

Informatica, Watson en de Screenagers

Universiteitsdag 2011
Alumnimiddag Informatica



Jan van Leeuwen

**Departement Informatica
Universiteit Utrecht**

The background of the slide is a photograph of the Jeopardy! game show set. It features large, illuminated blue letters spelling out 'JEOPARDY!' on a dark blue background. The letters are viewed through a grid of black lines, possibly representing a window or a screen. The lighting is dramatic, with blue and purple hues.

What is Watson?

Can a computer compete against the world's best Jeopardy! contestants?

Can machines think...?

Jeopardy!: van omschrijving → identificatie van kennis

- Dit wetenschapsgebied kreeg in 1964 zijn Nederlandse naam, ligt ten grondslag aan de gehele ICT en haar toepassingen in de maatschappij, en wordt gezien als een van de belangrijkste drivers van de innovatie in Nederland.
- Volgens de voorspelling zal de computer in dit jaar de mens voorbijstreven in kennis en slimheid.
- Het zorgt voor koorts, pijn in je linker been, ...

Vraagbaak → kennis → intelligentie ...?

A.M. Turing (1950)



“Watson can process 500 gigabytes, the equivalent of a million books, per second.

The content was stored in Watson's RAM for the game because data stored on hard drives are too slow to access. ”



David Ferruci (IBM)

Watson is a workload optimized system designed for complex analytics, made possible by integrating massively parallel POWER7 processors and the IBM DeepQA software to answer *Jeopardy* questions in under three seconds. Watson is made up of a cluster of ninety IBM Power 750 servers (plus additional I/O, network and cluster controller nodes in 10 racks) with a total of 2880 POWER7 processor cores and 16 Terabytes of RAM. Each Power 750 server uses a 3.5 GHz POWER7 eight core processor, with four threads per core. (Cf. *IBM, Wikipedia*)

Hoe ...

- Brad Rutter (deelnemer) over `Watson':
 - **“They figured out algorithms in a way that they can actually do what the brain does.”**
- Henry Liebermann:
 - ``It all depends on how it's done. Scientists judge by the principles and techniques, not by the contest results.”
- Commentaren:
 - **“The idea of an interface with a computer system that is capable of interpreting human language [and answering] is something that a lot of people think will be quite remarkable and pave the way for more interactive robots in health care, help desks, defense, and maybe even education.”**
 - ``A next step in search engine technology..



Het “object van studie” van de Informatica?

- Computers en software systemen?
- `ICT solutions' in industrie, bedrijfsleven, maatschappij?
- Intelligente systemen?
- Het web?
- Wat maar geautomatiseerd kan worden?
- Hoe de `echte' wereld in de digitale is voor te stellen, te begrijpen, bruikbaar te maken?
- `Mentale constructies' realiseren?
- All of the above?

Informatica ... begin

- **Beheersing van ...:**
 - “The exponential growth in numbers of human beings and data about them, and the increasing pace and complexity of our technological civilization create problems of information processing for which the electronic computer holds the only visible means of solution.” (Forsythe, 1967)
- **Informatie (als dimensie van de wereld):**
 - De (ontworpen) wereld te beschrijven, modelleren, simuleren, begrijpen, ...
 - De fijnstructuur van (natuurlijke en artificiele) systemen te kunnen bestuderen.
 - mbv algoritmen, processen, agents, expert simulaties, ..
 - (Geen) `natuur'-wetenschap?
 - “Science of the artificial” (Herbert Simon)?
- **Veel definities en beschrijvingen `in context':**
 - “extension of the human mind” (Forsythe 1967..)
 - “phenomena surrounding computers” (Newell, Perlis, Simon, 1967).
 - “what can (and should) be efficiently automated” (COSERS 1980)
 - “algorithmic processes that describe and transform information” (Denning, 1985)
 - “science of information processes and their interactions with the world” (Denning, 2003)
- **Oorsprong (veel) ouder dan de computer...**
 - Sinds informatie wordt opgetekend.
 - Vergelijkbare `informatierevoluties' al eerder, bijv tgv uitvinding van de boekdrukkunst.



Kerngebieden

- Networks
 - Internet communications
 - Protocols
- System/Web/Mobile code
 - High-level PL's research
 - Exchange languages
 - Software design
- Complexity
 - Algorithm design
 - Game theory
 - Computational complexity
- Data/knowledge analysis
 - Data mining
 - Semantic web
- Multimedia
 - Data compression algorithms
 - Image analysis
 - Computational geometry
 - Natural language technology
- Web science
 - Search engines
 - e-Markets
 - Social computing
- Computational systems
 - Algorithmics
 - Multi/many-core programming
 - Parallel compilation
- E-science
 - Virtual laboratories
 - Diagnostic systems
 - Life science informatics
- Cognitive systems
 - Computational theory of mind
 - Intelligent systems
 - Sensor networks and robots
 - Human-computer interaction
- Information systems
 - Transaction systems
 - Operations research
 - Value chain informatisation
- E-society
 - Information security
 - Enterprise architecture
 - E-services
- Philosophy...

