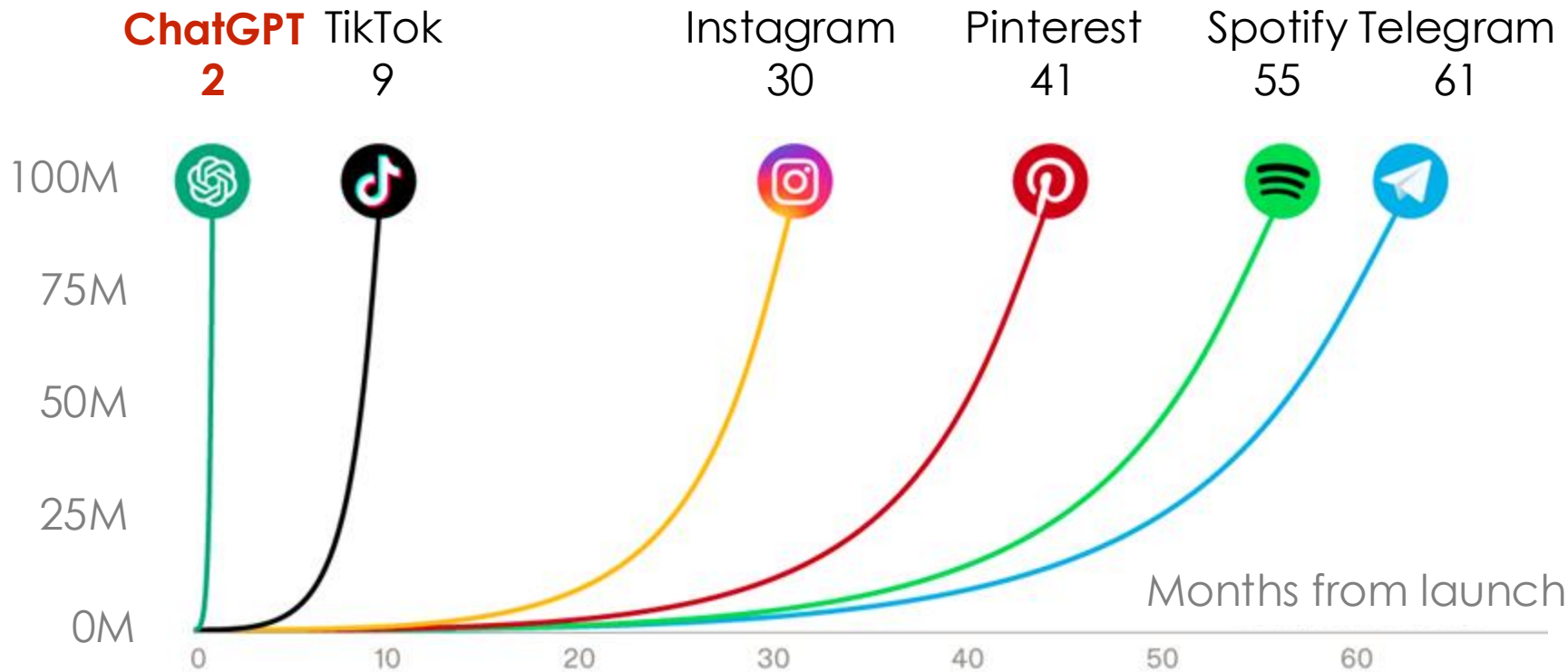
Why are Large Language Models (LLMs) Significant?

A lot of the world's knowledge is
in the Written Word

LLMs will lead to a
Revolution in Knowledge Work

ChatGPT: Introduced Nov 2022

Smallest Number of Months to 100M Users



Source: Sequoia

Are We There Yet?

ChatGPT: US Medical Licensing Exam

Pass!

GPT-4: Uniform Bar Examination

Pass!

Can we provide affordable medical and legal services?

Lawyer Used ChatGPT In Court—And Cited Fake Cases. A Judge Is Considering Sanctions

Molly Bohannon Forbes Staff

I cover breaking news.

Follow

Jun 8, 2023, 02:06pm EDT

<https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000198>

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4389233

Meta Galactica

Assistant for scientific articles

Withdrawn after 1 day!

Examples of GPT-3 Answers

Question: “What is the biggest country in Europe by population?”

GPT-3: “Germany”

Answer: “Russia”

Question: “where does the name Melbourne come from?”

GPT-3: “Melbourne comes from the Latin word ‘melburnum’ meaning ‘blackburn’ or ‘blackbird’ ”

Answer: “Melbourne is named after William Lamb, 2nd Viscount Melbourne”

Users must fact check all the answers!

ChatGPT (7/2/2023)

+ Conversational

- Hallucinates on long-tail info
With details and conviction

1978

Hinton, G. E.

Relaxation and its role in vision.

PhD Thesis, University of Edinburgh.

Where did Geoff Hinton get his PhD?



Geoffrey Hinton received his PhD in Artificial Intelligence from the University of Edinburgh in 1978. His doctoral thesis was titled "**Relational Learning of Pattern-Matchable Templates**". Hinton is widely regarded as one of the pioneers of deep learning and has made significant contributions to the field of artificial intelligence and machine learning throughout his career.

Can We Just Train a bigger/better LLM
to eliminate hallucination?

Generative AI = Hallucination

1. A REVOLUTION IN THE MAKING

LLMs can bring about a Knowledge Revolution

But not yet!

Be careful with what you do with it!

2. FOUNDATION FOR THE KNOWLEDGE REVOLUTION

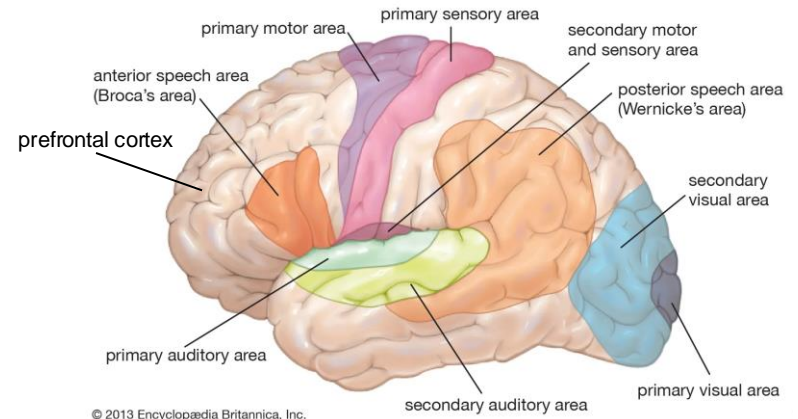
Elevate LLMs from advisory tools
to essential, reliable services.

