Exam Revision 1

INFR11158/11230 Usable Security and Privacy

Dr. Jingjie Li

01/04/2025



Exam Structure

- Three questions: you must do Question 1 and select either Question 2 or Question 3 to answer
- "NOTES PERMITTED, CALCULATORS NOT PERMITTED examination. Candidates may consult up to THREE A4 pages (6 sides) of notes. CALCULATORS MAY NOT BE USED IN THIS EXAMINATION"
- Past exams online: https://exampapers.ed.ac.uk/

Expectation

- Applying concepts and frameworks learned in the lecture
- Thinking and analyzing critically using logic and examples
 - What are the limitations/tradeoffs?
 - What are the experiment tasks and materials?
 - Any similar cases?
 - ...
- No statistics, calculation, and drawing tested in the exam

Topics

- USEC basics
- Study method and analysis
- Authentication
- Online fraud and phishing
- Security and privacy communication (warning, advice, etc.)
- Privacy framework, tools, and policy
- Ethics and consent
- Access control, vulnerability research, AI, IoT, at-risk users....
- Other coursework-related topics (framework application, dark patterns, etc.)

USEC Intro

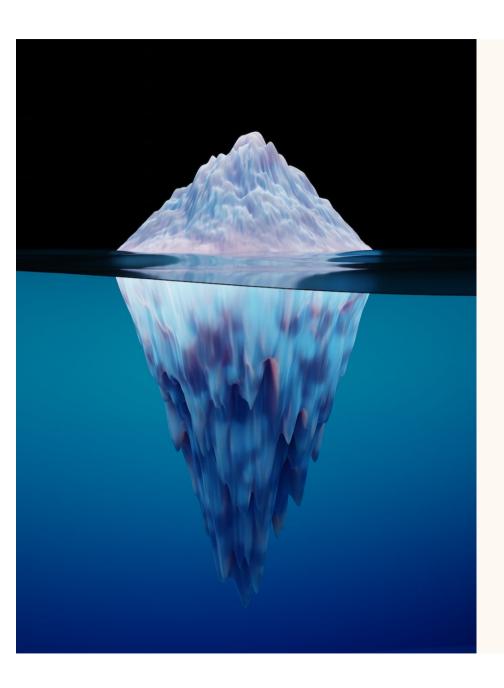
Defining security – CIA definition

Confidentiality	No improper information gathering
Integrity	Data has not been (maliciously) altered
Availability	Data/services can be accessed as desired
Accountability	Actions are traceable to those responsible
Authentication	User or data origin accurately identifiable

Usability and human factors

- **Learn-ability** The type for typical users to learn the actions relevant to a set of tasks.
- **Efficiency** How long it takes users to perform typical tasks.
- **Errors** The rate of errors users make when performing tasks.
- Memorability How users can retain their knowledge of the system over time.
- Subjective satisfaction How users like the various aspects of the system.





USEC is challenging because

- Interdisciplinary
- Seemingly familiarity
- Interrelations
- User evaluation
- Ecological validity
- Adversary model
- Technology velocity
- Customer

Threat Modelling: Adversaries

- Malicious actors
 - Hacker
 - Users (your family, your friend, your customer, etc.)
- Service providers
 - Company
 - App developers
- "Big brother"
- ... (depending on your position)

Assets

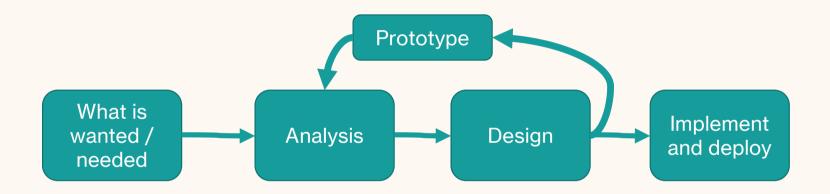
- Computer hardware: phone, laptop, server...
- Computer software: apps, operating systems, database...
- Physical assets: house, car.....
- Information: health record, your profile/identity, business info...
- Emotion, reputation, user experience....

