Part 1

- ***Research***
- ***Knowledge***
- ***Theory***
- ***Originality***
- ***Research Process Models***

Research Methods in Computer Science

Ullrich Hustadt

Department of Computer Science
University of Liverpool

Research

Research (Dictionary)

Noun

1. Scholarly or scientific investigation or inquiry.
2. Close, careful study.

Verb

1. To study (something) thoroughly so as to present in a detailed, accurate manner.
   (Example: researching the effects of acid rain.)

Research

Study (Dictionary)

Noun

1. The pursuit of knowledge, as by reading, observation, or research.
2. Attentive scrutiny.

Verb

1. To apply one's mind purposefully to the acquisition of knowledge or understanding of (a subject).
2. To inquire into; investigate.
3. To examine closely; scrutinise.

Research

Research (http://en.wikipedia.org/wiki/Research)

- an active, diligent, and systematic process of inquiry in order to discover, interpret or revise facts, events, behaviours, or theories, or to make practical applications with the help of such facts, laws, or theories.
- a collection of information about a particular subject.
- derives from the Middle French and the literal meaning is "to investigate thoroughly".

Research

Research (Higher Education Funding Council for England)

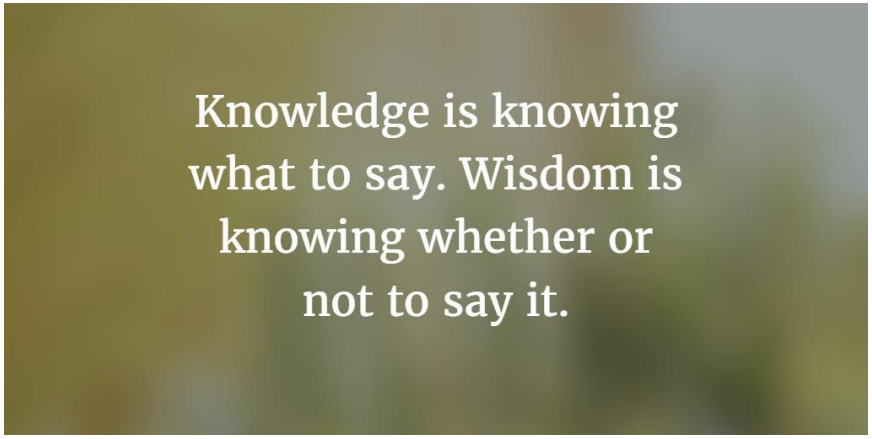
Original investigation undertaken in order to gain knowledge and understanding, including

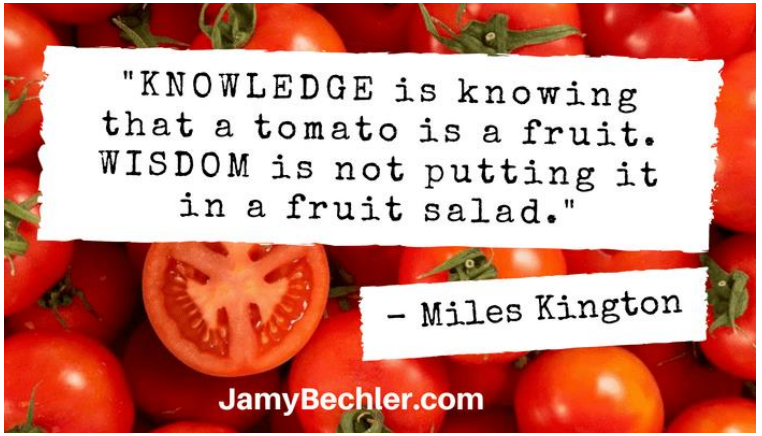
- work of direct relevance to the needs of commerce and industry and to the public and voluntary sectors
- scholarship (research infrastructure)
- the invention and generation of ideas, images, performances and artifacts including design, where these lead to new or substantially improved insights;
- the use of existing knowledge in experimental development to produce new or substantially improved materials, devices, products and processes, including design and construction.

Knowledge

Knowledge is a particular level in a hierarchy:

1. Data
2. Information
3. Knowledge
4. [Wisdom]





Data

Datum/Data

- statements accepted at face value (a 'given') and presented as numbers, characters, images, or sounds.
- a large class of practically important statements are measurements or observations of variables, objects, or events.
- in a computing context, in a form which can be assessed, stored, processed, and transmitted by a computer.