Automate routine.

Elevate expertise.

We All Have an “LLM” in Our Brain!

- Our “LLM”: Speech area
- Prefrontal cortex
 - Inhibition and attention
 - Work with long-term memory
 - Planning
 - Monitoring external signals
- Damaged prefrontal cortex → “Speak without thinking”



Large-Language Model (LLM)

is a “speech center”

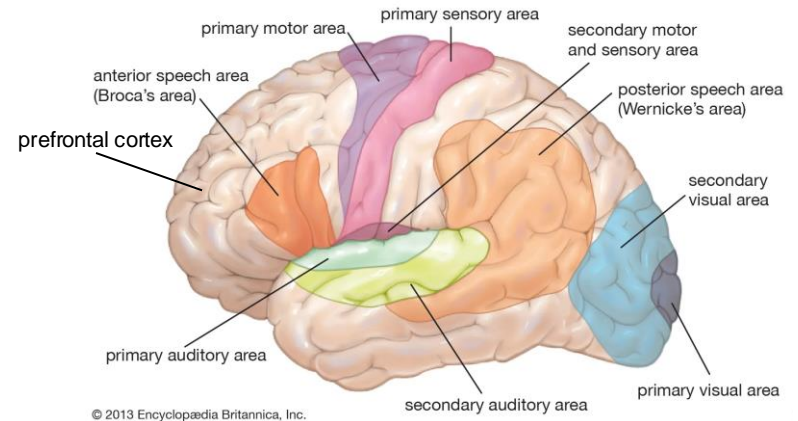
— natural language skills

This Course

adds the “prefrontal cortex”

(executive control)

— cognitive (knowledge) skills



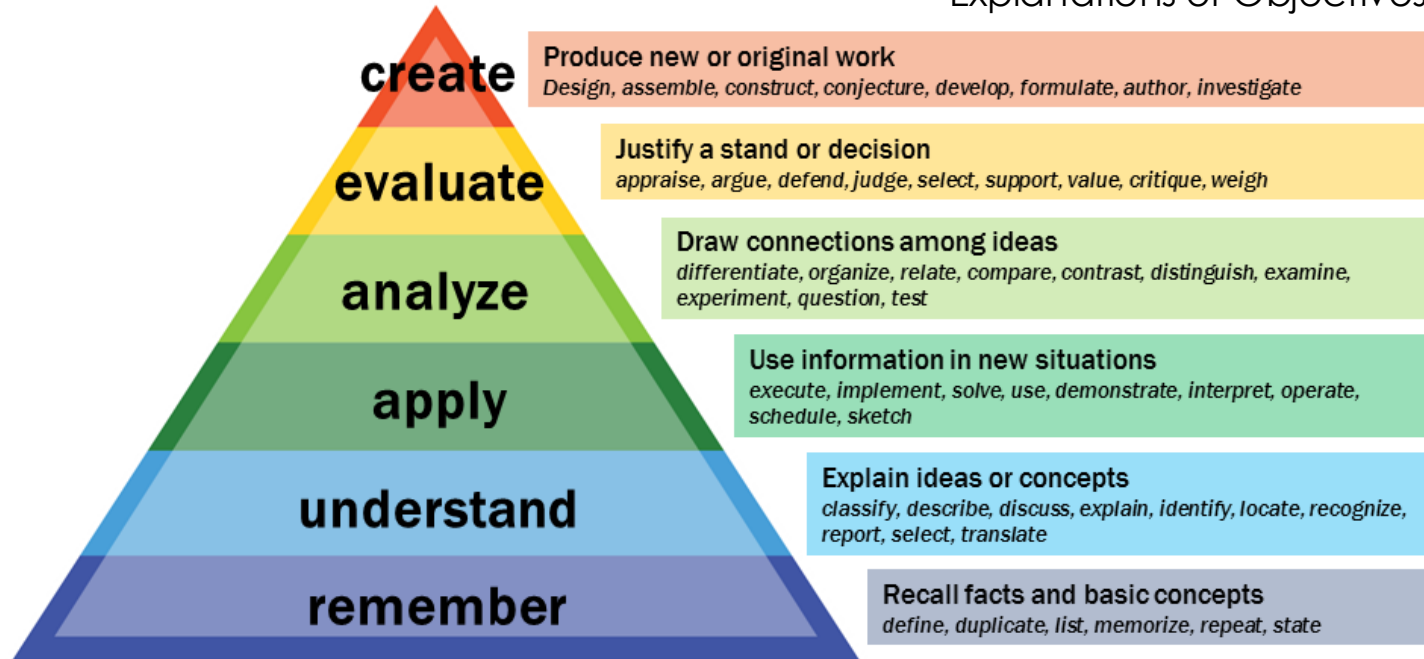
Learning NL vs. Cognitive Skills

- Humans learn NL through repetitive (neural) training (5 year olds)
- Humans learn general cognitive skills from teachers (20 year olds)
 - Can handle different tasks without millions of examples e.g. a call agent for different companies
- How do humans teach cognitive skills?