Risk, threat and vulnerability

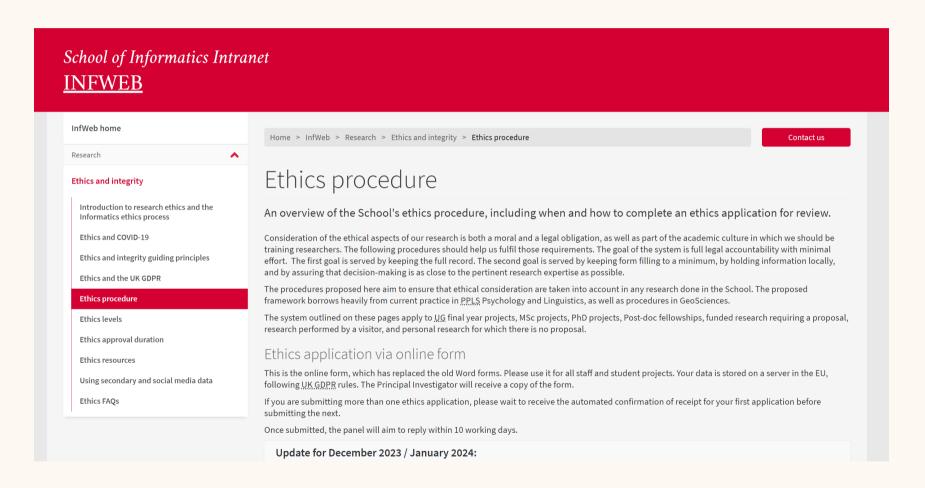
- Vulnerability: the weakness of X (system/human) that can be exploited
 - The program is overprivileged to access things
 - The user reuses their password across applications
- Threat is an action performed by the adversary to damage the asset by exploiting a vulnerability
- Risk = asset X threat X vulnerability

