Becoming and being a researcher

what I wish someone would have told me when I started doing research

Carlo Ghezzi
Politecnico di Milano

ECSA, Paris, September 2019

Outline

- Why do we do research?
- What is the role of research in society?
- Different kinds of research
- What is a *good* research?
- The role of research diffusion — publication and beyond
- The role of evaluation for research and researchers
- The role of ethics

Why do we do research? (1)

- Scientific research produces **new** knowledge
  - scientia (Latin) means knowledge
    - irrepressible aspiration to discover new knowledge is what makes humans different from other living entities
- Research empowers human society

Why do we do research? (2)

- Research empowers human society
  - The last 500 years witnessed a phenomenal and unprecedented growth in human power, mainly driven by scientific research — **Scientific Revolution**

- *During the last five centuries, humans increasingly came to believe that they could increase their capabilities by investing in scientific research. This wasn't just blind faith–it was repeatedly proven empirically. The more proofs there were, the more resources wealthy people and governments were willing to put into science.*

Sapiens–A Brief History of Human Kind (part 4, The Scientific Revolution) Y. N.Harari

Research driver of innovation

- The triple helix model of innovation
    - three main actors **academia**, **industry** and **government** interact to foster economic and social development
        - PhD students working with industry
        - PhD students producing artifacts/patents
        - PhD students involved in start-ups/spin-offs
        - PhD students moving to industry and transferring advanced knowledge

https://en.wikipedia.org/wiki/Triple_helix_model_of_innovation

Research is plural, multi-dimensional

Kinds of science/research

- **Basic** (or **Pure**) science/research as opposed to **applied** science/research
  - Engineering, like medicine, considered to be an applied science
- **Technology** refers to the application of science in a particular area for practical uses
  - comes from Greek, “techne” (skill)
- Distinction between science and the application of science often fuzzy and even snobbish
- Continuous spectrum of activities
- The notions of ***risk*** and ***gain***

Research paradigms

- **Formal research**
  - studies formal systems (math, theoretical CS)
- **Analytic research**
  - observes the world and builds theories that explain it and predict (typical of physics)
- **Empirical/experimental) research**
  - typically complements analytic research
- **Constructive research**
  - builds artifacts
- They may and often do co-exist in a given research

Informatics can be all of them

Essential attributes of research results

- **Originality**

Essential attributes of research results

- **Originality**
  - Outcomes/approach of the research are novel
    - specific new findings can be
      - new ways of thinking about a subject
      - new/improved solutions
- Different degrees of novelty
  - e.g., survey research

Essential attributes of research results

- **Originality**
- **Significance**

Essential attributes of research results

- **Originality**

- **Significance**
  - Outcome (potentially) exerts an influence on society, industry, or other research

- Multi-faceted concept
  - it solves an important open problem
  - solution was especially challenging
  - significant practical improvement in the way certain things used to be done
  - used by users

- From purely theoretical to very practical

Essential attributes of research results

- **Originality**
- **Significance**
- **Rigor**

Essential attributes of research results

- **Originality**
- **Significance**
- **Rigor**

- Refers to the intellectual integrity of the research process and the way the results are demonstrated
- Intellectual precision, robustness, and appropriateness of the concepts and methods adopted within the outcome
- Refers to the way experimental data are collected and analyzed, the rigor of the mathematical apparatus, …

Essential attributes of research results

- **Originality**
- **Significance**
- **Rigor**

- quality of research defined by degree of OSR
- S most difficult to assess

How to achieve results?

Two essential skills

Creativity/
Ingenuity

Rigorous, systematic
methods

description and
validation of
research results

Absolute truths?

