Types of CS research

Honors seminar

Types of CS research

Computer science is...

"...the study of algorithmic processes that describe and transform information"

"...the study of phenomena related to computers"

"...the study of information structures"

"...the study and management of complexity"

"...the mechanization of abstraction"

"...a field of study concerned with [...] the development and use of computers for information storage and processing, mathematics, logic, science, and many other areas"

Engineering?

	Science	Engineering
Goal	Knowing the general	Knowing the concrete
Activity	Develop/test theories	Building artifacts
Method	Abstraction	Embracing complexity
Process	Conceptualization	Optimization
Innovation	Discovery	Invention
Take-away message	Law-like	Rule-like
Time perspective	Long-term	Short-term

How do I science?

Science is…

- an attempt to provide an ordered description and explanation of the world
- based on experience and observation
- self-correcting

Replicable & open

Goal of science:

Discover new and useful information...

...in the form of verifiable data

Findings which can be repeatedly ascertained by other investigators using similar methods

To this effect, scientists should make their methods public, and their data available to others

Many cases of scientific fraud suspected from secrecy and confirmed through openness

Values vs. facts

Subjective statements
(cannot be shown to be true or false)

- Handguns should be outlawed
- People should wear masks
- Blue is beautiful

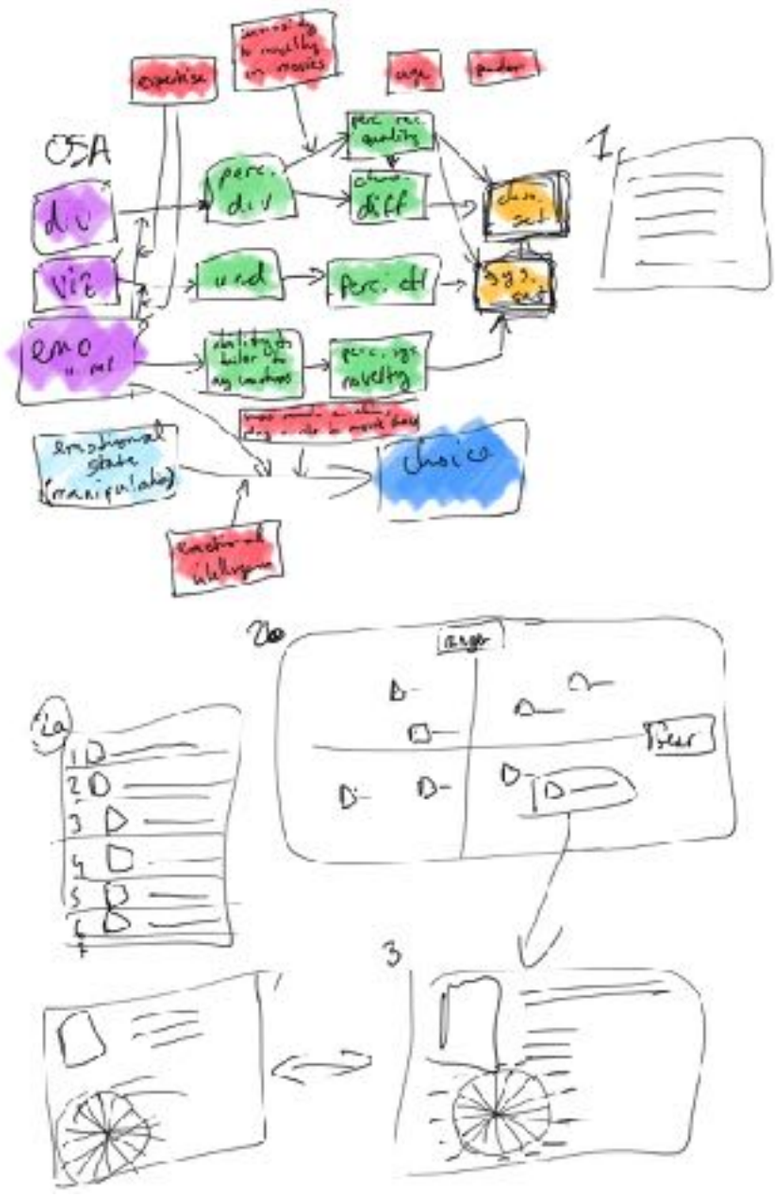
Objective statements
(can be shown to be true or false)

- Handguns caused over 30k US deaths in 2002
- Masks significantly reduce the spread of COVID-19
- 35% of college students say blue is their favorite color

Scientific process

Ideal world:

- exploration (observation)
- theory (hypothesis)
- testing (further observation or experimentation)
- accepted theory (evidence)



Scientific process

Real world:

- exploration
- theory
- testing
- **completely unexpected results**
- interpretation
- revision
- new theory (etc.)