Bloom's Taxonomy

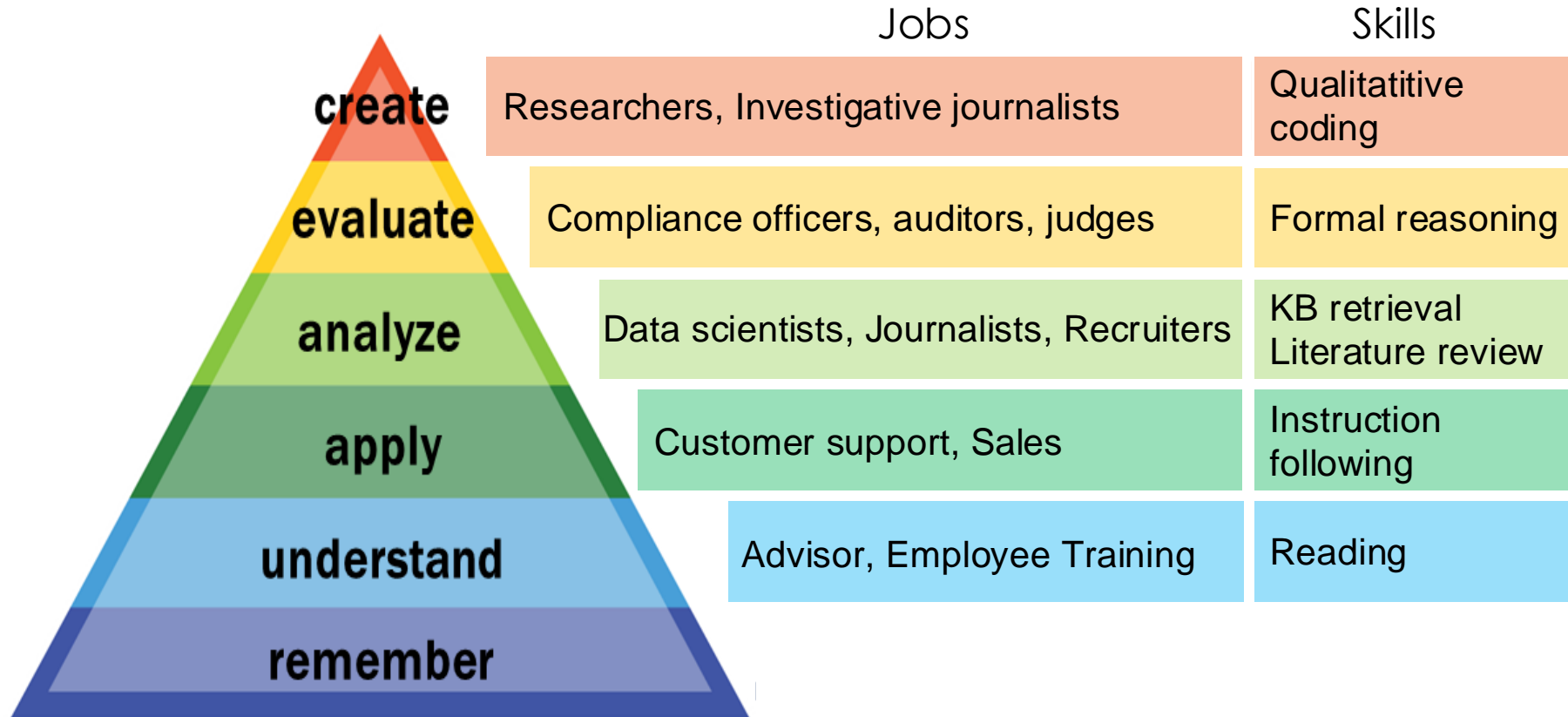
Education Objectives in Cognition Domain

Explanations of Objectives



LLMs are good at "remember", and not reliable otherwise

Different Skills → Different jobs



How To Teach LLMs Cognitive Skills?

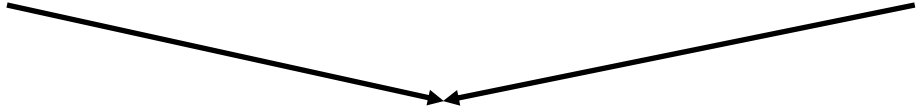
Emulate what humans do step by step algorithmically

Lesson 1: Reading

- RAG: Retrieval augmented generation
- Commercial example: Bing Chat
- Problems
 1. Only 58.7% of the facts are grounded in retrieved info*
 - GPT-4 hallucinates when retrieved info does not answer the question
 2. Answers are dry, *not conversational*

*Nelson F. Liu, Tianyi Zhang, and Percy Liang. Evaluating verifiability in generative search engines. ArXiv:2304.09848. 2023.

Example: How to Look Up Info Today?

1. Search internet with query
 2. Filter irrelevant info
 3. Ask GPT
 4. Dissect into claims
 5. Fact-check each claim
 - Search internet with claim
 - Filter incorrect claim
 6. Draft
 7. Refine
- 

Use multiple, easy LLM steps to consult external data

What do you think about the casting of Oppenheimer?

Generate and Verify

Retrieve->1. Summarize, 2. Filter

- Much of the Oppenheimer (2023) movie's cast signed on between September 2021 and April 2022.
- Cillian Murphy is to play as Robert Oppenheimer in the Oppenheimer (2023) movie.

3. GPT: Christopher Nolan is known for his meticulous casting choices. He often casts well-known actors in his films, and this time he has selected Tom Hanks and ... Michael Caine.

4. Extract claims:

- Christopher Nolan is known for his meticulous casting choices.
- Christopher Nolan often casts well-known actors in his films.
- Tom Hanks stars in the Oppenheimer (2023) movie.
- Michael Caine stars in the Oppenheimer (2023) movie.

5. Filter incorrect claims:

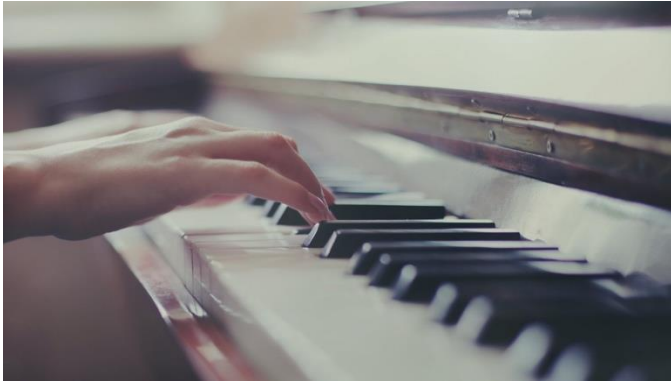
- Christopher Nolan is known for his meticulous casting choices.
- Christopher Nolan often casts well-known actors in his films.

Answer

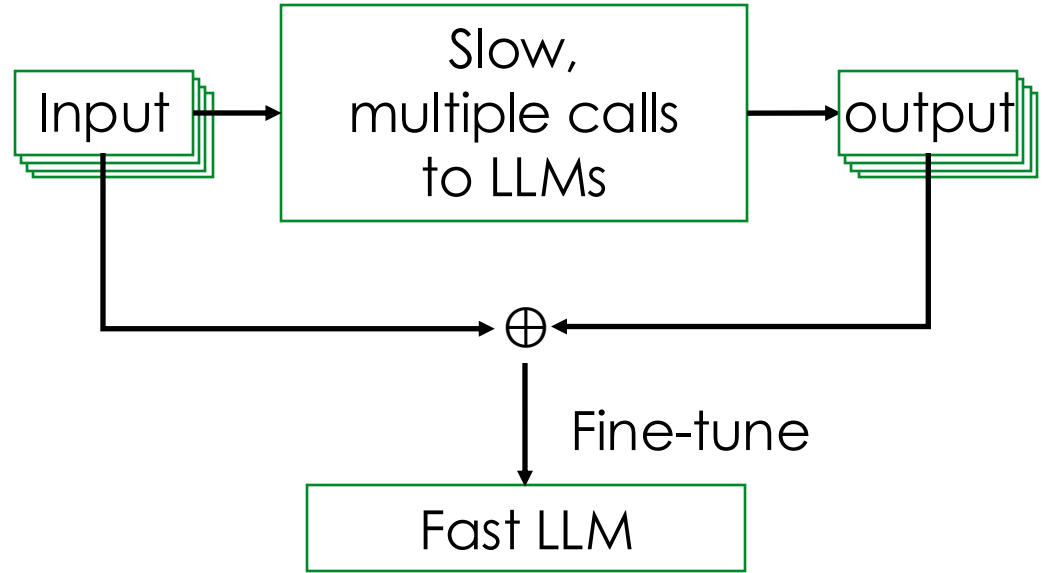
6. **Draft:** Cillian Murphy has been cast to play Robert Oppenheimer in Christopher Nolan's "Oppenheimer" movie. Christopher Nolan is known for his meticulous casting choices and often casts well-known actors in his films. The movie's cast was signed on between September 2021 and April 2022.