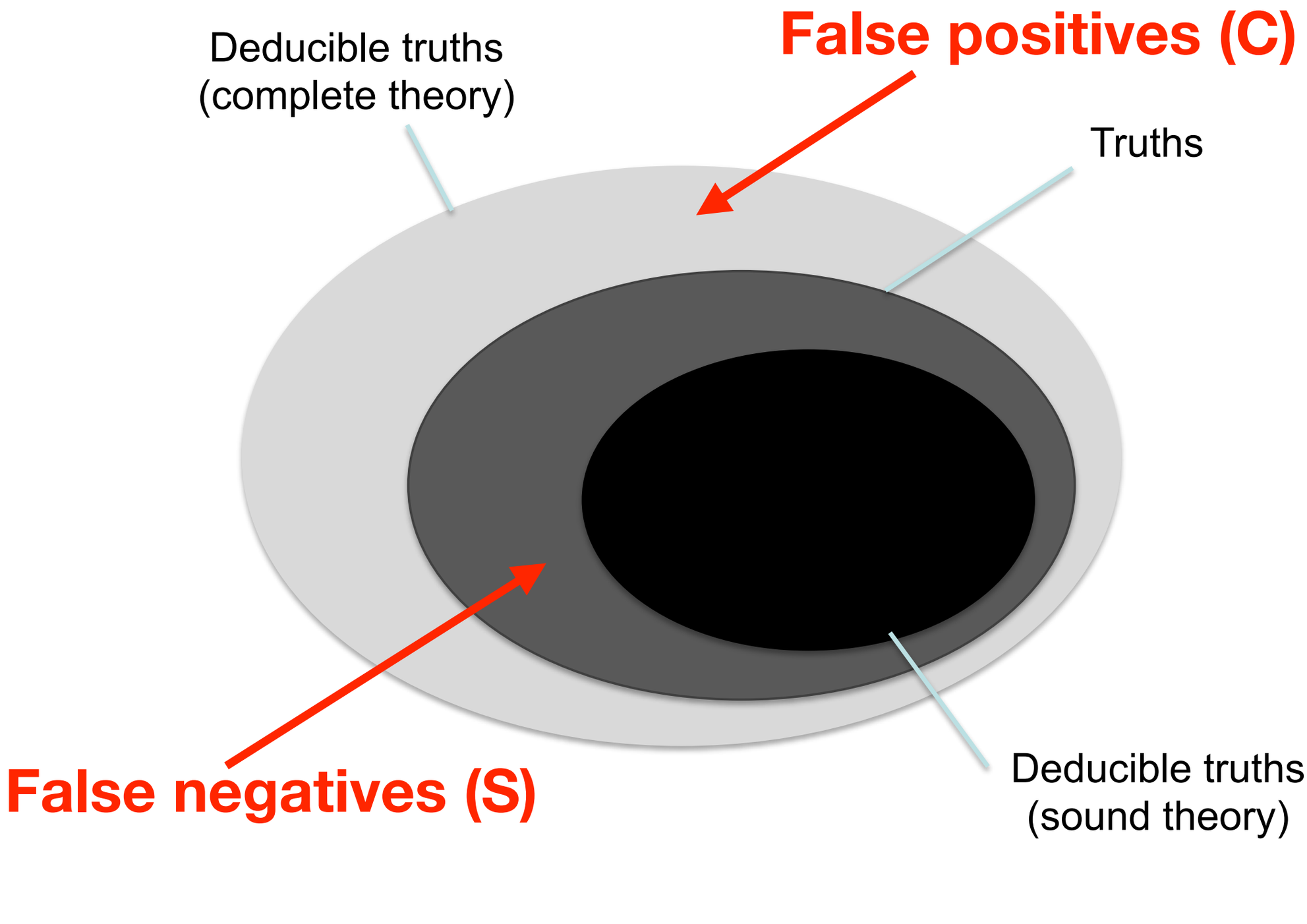
- Scientific discoveries in analytic and constructive research never expressed as absolute truths
  - Theories true until disproved by further observations
  - New kinds of artifacts good until better ones invented
    - Outputs of constructive research are often "approximate" solutions
      - e.g, algorithms solving problems for which an exact automatic solution is not known, or does not exist (undecidable problem), or it is impractical (too costly)

How can research results be validated?