Keep in mind...

Scientists are human

- they can become very attached to theories
- they bring their own opinions to the interpretation of data
- they rarely change their basic approach or paradigm
- they can be swayed by authority

Science is influenced by social and political pressures

- funding goes to issues deemed more "important"
- popular theories are aligned with societal values

Research

What is CS research?

Research

CS research methodologies

- Formal ("theoretical CS"... turns CS into math)
- Experimental (ideal for studying humans and data)
- Build ("case studies" to go beyond engineering)
- Process (anything to do with human subjects)
- Model (abstraction of the real world)

Research

Case studies (scientific way to use the "build" methodology):

- Feasibility study (Can X be built? Lessons learned)
- Pilot case (Does X work in situation Z? Check criteria)
- Comparative study (Which is better, X or Y? Test criteria)
- Observational (How does X fare in the real world? Observations, subjective feedback)

The "case" should be a real-world example (not a toy or exemplar)

It can be a "challenge" or a benchmark

Research

Combinations:

- Formal + process = verification, proof of correctness
- Process + experiment = HCI study
- Process + model = cognitive modeling, system modeling
- Process + observational = ethnographic study
- Model + experiment = simulation

Research questions

What are they and where do they come from?

Research questions

A research question defines the objective of a study

Benefits for you, before doing the research:

- It distills your research topic area into a specific goal
- It motivates the research methodology
- It drives the development of hypotheses and/or criteria

Benefits for you, after the research:

- It motivates the analysis of collected data
- It scopes the argumentation of your paper

Research questions

A research question defines the objective of a study

Benefits for others, before the study:

- It explains what question you set out to answer in a study proposal (to IRB, grant reviewers, potential collaborators)

Benefits for others, after the study:

- It explains to readers what they will learn from your paper
- It explains to listeners what they will learn from your presentation

Research questions

Science = “stand on the shoulders of giants”

Goal: build an organized body of knowledge about how people interact with technology

A good research question has the potential to expand the existing body of knowledge

- It must be grounded in the existing knowledge base
- It must be researchable
- It must be important

Hypotheses

Research questions:	**Hypotheses:**
Questions	Statements
Usually open-ended	Supported or rejected
General (each paper usually has 1–3 research questions)	Specific (each research question can lead to multiple hypotheses)
Describes broad relationships	Describes relationships between specific variables

Example

Research question:

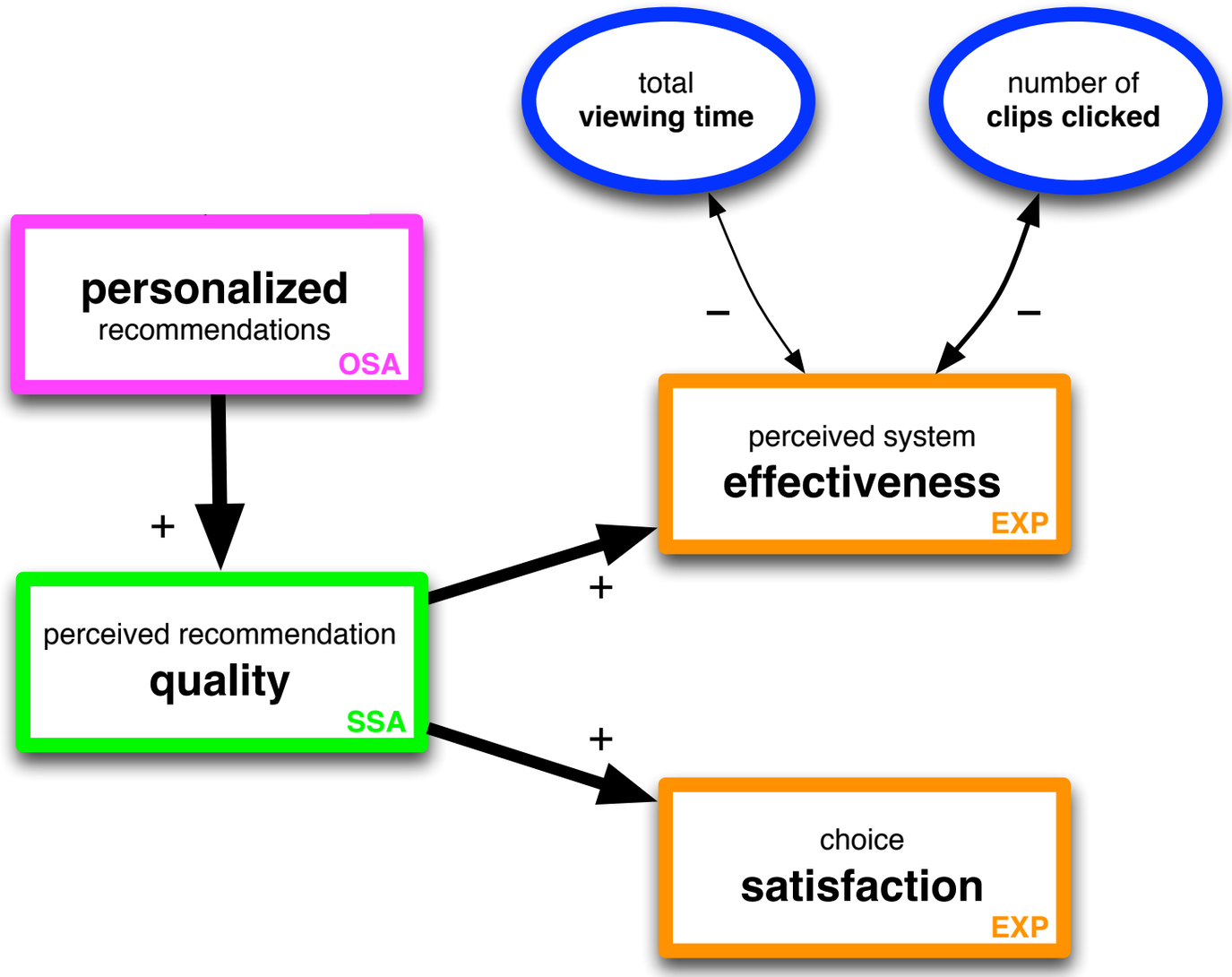
How come personalized video clip recommendations **reduce** users' overall viewing behavior (compared to random clips)?

Example

Hypotheses:

- Users perceive personalized recommendations to have a higher quality than random recommendations
- Recommendation quality is positively associated with perceived system effectiveness
- Recommendation quality is positively associated with choice satisfaction
- Perceived system effectiveness is **negatively** related to users' total viewing time
- Perceived system effectiveness is **negatively** related to the number of clips people watch

Example


total
viewing time
number of
clips clicked
personalized
recommendations
OSA
perceived system
effectiveness
EXP
perceived recommendation
quality
SSA
choice
satisfaction
EXP
+
+
+
–
–

Source

Research topic area

Specific research question(s)

Topic area

Topic areas can come from anywhere!

- Your curiosity
- Your advisor's grant topic
- "Hot" topics
- Really anywhere!
  - Seminars, conversations with others, newspapers, podcasts

Topic area

Valid topic areas depend on your CS subfield. E.g, in HCI:

- A (feature of) technology that fulfills a certain human or societal need (e.g. diversity in video games to reduce bias)
- A situation where technology creates a new social dynamic or a social/societal problem (e.g., defaults and framing lead to worse privacy decisions)
- A new way of interacting with the digital world (e.g. studying the effects of virtual agents)
- A situation that could benefit from (a fundamental rethinking of) digital assistance (e.g. recommender systems for self-actualization)

Research questions

They can come from:

- Again, your curiosity
- Practical problems (evaluate an assumption or new technology)
- Literature (previous studies or theory)
- Requests for proposals

Prior research

Initial work

Pilot studies, case studies, studies with a small N

Questionable assumptions

E.g.: “More sharing options is better for privacy”

Conflicting findings

E.g.: Facebook use and depression

Overlooked variables

E.g.: The moderating effect of culture

Proposal requests

Advanced Distributed Learning

BAA Update

Topic: Activity Registry and Common Course Catalog

Across the Department of Defense (DoD), training and education systems historically use pre-determined, point-to-point connections and manual intervention to share actionable information relating to requirements, curricula, and student performance across systems. This requires integrators to customize each system to get the data out of one and into another, and human processes to analyzes and correlate it. This is especially true with the way we catalog and archive different types of instructional activities including digital content, instructor-led courses, simulations, and other learning opportunities. This is complicated in that a single entity does not own all training and education content and therefore can only publish catalogs (e.g. CANTRAC) for the courses they own.