Information

Information

- Data on its own has no meaning, only when interpreted by some kind of data processing system does it take on meaning and becomes information

Example:

The human genome project has determined the sequence of the 3 billion chemical base pairs that make up human DNA

$\rightsquigarrow$ identifying base pairs produces data

$\rightsquigarrow$ information would tell us what they do!

Knowledge

Knowledge (Dawson 2005)

- higher level understanding of things
- represents our understanding of the 'why' instead of the mere 'what'
- interpretation of information in the form of rules, patterns, decisions, models, ideas, etc.

In natural sciences, understanding 'why' is too ambitious most of time; understanding 'how' is usually what we aim for

In other areas, understanding 'how' is trivial, understanding 'why' is challenging

Knowledge

Knowledge (http://en.wikipedia.org/wiki/Knowledge)

- the awareness and understanding of facts, truths or information gained in the form of experience or learning (a posteriori), or through deductive reasoning (a priori)
- an appreciation of the possession of interconnected details which, in isolation, are of lesser value
- both knowledge and information consist of true statements, but knowledge is information that has a purpose or use (information plus intentionality)

Theory

Scientific knowledge is often organised into theories.

Theory (http://en.wikipedia.org/wiki/Theories)

- a logically self-consistent model or framework describing the behaviour of a certain natural or social phenomenon, thus either originating from observable facts or supported by them
- formulated, developed, and evaluated according to the scientific method

Theory

Theory (http://en.wikipedia.org/wiki/Theories)

A body of (descriptions of) knowledge is usually only called a theory once it has a firm empirical basis, that is, it

1. is consistent with pre-existing theory to the extent that the pre-existing theory was experimentally verified, though it will often show pre-existing theory to be wrong in an exact sense,
2. is supported by many strands of evidence rather than a single foundation, ensuring that it probably is a good approximation if not totally correct,

Theory

Theory (http://en.wikipedia.org/wiki/Theories)

A body of (descriptions of) knowledge is usually only called a theory once it has a firm empirical basis, that is, it

3. makes (testable) predictions that might someday be used to disprove the theory, and
4. has survived many critical real world tests that could have proven it false,
5. is a/the best known explanation, in the sense of Occam's Razor, of the infinite variety of alternative explanations for the same data.

Fact

'This (e.g. evolution) is only a theory not a fact'

Fact

1. a truth (statement confirming to reality)

   or

2. data supported by a scientific experiment

- Status of a 'truth' is by and large unachievable
- A theory is formulated, developed, and evaluated according to the scientific method

  Given enough experimental support a theory can be (a scientific) fact

Originality

Research (HEFCE): Original investigation undertaken in order to gain knowledge and understanding

Originality

Doing something that has not been done before

Dawson (2005):
There is no point in repeating the work of others and discovering or producing what is already known

- Theories make predictions, which need to be tested
- The people performing those tests are neither infallible nor trustworthy
- Tests need to be repeated and results replicated

Originality

Areas of originality (Cryer 1996)

- Exploring the unknown
  Investigate a field that no one has investigated before
- Exploring the unanticipated
  Obtaining unexpected results and investigating new directions in an already existing field
- The use of data
  Interpret data in new ways
- Tools, techniques, procedures, and methods
  Apply new tools/techniques to alternative problems
  Try procedures/methods in new contexts

Fallibility

Cold fusion
(http://en.wikipedia.org/wiki/Cold_fusion)

- Cold fusion: Nuclear fusion reaction that occurs well below the temperature required for thermonuclear reactions, that is, near ambient temperature instead of millions of degrees Celsius
- First reported to have been achieved by Pons (University of Utah) and Fleischmann (University of Southampton) in 1989
- Scientists tried to replicate their results shortly after initial announcement
- Teams at Texas A&M University and the Georgia Institute of Technology first confirmed the results, but then withdraw those claims due to lack of evidence
- Vast majority of experiments failed

Investigation

'An active, diligent, and systematic process of inquiry' (Wikipedia)

- Scientists use observations and reasoning to develop technologies and propose explanations for natural phenomena in the form of hypotheses
- Predictions from these hypotheses are tested by experiment and further technologies developed
- Any hypothesis which is cogent enough to make predictions can then be tested reproducibly in this way
- Once it has been established that a hypothesis is sound, it becomes a theory.
- Sometimes scientific development takes place differently with a theory first being developed gaining support on the basis of its logic and principles