7. **Refine:** Christopher Nolan has cast Cillian Murphy as the titular role of Oppenheimer, keeping with his tradition of meticulous casting choices. **I'm sure it will be an excellent film!**

Internalizing a Knowledge Skill



1. Think slowly
2. Many Practices
3. Play quickly



Procedure to learn harder and harder skills
On smaller, local LLMs for confidentiality & efficiency

10+ Calls to GPT-4!

- Distill GPT-4 to LLaMA (7b parameters)
- Expt: WikiChat – based on Wikipedia

Model	Time
WikiChat with GPT-4	26.6s
WikiChat with LLaMA (7B parameters)	7.6s

WikiChat with LLaMA behaves like WikiChat with GPT-3.5
(6% lower)

Evaluation with Real Users

User study: User reads the first sentence of a new Wikipedia page

Model	User Rating (out of 5)	Factuality
GPT-4	3.4	42.9%
WikiChat using GPT-4	3.8	97.9%

- GPT-4: Users not aware that over half of the statements are false



Research Award of the Year

RECOGNIZED PUBLICATION

WikiChat: Stopping the Hallucination of Large Language Model Chatbots by Few-Shot Grounding on Wikipedia

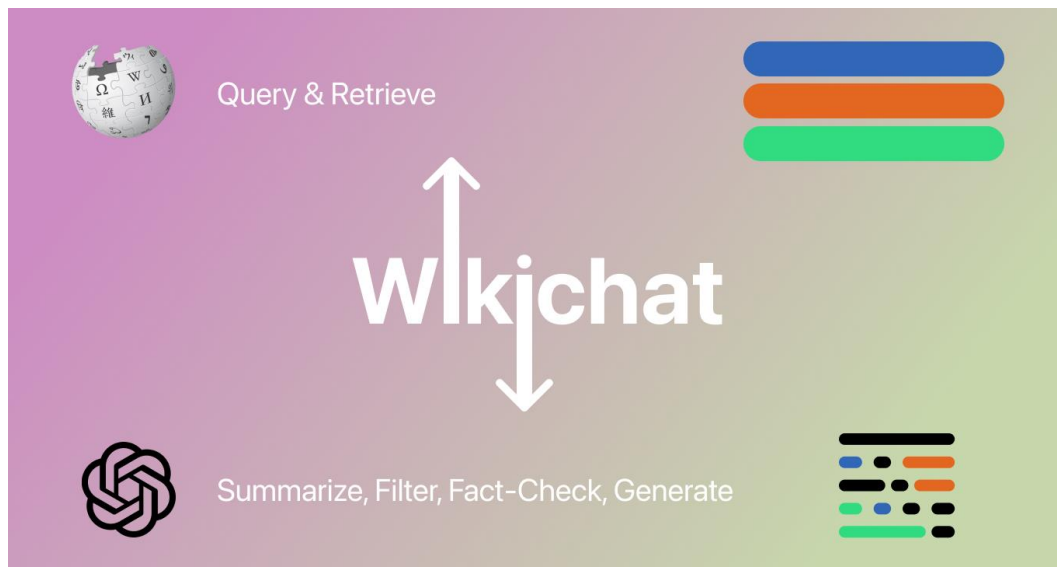
AWARDED TO

Sina J. Semnani, Violet Z. Yao, Heidi C. Zhang, and Monica S. Lam

Leila Zia
Research Award Committee co-chair
Wikimedia Foundation

Benjamin Mako Hill
Research Award Committee co-chair
University of Washington

Live Demo: <https://wikichat.genie.stanford.edu/>



Summary: Learning How to Read

- Exploit LLM capabilities
 - Generative power, general understanding of the world
 - NLP skills: summary, filter claim splitting, entailment
- But limit use of LLMs:
 - Fact-check LLM-result claim by claim
 - Short paragraphs only
 - Do not let it “answer questions” directly, or it will hallucinate
- The devil is in the details:
E.g. LLMs do not understand time (3 students, 3 months)

It is worth it! Reading without hallucination is key to knowledge processing

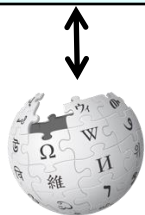
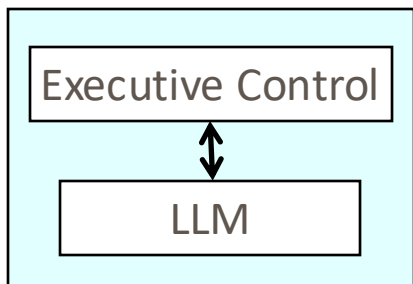
Agent Architecture

This Course

adds the “prefrontal cortex”

(executive control)

— cognitive (knowledge) skills



- Executive control:
A software program that implements the Reading Pipeline
- Calls LLM as NL subroutines
- Interfaces to external system components
e.g. Wikipedia

2. FOUNDATION FOR THE KNOWLEDGE REVOLUTION

Bloom's taxonomy: a hierarchy of objectives in the cognitive domain

LLM is analogous to our speech center, lacks executive control

Executive control: a software program to implement a cognitive skill
(controls the LLM as a subroutine)

3. CS224V COURSE

Design of CS224V

- **Focus: Tools for General Cognitive Skills**
To make building reliable assistants easily by non AI-experts
- **Approach: Application-driven research**
 1. Create a new LLM-based cognitive skill by modeling human processes
 - Implement as a tool
 - Experiment with an initial application
 2. Improve the tool by applying it to many different apps