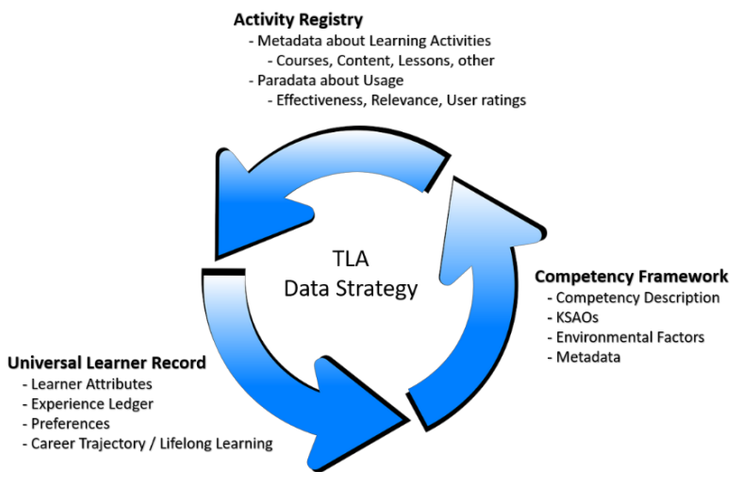


Figure 1. *The TLA Data Strategy aligns ongoing modernization efforts across the DoD's training and education enterprise*

In the face of numerous, varied, and complex information sharing challenges facing the DoD, the DoD Chief Information Officer (CIO) has set a vision to deliver an Information Enterprise (IE) that enables DoD and partners to securely access data and information services they need at the right time and place, and on approved devices of their choosing. The DoD Net-Centric Data Strategy provides the basis for implementing and sharing data in a net-centric environment. It describes the requirements for inputting and sharing data, but it does not accommodate the myriad of complexities associated with today's training and education environment. The Total Learning Architecture (TLA) data strategy summarized in Figure 1, along with the related common course catalog from the DoD Chief Management Officer, build upon the DoD Net-Centric Data Strategy's goals of making data assets visible, accessible, secure, and understandable. This approach establishes services as the preferred means by which information producers and capability providers can make their data assets and associated operational activities available across DoD and beyond. It also establishes services (i.e. data contracts and manipulation business logic) as the preferred means by which consumers can access and use these data assets and capabilities.

Proposal requests

As envisioned, this research will focus on identifying, enhancing, or developing the standards and specifications necessary to create a federated data framework that allows individual DoD components to manage their own learning resources via an Activity Registry that rolls up (i.e. federates) into a Common Course Catalog. The key attributes of this research include:

- Ensure data about learning resources is visible, secure, available, and usable when needed and where needed to accelerate decision-making. Predict and inventory requirements for learning activities (e.g., courses) to meet the demands of DoD consumers while minimizing the duplication of course development efforts.
- Align with the DoD Discovery Metadata Specification (DDMS) which specifies how to use metadata to make data assets visible across the enterprise.
- Automate the ability to align content with competencies, generate and manage paradata, perform semantic search services, and create machine-actionable metadata that roll up from local Activity Registries into a single Common Course Catalog.
- Align ongoing initiatives from the DoD CIO, CMO, and the Chief Data Officer with the needs of the training and education community, while providing guidance to ongoing training and education modernization efforts across DoD.

The ADL Initiative invites white paper proposals to address these issues through the following tasks:

- **Task 1: Definition/Validation of the learning activity metadata elements including content, courses, exercises, and other formal learning activities common across the DoD.** Identify an approach for describing various learning activities a learner will encounter across the continuum of lifelong learning. Investigate various metadata formats, standards and specifications to inform an approach that describes content in support of the Future Learning Ecosystem. Create an approach that considers machine-learning created metadata (for example, relevance and engagement) beyond the more commonly available human-specified attributes. This task includes consideration of metadata storage and retrieval methods across the myriad of T&E commands within DoD (e.g., Activity Indices). The deliverables would include both Metadata Strategy and Metadata Guidance.

- **Task 2: Definition/Validation of learning activity paradata elements (data about usage)**. Identify a taxonomy and approach for gathering, verifying, storing, and sharing paradata within an Activity Registry. While metadata classifies properties about the learning resource, paradata involves the collection and open sharing of in situ information about users' actions, preferences or experiences related to the resource. Paradata is generated through user processes of searching for

Proposal requests

“Homework”

Research the SoC faculty

- What research do they do?
- Example papers, presentations, etc.

Pick 4-5 faculty members you’d like to work with

- What type of research would you like to work on with them?

Create 4-5 slides (one for each faculty member)

- Present the information listed above

Questions…