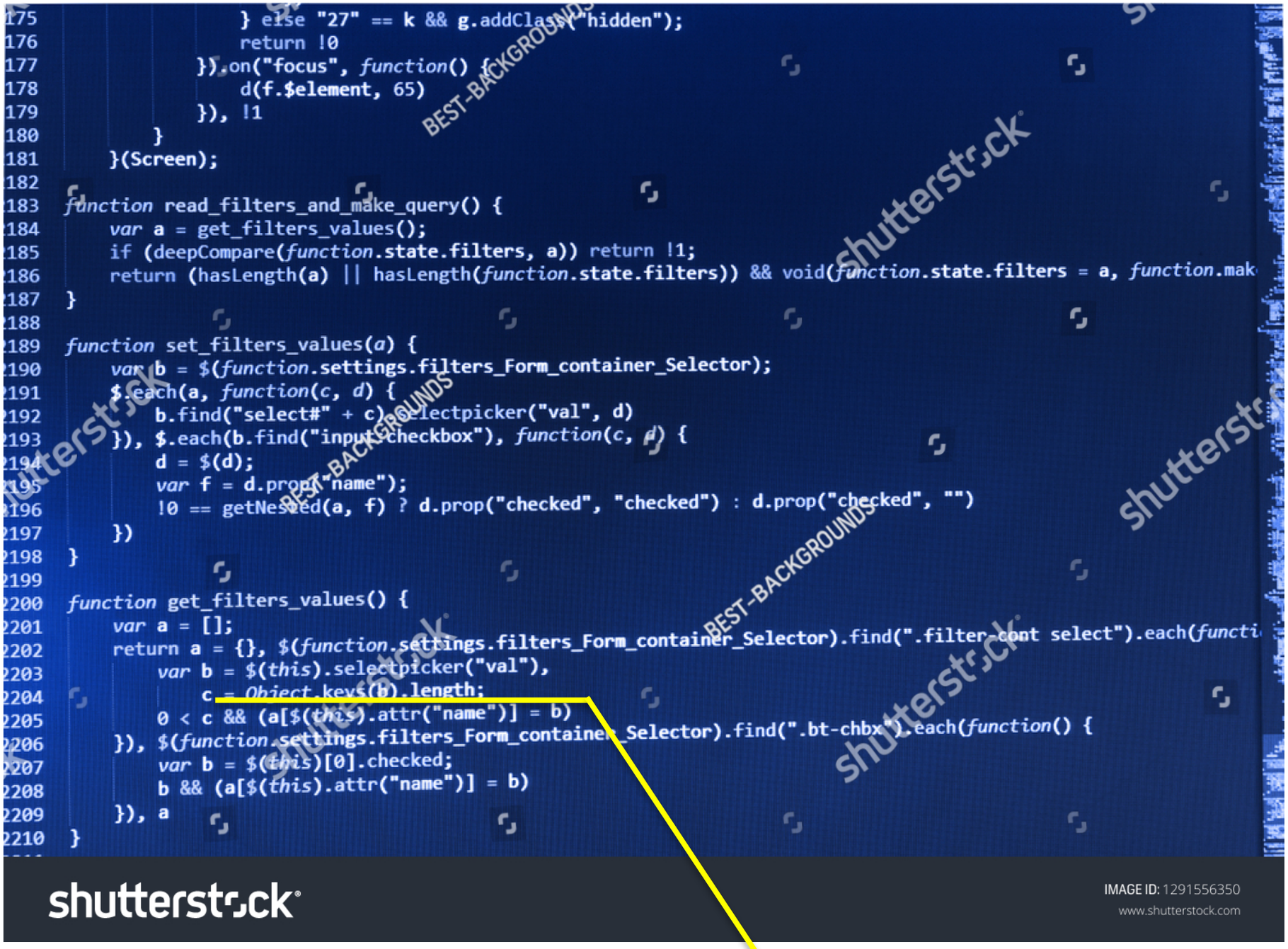
Soundness and completeness

- Concepts coming from logic
- Sound theory
    - all deductions are truths (can't deduce wrong things)
        - the potential predictions made by the theory are observable for the real phenomenon
- Complete theory
    - all truths can be deduced
        - everything that can be observed is predicted by the theory