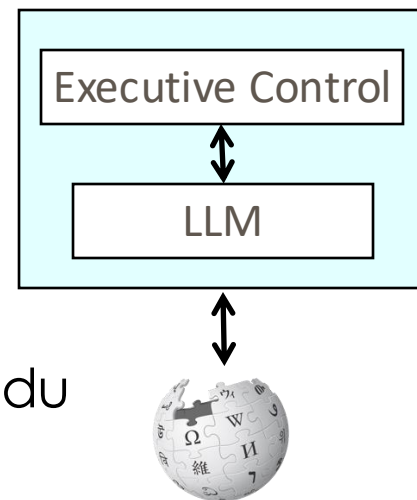
Course Objectives

Learn and advance the state of the art (SOTA)

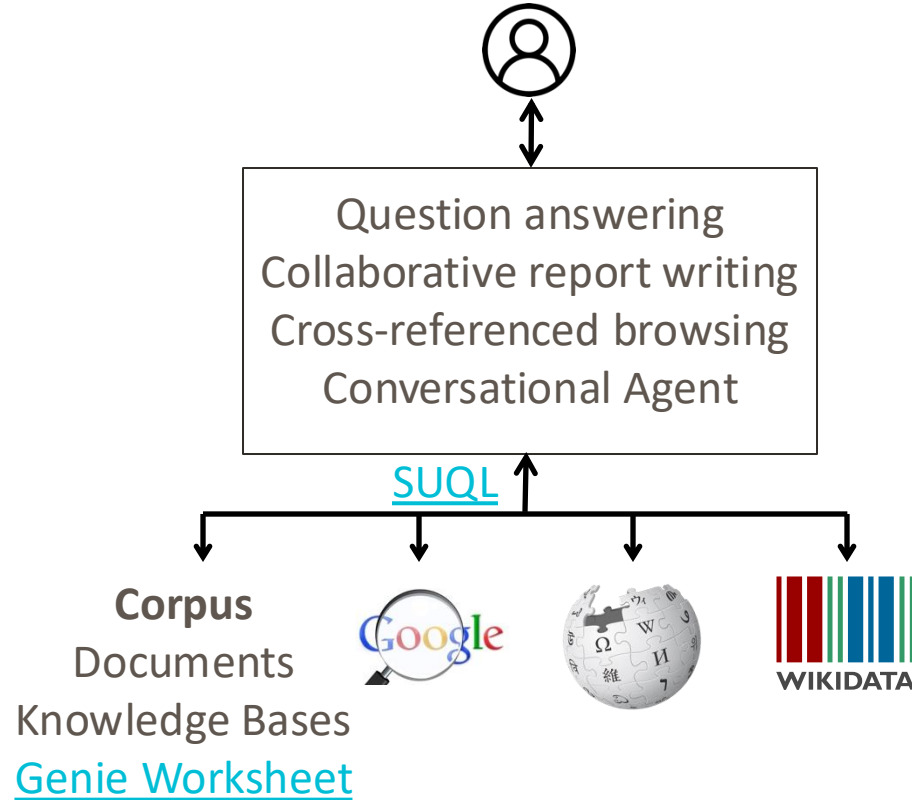
1. 2 Homeworks to bring everybody onboard with SOTA tools
2. Lectures on techniques of LLM-based conversational agents
3. Supervised quarter-long project
 - Develop new tools; enhanced existing tools (with app)
 - Create a new app and evaluate

CS224V in 2023

- WikiChat tool was available to the class
 - Homework in week 1: a non-hallucinating chatbot
- Result of the course
 1. 2 publications
 2. Ground work for 3 other publications
 3. Live demos on <https://oval.cs.Stanford.edu>



Tools Available for CS224V in 2024



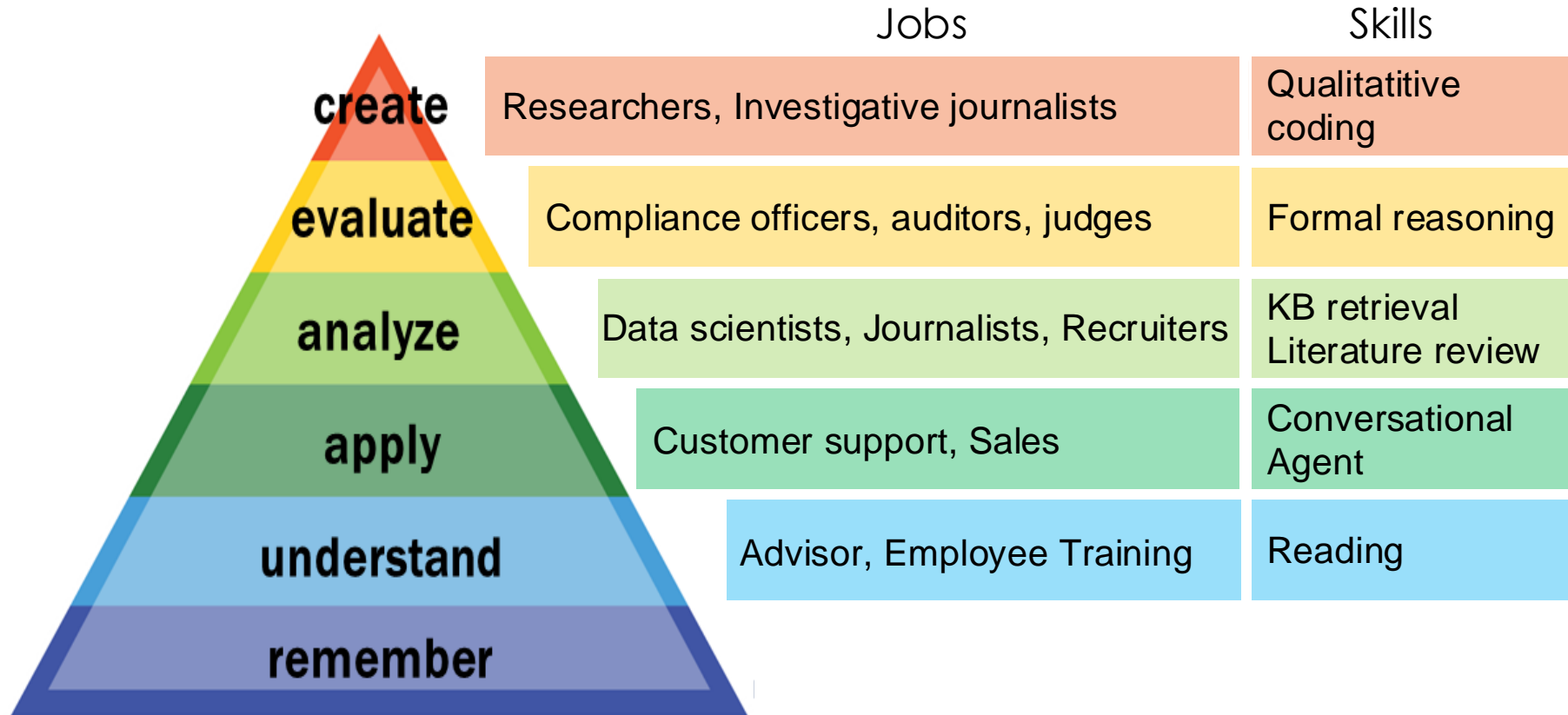
SUQL:

The first query language for structured & unstructured data.