Contribution

Contribution

Research is supposed to add to the world's body of knowledge and understanding (in contrast to adding to the researcher's knowledge and understanding)

Research process models

All definitions agree that research involves a systematic or methodical process

Dawson (2005), following Baxter (2001), identifies four common views of the research process:

- Sequential
- Generalised
- Circulatory
- Evolutionary

Sequential research process model

Research process as

- Series of activities
- Performed one after another (sequentially)
- In a fixed, linear series of stages

Greenfield (1996):

1. Review the field
2. Build a theory
3. Test the theory
4. Reflect and integrate

Sharp et al (2002):

1. Identify the broad area of study
2. Select a research topic
3. Decide on an approach
4. Plan how you will perform the research
5. Gather data and information
6. Analyse and interpret these data
7. Present the result and findings

Problems with the sequential

1. Stages not subject specific
2. No repetition or cycles
3. Starting point and order fixed

Generalised research process model

- The generalised research process model recognises that the stages of the research process depend on the subject and nature of the research undertaken

  Example:
  Data gathering and data analysis play no role for research in pure mathematics and large parts of computer science
  Instead researchers make conjectures which they prove mathematically
- The generalised research process model provides alternative routes depending on the subject and nature of the research undertaken
- But each route is still sequential

Generalised research process model

Example:

(1) Identify the broad area of study
(2) Select a research topic

In natural sciences:	In mathematics:
(3) Decide on an approach	(3') Make a conjecture
(4) Plan the research	(4') Prove the conjecture
(5) Gather data and information	
(6) Analyse and interpret these data	

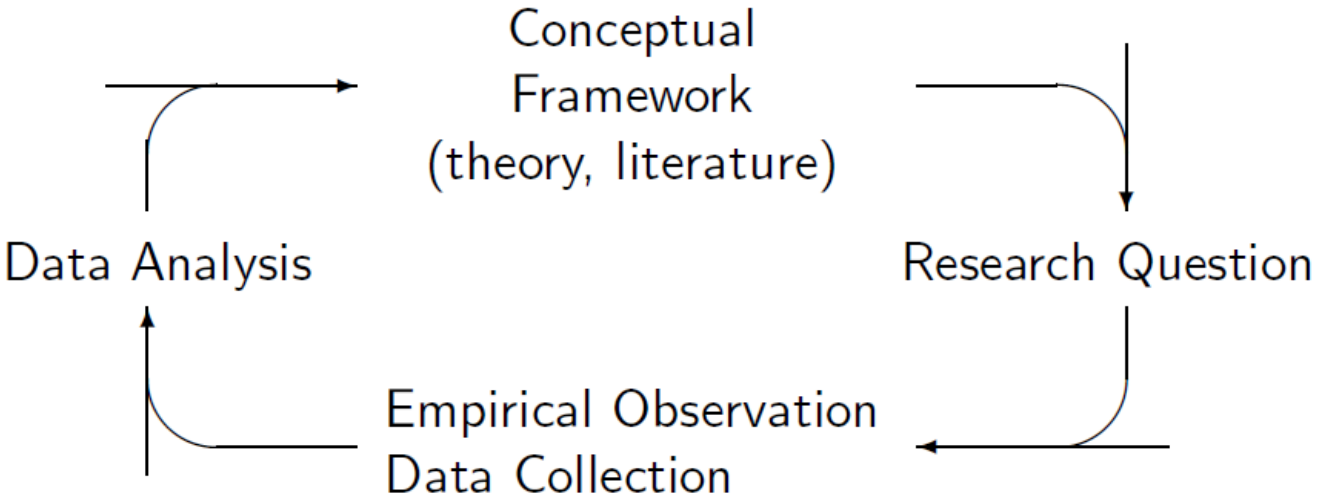
(7) Present the result and findings

Problems with the generalised process model:

1. No repetition or cycles
2. Starting point and order fixed

Circulatory research process model

- The circulatory research process model recognises that any research is part of a continuous cycle of discovery and investigation that never ends
- It allows the research process to be joined at any point
- One can also revisit (go back to) earlier stages



Circulatory research process model

An example of a circulatory process is Rudestam and Newton's *Research Wheel* (2007: 5),