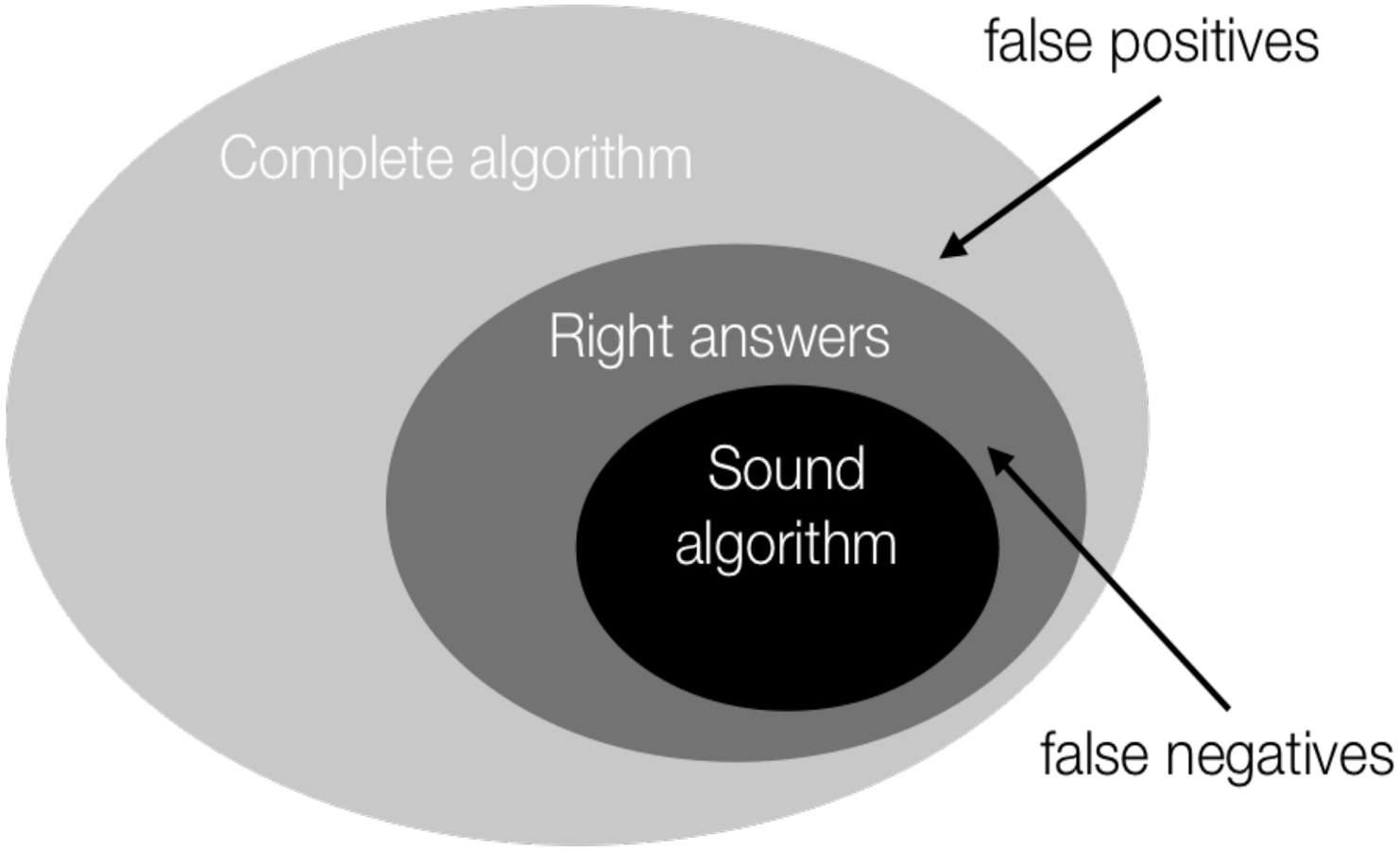
Soundness and completeness



Approximate algorithms: program analysis



is this statement reachable?