Study and Analysis Methods

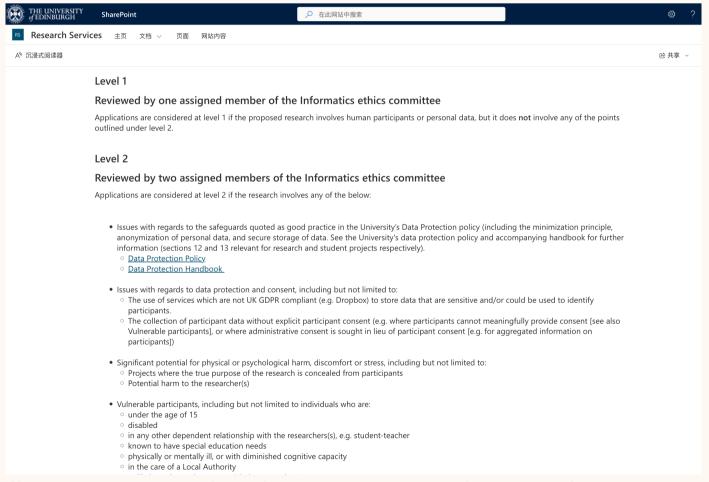
Project lifecycle



Ethics guidelines

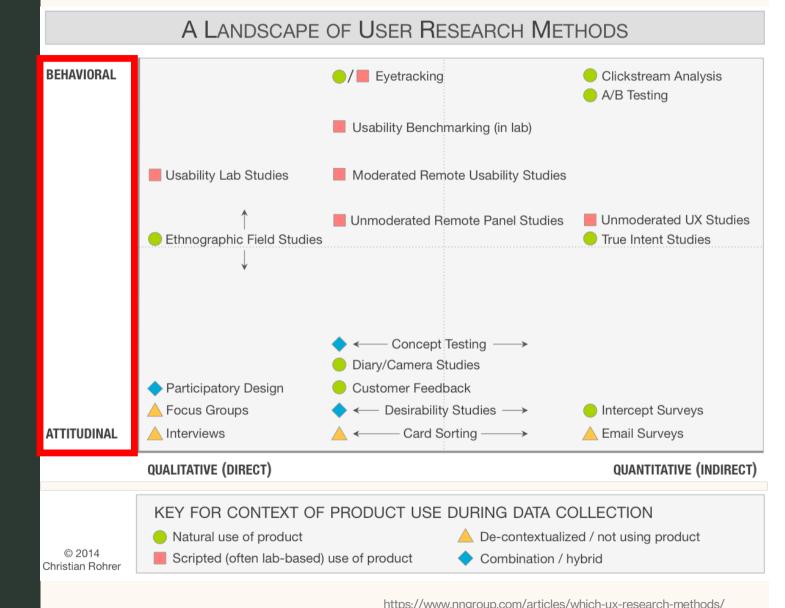


Ethics guidelines



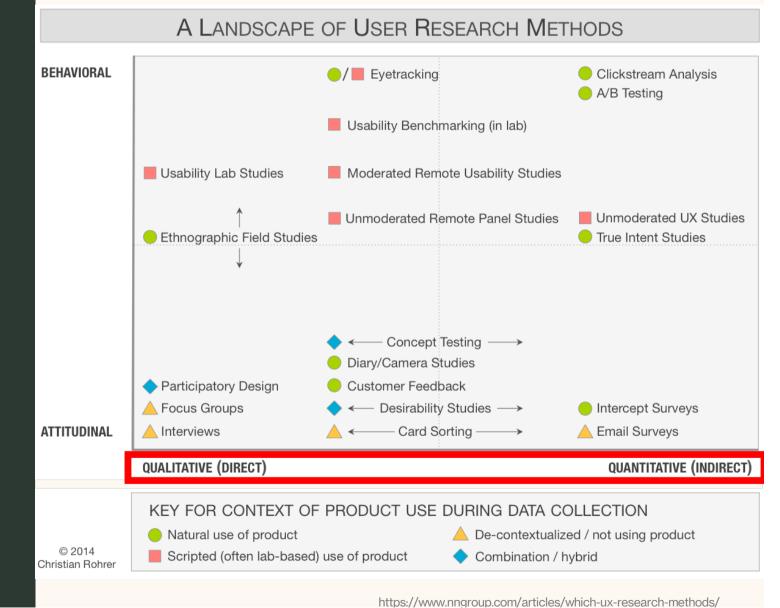
Behavioral – measures how people actually behave, what they do.

Attitudinal – measures what people say they think or how they say they behave.



Qualitative – unstructured data such as natural language.

Quantitative – numerical data. Anything that can be counted or measured with numbers.



Lab Study

- Basic idea: Have a participant come to a physical place (lab) and interact with the interface there
- You setup the lab so it mimics the situation you want to test
- Pros
 - Full control over the environment so limited confounds
 - Detailed data from each subject
 - Ability to ask them why they did something
- Cons
 - Small sample sizes
 - Being in the lab changes user behavior. They feel safer and their normal distractions are gone. That can be bad for deception studies.

Think aloud

- Basic idea: Have a participant use the interface and speak aloud while they do so
- Think aloud is a very versatile, can be long or short, detailed or minimal, planned or ad-hoc
- Pros
 - Learn what the user is trying to do and why they click on some things
 - Very detailed information
 - Testing with about 5 users will find the majority of major (usability) issues
- Cons
 - Biasing user behavior, making the situation unatural
 - (Concurrent) Talking aloud changes how long a user spends on tasks so this method cannot be combined with timing

https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/

Planning a survey

- Surveys normally answer multiple research questions. With each research question tied to one or more survey questions.
- **Descriptive** learn something about the whole population.
 - How many people have heard of the term "phishing"?
 - What words do people use to describe cookie tracking?
- **Testing for correlation or causation** show that two things are related or one thing causes the other thing.
 - If someone has been trained on phishing in the past, are they better at differentiating phishing emails?
 - We have three training options, each user goes through one training, which training causes people to identify phishing emails the best?

Survey scales

 Basic idea: A set of questions that have been previously shown to measure a property.

Pros

- Easy to copy-and-paste into a survey.
- Allows you to measure hard-to-measure concepts like risk seeking behavior or attitude towards privacy.

Cons

- Making a new scale is very challenging.
- Can contain an annoyingly large number of questions.

Testing: Correlation vs. Causation

Correlation

- Two things tend to behave in a way that seems inter-related, where if one thing changes the other thing will also change in a related way.
- For example, if the price of rice goes up at the same time as the price for beans.

Causation

- · When one thing changes it causes the other thing to change.
- For example, when the weather gets cold more people wear coats.
 Cold weather causes more people to wear coats.

Testing: What are you going to measure?