Adapted from R. Constable (2000)



Other People's Responses to the fact that I am a Computer Science Major



*To which my typical response is "Yeah, mostly ones and zeroes"

Wetenschapsfilosofie ... begin

- **Kenmerking van de Informatica**

- G. Zoutendijk (1964):“... de informatica, *de wetenschap die zich tot taak stelt het proces van informatieverwerking en de daarop gebaseerde beslissingen te bestuderen.*”
- **Belangrijk debat**
 - Waar gaat het over? (Objecten, verschijnselen.)
 - Wat willen we daarover weten? En waarom? (Doel.)
 - Wat verklaart het? Hoe? En klopt dat? (Modellen, hypothesen, theorie, experiment.)
 - Wat is eigen? (Methodologie, wijze van modelleren, ...)
 - Waar leidt het toe: wat, welke kennis, welke waarde? (Technologie, spiraaleffect.)
 - Hoe ontwikkelt zich deze kennis? (Invloeden.)

- **Wetenschap versus technologie versus ...**

- **Natuurkunde, wiskunde, ...**
 - Wetenschap als doel, technologie als middel om die te ontwikkelen, verbeteren, ...
- **Informatica**
 - Technologie als doel, wetenschap als middel om die te ontwikkelen, verbeteren, ...
 - Of toch omgekeerd?
 - `Computer' is niet (meer) de basis noch eigen; gebied al voor de computerrevolutie
- **Essentie van (informatie-)technologie?**
 - W.B. Arthur: “Technology: phenomena captured and put to use, leading to design of systems. Science needed to uncover and understand phenomena, technology needed to advance science’. *Science does not lead.* Everything emerges out of technology.”

Wereld van ontdekkingen

- Von Neumann architecture
- Computer can be used for human to human communication
- Definition of [classical] computability
- Core of intractable problems
- Formal frameworks for problem mapping and solving: calculi, LP, ..
- Public key cryptography
- Use of randomness as a computational resource
- Can formalize arguments about code
- Anthropomorphic/cognitive modeling of systems through 'agents'
- Distributed systems can [be programmed to] self-organize
- The web as PC
- How to search massive data spaces effectively (PageRank)
- Structure of social networks (power laws)
- DNA can compute
- Principles of quantum computing
- Embedding of physical problems and digital simulation
- Virtual businesses (.com)
- Business processes can be outsourced
- Principles of [electronic] markets
-



Filosofie van de Informatica

- **Historische ontwikkeling**
 - Centric views: logic, mathematics, programming, engineering, specific applications, etc
 - Invloeden: scientific, technological, social, societal, political, ...
 - Van 'tool perspective' tot 'science perspective' tot 'innovation perspective' tot....
- **Theorie van 'perspectives'**
 - (1) Information-oriented, (2) Computing oriented, (3) Communication-oriented, (4) Cognition-oriented, (5) Design-oriented, (6) Behaviour-oriented.
- **Filosofie van**
 - De perspectives
 - Specifieke gebieden (AI, agents, information, design methodology, risk, ...)
 - Rol van klassieke (wetenschaps-)filosofie, paradigmas
- **Wetenschappelijke methode**
 - Stapsgewijze modellering, algoritmen, talen, formele precisie, ontwerpmethoden?
 - 'Third way': theory, experiment, *program*?
- **Impact op denk- en ontwerpprocessen**
 - Algorithmic (computational) thinking, design science
 - Bewustwording van complexiteit, organisatie van kennis, ...
 - "Can machines think, have emotions, etc" (mimicking behaviour by mass computation)
- **Impact op onze waarden:**
 - Cultuur, democratie, politiek, milieu, ethiek, ..
 - Screenagers
- **Impact op onze toekomst**
 - Beheersing, socio-technische innovatie
 - Kurzweil's 'Singularity'

Enkele kernvragen

- **Hoe is de (echte of man-made) wereld in de computer voor te stellen?**
- **Wat is informatie?**
- **Wat zijn de information system principes van (natuurlijke of gecreeerde) systemen?**
 - Van natuurlijke systemen (bv cel), organisaties, het Web, communities, social networks
- **Wat is intelligence?**
 - Begrijpen van cognitie en andere menselijke kwaliteiten, affective computing.
- **Wat is berekenbaar?**
 - Nieuw inzichten na Turing's 1936 model? Computational modeling van fysische of biologische processen ('quantum computing'), van computing artefacts ('amorphous systems')
- **Wat zijn de grenzen van het eindige?**
 - Using current or imagined technology, classifications ('P vs NP').
- **Hoe kunnen complexe systemen (adequaat) ontworpen/gemaakt worden?**
 - Specificeerbaarheid, programmeerbaarheid, bruikbaarheid, onderhoudbaarheid, evolueerbaarheid; inbedbaarheid, afstemming (alignment), kwaliteit.
- **Wat is (efficiënt) te automatiseren?**
 - En wat zou wel/niet geautomatiseerd moeten worden?
- **Wat zijn de grenzen aan information system acceptatie?**
 - Games, human-media interaction, ethiek, inclusie.
- **Meer ... voor zover de verbeelding reikt!**
 - Grand challenges.

De App-generatie...

Kinderen zijn vanaf hun 3e jaar actief met interactieve media.

Stichtingen:
Mijn Kind Online
Z@ppelin
NTR Jeugd