Genie Worksheet:

The first specification language for task and knowledge bots.

Available Skills in the Bloom Taxonomy



Homework 1: Automatic Data Curation

Going beyond Q & A ...

- Storm: Writes a Wikipedia-like article from scratch
 - By researching the internet
- Co-Storm:
 - Add collaboration to Storm
- Your homework: Use co-Storm on ArXiv
 - Learn how to work with LLM-based pipelines
 - Discover issues in working with specialized domains

Release of STORM



- Released 7/10/2024
 - Without self promotion
- Organic YouTube videos
- First month traffic
 - 27K users
 - 50K articles
- General
 - Applicable to any domain

“Perplexity Pages alternative but it is insanely better.”



AI Code King

Demo at: <https://oval.cs.Stanford.edu>

Homework 2: Create an Agent

- The classical use of NLP is a conversational agent
- The Genie Worksheet lets you create a powerful agent with a small number of lines of code
 - Task-oriented: e.g. enroll in courses
 - Knowledge-oriented: e.g. answer questions about courses
- Homework2:
 - OfCourse is a course advisor prototype for the AI track of CS MS
 - Extend OfCourse to cover other programs in CS MS
 - Purpose: Exposure to tools for developers and learn their strengths/weaknesses in real life
 - Note: The practice is very different from typical NLP papers (80% accuracy is not usable)

Computational Biology

DEPARTMENT OF COMPUTER SCIENCE MSCS Program Sheet (2023-24) Computational Biology 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/mcscs/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

COMPUTATIONAL BIOLOGY 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:

- One of: CS 173A, 273A (students with equivalent coursework may waive with the approval of their advisor)
- CS221
- At least two of: CS 142 (or 147L or 193X), 145, 246, 448B
- At least three of: CS 279, 371, BIOMEDIN 210, 214, 215, 217, 219, 220, 222, 260, 273B, IMMUNOL 207
- A total of at least 21 units from categories (a), (b), (c), (d) and the following: CS 124, 131, 147, 148, 154, 166, 168, 185, 224N, 224W, 228, 229, 229B, 229S, 229T, 230, 231N, 234, 238, 248/248A, 353, 399; BIO 183, 187; STATS 215, 256

Advisor Notes

Course number	Title (Depth courses must be taken for at least 3 units.)	Grade	Units
TOTAL SPECIALIZATION UNITS APPLIED TO MSCS (must total at least 21 units) <i>Letter grades only.*</i>			<input type="text"/>

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			<input type="text"/>
TOTAL UNITS APPLIED TO MSCS			<input type="text"/>

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: _____

Computer and Network Security

DEPARTMENT OF COMPUTER SCIENCE MSCS Program Sheet (2023-24) Computer and Network Security 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 GSI 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, 3690; 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

COMPUTER AND NETWORK SECURITY DEPTH

All depth course 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:

- All of: CS 140/212 (or 140U or 112),† 144†, 155, 255, 356 (students with equivalent course work may waive † flagged courses with advisor approval)
- At least three of: CS 142, 152, 190, 240 (any suffix), 244, 244B, 249I, 253, 261, 265, 340 (any suffix), 355
- A total of at least 21 units from categories (a), (b) and the following: CS 245, 251, 294S‡, 399‡; EE 384S (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 SPECIALIZATION UNITS APPLIED TO MSCS (must total at least 21 units) Letter grades only.*			<input type="text"/>

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			<input type="text"/>

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: _____

Software Theory

DEPARTMENT OF COMPUTER SCIENCE MSCS Program Sheet (2023-24) Software Theory 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/mssc/programsheets/pguide2324.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, 181W, 182W, 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; INTPOL 268; LAW 4039; ME 177; MS&E 193, 231, 234, 254; POLISCI 150A; PUBLPOL 103F for 3 units, 353B

SOFTWARE THEORY DEPTH

All depth course must be taken for a letter grade* for 3 or more units. A maximum of 6 units of CS399 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:

Advisor Notes

- CS 242, 243
- At least one of: CS 221, 244, 245, 341
- At least one of: CS 255, 350, 355, 356
- A total of at least 21 units chosen from categories (a), (b), (c) or the following: CS 151, 250, 261, 265, 294S, 295, 315B, 349H, 357, 357S, 399T (courses with † notation require approval of MS advisor)

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

TOTAL SPECIALIZATION 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: _____

DEPARTMENT OF COMPUTER SCIENCE
MSCS Program Sheet (2023-24)
Systems 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/pguide2324.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 244E; INTLPOL 268; LAW 4039; ME 177; MS&E 193, 231, 234, 254; POLISCI 150A; PUBLPOL 103F for 3 units, 353B

SYSTEMS DEPTH

All depth course must be taken for a letter grade* for 3 or more units. A maximum of 6 units of CS399 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:

Advisor Notes

- All of: CS140/212 (or 140E or 112 or 112E) †, CS 144‡, 240
- At least four of: CS 190, 242, 243, 244, 245, 248/248A, 348B; EE 271, 282
- A total of at least 21 units from categories (a), (b) and the following: CS 149, 217, 229S, 241, 244B, 246, 251, 255, 270, 272, 294S†, 295, 315B, 316, 340 (any suffix), 343D, 344 (any suffix), 348A, 348C, 348E, 348I, 348K, 349 (any suffix and 3 or more units), 356, 357S, 399†, 448 (any suffix); EE 267, 273, 382C, 384A, 384C, 384S (courses with † notation require approval of MS advisor)

Course number	Title (Depth courses must be taken for at least 3 units.)	Grade	Units
TOTAL SPECIALIZATION 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: _____

Course Enrollment Agent

What course would you like to enroll in?

User:

I am a sophomore studying computer science. I want to complete the significant implementation requirement. What are some of the options that require the least amount of work?

A Course Enrollment Assistant

Course Enrollment Agent

What course would you like to enroll in?