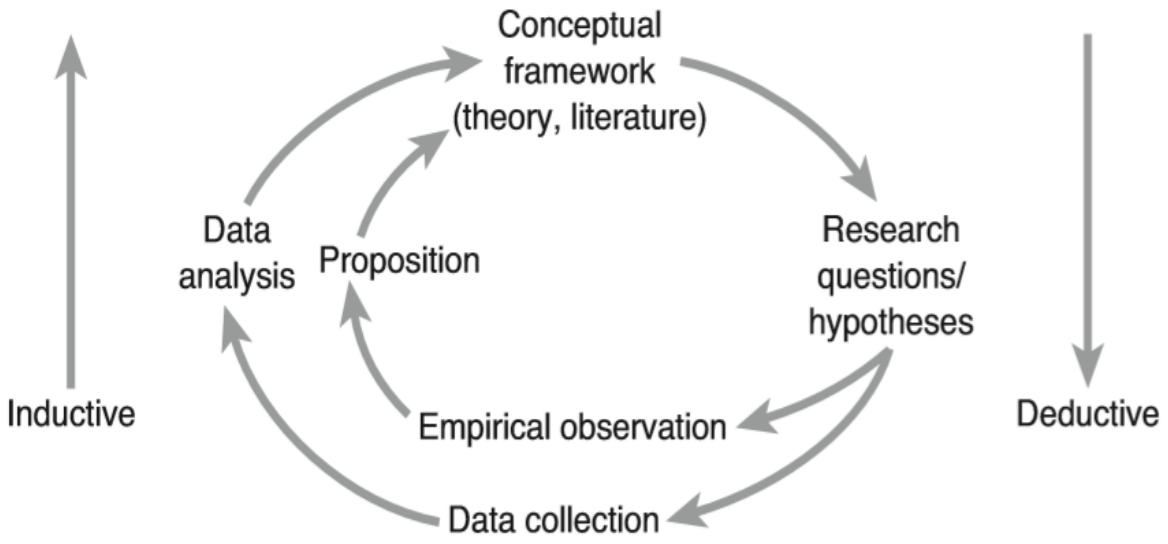


Figure 1.1 The research wheel

- Quite often, research will uncover more questions than it answers.

Evolutionary research process model

- The evolutionary research process model recognises that research (methods) itself evolve and change over time

  That is, over time our concept of
  - What research questions are admissible
  - What extend and methods of data collection are possible, necessary, ethical, or reliable
  - What methods are data analysis are available
  - What constitutes sufficient evidence for a hypothesis
  - What we mean by a systematic approach to research changes

Evolutionary research process model

- As an example, we can consider research in mathematics, in particular, its use of computers
- With respect to mathematical proofs we can make the following distinctions:

  (1) Proofs created solely by humans
  ⇝ typically 'sketchy', omitting steps that are considered 'obvious'

  (2) Computer-aided mathematical proofs
  ⇝ Structure and deductive steps still provided by humans, but certain computations are delegated to a computer

  (3) Fully formal, computer generated and validated proofs
  ⇝ Every step of a proof is conducted and validated by a computer, possibly under guidance by humans

Evolutionary research process model

Computer-aided mathematical proofs (1)

Four colour theorem

> *Any planar map can be coloured with at most four colours in a way that no two regions with the same colour share a border.*

Conjectured in 1852 by Guthrie. Proved in 1976 by Appel and Haken. Proof involves a case analysis of about 10,000 cases for which the help of a computer was used

Proof seems generally accepted, but not by all Mathematician

Evolutionary research process model

Computer-aided mathematical proofs (2)

Sphere packing theorem

Close packing is the densest possible sphere packing.

Conjectured in 1611 by Kepler. Hayes published a proof plan in (1997). Execution of the plan involved solving about 100,000 linear optimisation problems using a computer. The computer files for the related programs and data requires more than 3GB of space

At one point it was suggested that the proof will be published with a disclaimer, saying that it is impossible for a human to check its correctness

Research process model

- Among the four common views of the research process
  - Sequential
  - Generalised
  - Circulatory
  - Evolutionary

  the evolutionary research process model best describes the 'real' research process
- While the evolutionary research process model allows for the 'rules of the game' to change over time, this does not imply there aren't any rules
- For a young researcher it is best to follow the current established research process

End of part 1

- *Research*
- *Knowledge*
- *Theory*
- *Originality*
- *Research Process Models*

Research Methods in Computer Science

Ullrich Hustadt

Department of Computer Science
University of Liverpool