- In statistics there are classically two types of measurements (variables): dependent and independent
- Dependent
 - Also known as the outcome variable
 - "Dependent" on the study
 - Measures the usability goal
- Independent
 - Anything you are directly manipulating
 - An element of the study which is under your control
 - A pre-existing feature of your participant

Testing: Between vs. Within subjects

- Between subjects
 - Your study only shows one interface to one person
 - You are measuring how well the people randomly assigned to the A interface did compared to the people randomly assigned to the B interface
 - Lots of variability with this method
- Within subjects
 - Your study shows all interfaces to all people
 - You are measuring the difference in how they do on the two interfaces
 - Less variability (same person) but more learning effects and priming

Testing: Types of data

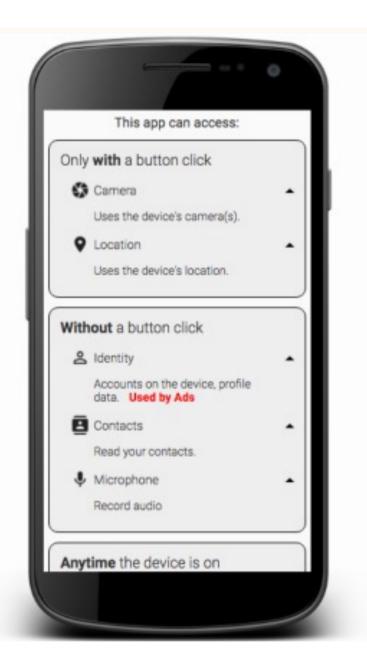
- Numeric
 - Continuous Any value on the range is possible including decimal (1-5)
 - **Discrete** Only certain values on the range are possible (1,2,3,4,5)
 - Interval Only certain values on the rage are possible and each has equal distance from its neighboring values (strongly agree, agree, neutral, disagree, strongly disagree)
- Categorical
 - Binary Only two possibilities (true, false)
 - Ordinal The values have an ordering (slow, medium, fast)
 - Nominal The values have no ordering (apple, pear, kiwi, banana)

Some research questions:

- Can people differentiate between a subdomain and a domain when reading a URL?
- Does [my new system] help people differentiate between malicious URLs and safe ones?
- Can users use [my new password manager] faster and with less errors than [the old password manager]?
- Does knowing how an app will use its permissions impact app installation decisions?
- What factors impact end-users' willingness to update software?
- Using [website], can users successfully opt-out of cookie tracking without forming inaccurate mental models?

Study design

- RQ: Does [my new interface] enable people to accurately determine what permissions an app will use?
- A/B test between the existing and new interface
- Between subjects
- 10 Tasks shown in the same order to all participants
- Dependent variables
 - Accuracy on task
- Independent variables
 - Which interface (A or B)

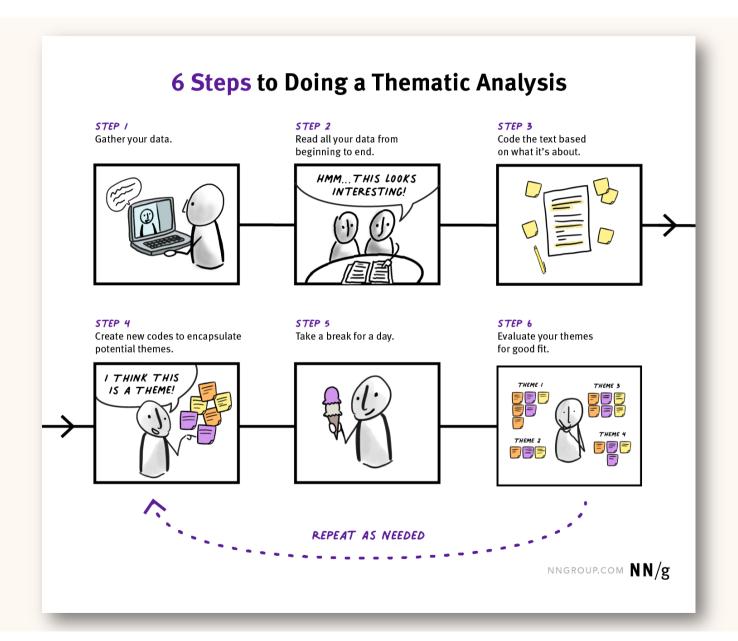


Inductive coding vs deductive coding

- Inductive (bottom-up): look for any ideas that interest you from different aspects
 - Snapshot of an app on a phone
 - Child playing with dog
 - Edited picture
 - Motion detection enabled

. . . .

- **Deductive (top-down):** start with some hypothesis
 - Children being monitored by app (privacy concern)
 - Camera placed in the living room (place of the scene)



Framework and Topics (next lecture)