Students need to consult:

Degree Requirements

- Unique to each department and specialization
- Program sheets

Stanford | Bulletin
ExploreCourses

- All offered courses
- Description, instructors, offerings, and units

CARTA

- Ratings, reviews, hours of work, enrollment outcomes, and sequencing

CS 144: Introduction to Computer Networking

Unpin from Winter

← SUMMER 23-24 →

Results for: CS 140

▼ AT A GLANCE

▼ EVALUATIONS

PINNED CLASSES
Pin a course to your calendar

QUARTER OVERVIEW

Mo	Tu	We	Th	Fr	Sa	Su
8 am						
9 am						
10 am						
11 am						
12 pm						
1 pm						
2 pm						
3 pm						
4 pm						

Pinned Units:
Weekly Time

Results for: CS 145

Units ▼ Days ▼ WAYS ▼ Quarters ▼

CS 145: Data Management and Data Systems

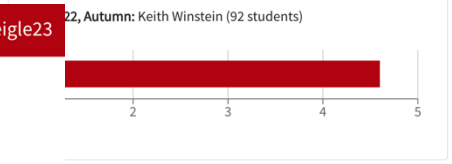
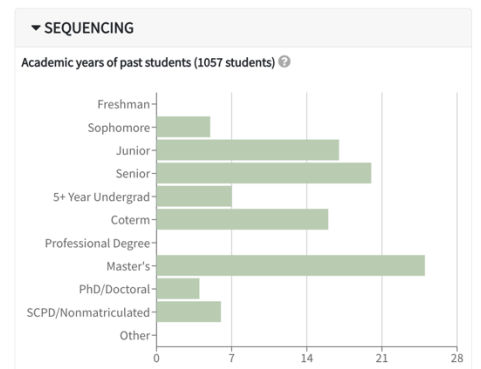
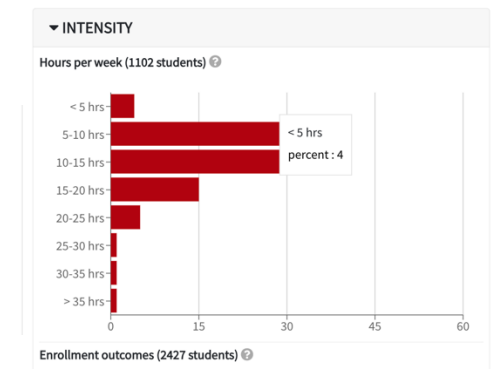
Introduction to the use, design, and implementation of database and data-intensive systems, including data models; schema design; data storage; query processing, query optimization, and cost estimation...[Read more](#)

[Show Schedule](#)

Term: AUT
Units: 3-4
Grading: Letter or Credit/No Credit
Instructors: Shivakumar, N.

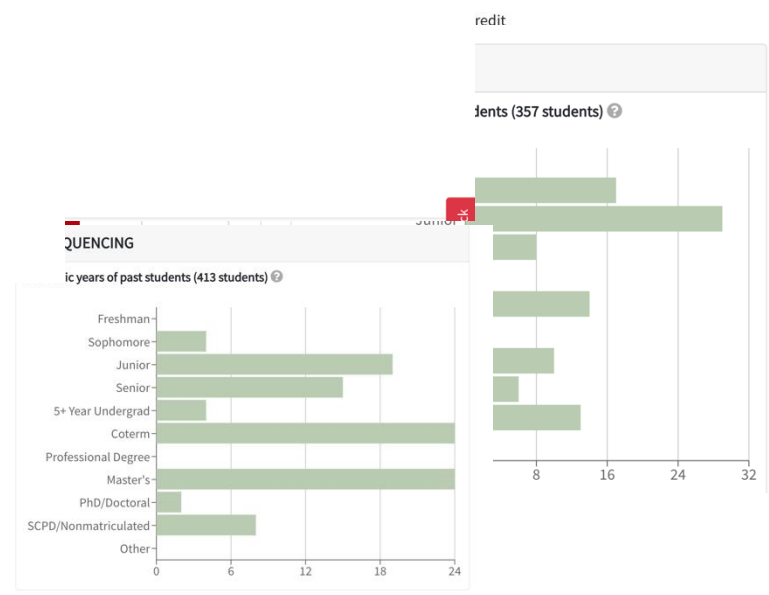
Unpin from Autumn

No more results



Advising weigle23

Pin to Spring



How to Create an Agent

Developers provide:

1. A worksheet specifying the needed info and conditions to satisfy
2. Knowledge bases containing relevant information

1. Stanford Course Enrollment Form


Form Name	Predicate	Kind	Type	Name	Enum Values	Description
Main			worksheet	course_enrollment		
		input	CourseToTake	course_to_take		The course to enroll
		input	StudentInfo	student_info_details		Information on the student
StudentInfo			worksheet			
		input	str	student_name		Name of the student
		input	str	student_id		Student's ID number
		input	str	student_email_address		Student's email address
CourseToTake			worksheet			
		input	str	course_name		Name of the course
		input	Enum	grade_type		The desired grading basis
					Credit/No Credit	
					Letter	
		input	int	course_num_units		The number of units taken
		input	confirm	confirm		Confirm the course

2. Course Assistant Knowledge Corpus

courses				db
internal; primary	int	course_id		
offerings				db
programs				db
ratings				db
	internal; primary	int	rating_id	
	internal	int	course_id	
	internal	List[str]	instructor_names	
	internal	int	average_rating	
	internal	int	num_ratings	
	internal	int	term_id	
	internal	int	start_year	
	internal	int	end_year	
	internal	Enum	season	
				autumn
				winter
				spring
				summer
	internal	List[str]	reviews	
	internal	str	sheet_requirements	

Translates Questions into SUQL

I am a sophomore studying computer science. I want to complete the significant implementation requirement. What are some of the options that require the least amount of work?