Solutions sometimes neither sound nor complete
—> Adequacy of approximation

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

Validation by case study

- Up-close, in-depth, and detailed examination of a subject of study (the case), as well as its related contextual conditions

- Common in constructive research, but also in social sciences and in life sciences

Observational case study

Feasibility case study

Comparative case study

Community case study

Replicability case study

Research diffusion

- With few exceptions, research is an intrinsically open process

- Diffusion essential for validation

- Researchers must play both the role of "producers" and "assessors"

- Researchers' progress depends heavily on diffusion of their results

Diffusion

- Publications (technical papers)
- Artifacts (tools, prototypes, data sets, …)
- Presentations
- …

Publications—*the bleeding edge*

- Publish or perish: ask the right questions
- We should never confuse **means** with **ends**
- We should never privilege **quantity** over **quality**

Tips: do's

- Be critical on your outputs and target venue based on honest assessment of contribution
- Adopt a diffusion strategy that is incremental
  - specialized workshops/conferences, major conferences, journals
- Understand exactly where you are positioned with your work
- Invest effort on development of artifacts
- Be an active member of the community, collaborate
- *… and many more…*

Tips: don't do's

- Don't be obsessed by publication… good work eventually recognized
- Don't follow easy paths, don't get trapped by predatory publishers *(strive for quality over quantity)*
- Beware of fashion and hype
- *…and many more…*

Evaluation

- Evaluation intrinsic in research
  - supports merit, quality, accountability
- Can apply to research
  - for funding
- Can apply to researchers
  - for hiring and promotion

Survive and triumph over evaluations

Prepare for evaluation

- Accept the fact that you live in a competitive arena
- Prepare material and dossiers you can use to progress in you career
- for research funding
- for hiring/promotion

Be aware of evaluation processes

- Understand the criteria
- Make your enthusiasm/commitment/fitness evident
- No long, flat, boring laundry lists
- Stress **quality** and **impact** over quantity

Tips

- Research
  - **quality**: best paper nominations, prestige and selectivity of venue (give data)
  - **Impact**: citations, downloads, patents, adoptions, …
    - networking, funding
- Teaching
  - student evaluations, new teaching material, new course
- Supervising/mentoring
- Service

Understand bibliometrics and rankings

- Develop rational criticism, but do not ignore them
  - understand what indicators try to indicate
    - citation counts, impact factor, H-index, …
    - and their limits
- Use them wisely to explain your points

Ethical issues

- Independent of your research
  - in doing/diffusing research
    - falsification, fabrication, plagiarism, …
  - in relation with others
    - fairness in judgement and evaluation, recognition of work, …
- Dependent on your research

Research-dependent ethical issues

- Research increasingly directly involves and/or (potentially) affects humans
- Traditionally well-understood in life sciences and social sciences
- Increasingly and disruptively so in Informatics

Humans participating in research

- Case studies, usability studies, "in-the-wild" tech assessments, …
    - the notion of *informed consent*
        - compliance with protection laws
            - GDPR in the EU
- Beware of the need for clearance from your organization's ethics committee

Ethics-sensitive research

- Physicists have long been concerned with uses of their research for destructive purposes
- The technology we create shapes our society and even our lives: can we ignore that?
  - We create new cyber spaces
  - We embed intelligence through algorithms into autonomous entities
- We cannot ignore questions like "why?", "what for?"
- We cannot forget that we do research for the good of mankind
- We cannot afford delaying these questions to "after deployment"—*it is damned too late*

Just a few examples
do these raise ethical concerns?

- My research focuses on game design, more specifically on usability of a new interactive device.

- My research focuses on the use of humanoid robots as a possible aid to support interaction with autistic children.

- My research focuses on sentiment analysis. I try to identify and extract subjective information by monitoring online conversations in social media and help business to understand the social sentiment.

- My research focuses on autonomous vehicles.

- *… and many many others*

Have a wonderful future in research

Watch out for a forthcoming book published by Springer in 2020

"Being a researcher—an Informatics perspective"