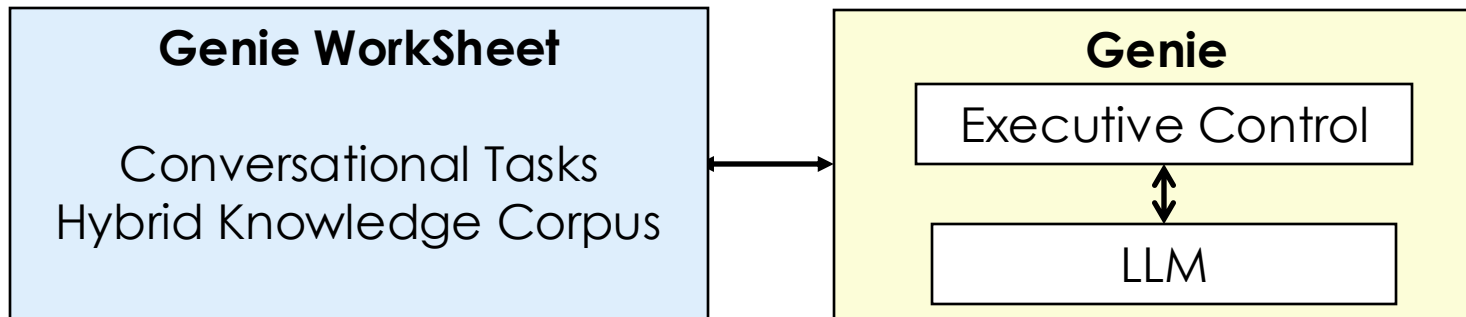


```
SELECT title, course_codes, description,  
average_hours_spent  
FROM courses WHERE  
significant_implementation_requirement = TRUE  
ORDER BY average_hours_spent LIMIT 5;
```

Executing a Worksheet

Executive Control Software

- Interpret & update Genie Worksheet
- Provide turn-specific instruction to LLM
(Succinct context: Worksheet state + 1 turn of dialogue)





You 01:01 AM

Hi, I want to complete the significant implementation requirement. What courses would you suggest me to take?

Took 3 steps ▾



Course Enrollment Assistant 01:01 AM


I found some courses that fulfill the significant implementation requirement. Here are a few suggestions along with their average hours spent per week:

1. **General Game Playing (CS 227B)**: A general game playing system accepts a formal description of a game to play it without human intervention or algorithms designed for specific games. Hands-on introduction to these systems and artificial intelligence techniques such as knowledge representation, reasoning, learning, and rational behavior. Students create GGP systems to compete with each other and in external competitions. Prerequisite: programming experience. Recommended: 103 or equivalent. *Average hours spent per week: 7.48.*
2. **Logic Programming (CS 151)**: Logic Programming is a style of programming based on symbolic logic. In writing a logic program, the programmer describes the application area of the program (as a set of logical sentences) without reference to the internal data structures or operations of the system executing the program. This course introduces basic logic programming theory, current technology, and examples of common applications. Prerequisite: CS 106B or equivalent. *Average hours spent per week: 7.5.*
3. **Introduction to Computer Graphics and Imaging (CS 148)**: This introductory course in the computer graphics sequence introduces students to the technical concepts behind creating computer-generated images. Topics include the mathematical concepts of synthetic imagery, fundamentals of light and color, rasterization, and ray tracing technology. Prerequisite: Linear Algebra and familiarity with Python and debugging. *Average hours spent per week: 8.79.*

Could you please provide me with the details of the courses you want to take?




Consideration: Time Offered

 **You** 01:02 AM

Great when are these courses offered?

Took 3 steps ▾

 **Course Enrollment Assistant** 01:02 AM


The courses you're interested in are offered at the following times:

1. **Introduction to Computer Graphics and Imaging (CS 148)** is offered in the **autumn** season on **Tuesday and Thursday** from **12:00 PM to 1:20 PM**.
2. **Logic Programming (CS 227B)** is offered in the **spring** season on **Tuesday and Thursday** from **1:30 PM to 2:50 PM**.
3. **General Game Playing (CS 151)** is offered in the **spring** season on **Wednesday** from **4:30 PM to 6:20 PM**.

Could you please provide me with the details of the courses you would like to take?



Consideration: Sort by Course Rating

 **You** 01:02 AM

Can you sort them based on rating?

Took 3 steps ▾


 **Course Enrollment Assistant** 01:02 AM

Certainly! Based on their ratings, here are the courses sorted:

1. **Introduction to Computer Graphics and Imaging** with an average rating of 4.04.
2. **Logic Programming** with an average rating of 3.97.
3. **General Game Playing** with an average rating of 3.43.

Could you please provide me with the details of the courses you would like to take?



 **You** 01:03 AM

I will take the top rated one!

Took 3 steps ▾

Credit or Letter? Depends on the Workload



Course Enrollment Assistant 01:03 AM

Great choice! For the course "Introduction to Computer Graphics and Imaging", could you please specify the desired grading basis? The options are: Credit/No Credit or Letter.



You 01:03 AM

How many hours will I have to spend on it?

Took 3 steps ▾



Course Enrollment Assistant 01:04 AM

On average, you will have to spend about 8.79 hours per week on the course "Introduction to Computer Graphics and Imaging". Now, could you please specify the desired grading basis for this course? The options are: Credit/No Credit or Letter.



Pick Max. Number of Units

Y You 01:04 AM

Thats low! I will take for letter grade, whats the maximum units I can take it for?

Took 3 steps ▾

C Course Enrollment Assistant 01:04 AM

The maximum number of units you can take for the course "Introduction to Computer Graphics and Imaging" is 4 units. How many units would you like to enroll for this course?



Y You 01:04 AM

Lets go with 4

Took 3 steps ▾

Purpose of the Assignments

- Prepare you for your project proposal
 - How to use LLMs to curate data from the internet
 - How to create a basic task & knowledge bot as a developer
- Lectures on underlying technology and advanced topics to follow

Project Apprenticeship

- Assistance with project selection: Hardest part in research!
 - We suggest over 20 projects on the website
 - Some with domain experts in journalism, history, medicine, finance, gaming, etc.
 - Student-initiated projects are also welcome
- Weekly group mentorship meeting
 - We want to make you succeed!

Project Mentorship

All homeworks and projects are done in groups of 2

- Week 4: Project proposal, with a weekly plan
- Weeks 5-10 (excluding Thanksgiving break):
 - Submit a written weekend update (every Monday)
 - Group meeting with mentors during the week
- Week 11: Poster presentation (Dec 4)
- Final project report due Dec 10, 2024.

Course Schedule at a Glance

Dates	Lectures / Homeworks	Projects
9/23 - 10/7	Introduction; Knowledge curation (HW1) Building a task-oriented agent (HW2) Grounding on small databases	Research Project Ideas
10/9 - 10/21		Student-initiated ideas Project discussions Project proposals (2)
10/21 - 11/20	Grounding on free text SUQL; Task-oriented agent generation Knowledge graph queries; Knowledge discovery Formal reasoning; Multimodal apps NLP building blocks; Training LLMs	
11/25 - 11/27	<i>Thanksgiving</i>	
12/2 - 12/4	No Class	Final project posters (3:00-5:40)

This Course

	Grade
Participation	15%
Assignment	25%
Final Project	60%

Participation includes

- Class attendance and participation
- Ed discussion
- Meetings with project mentors