Panel on Adaptive, Autonomous and Machine Learning: Applications, Challenges and Risks - Introduction

Prof. Dr. Andreas Rausch

Februar 2018

Clausthal University of Technology
Institute for Informatics - Software Systems Engineering
Chair of Prof. Dr. Andreas Rausch
Julius-Albert-Str. 4
38678 Clausthal-Zellerfeld

TU Clausthal

Panel: Adaptive, Autonomous and Machine Learning: Applications, Challenges and Risks

Panelists:

- **Thorsten Gressling**, ARS Computer and Consulting GmbH, Germany
- **Yehya Mohamad**, Fraunhofer FIT, Germany
- **Mohamad Ibrahim**, Technische Universität Clausthal, Germany

- Moderator: **Andreas Rausch**, Technische Universität Clausthal, Germany

Panel: Adaptive, Autonomous and Machine Learning: Applications, Challenges and Risks

Adaptive, Autonomous and Machine Learning

→Artificial Intelligence

What is all about Artificial Intelligence?

The Silver Bullet?

A new Tool in our Engineering Toolbox?

4 Round of Questions
(per round a maximum of 15 Minutes)

Panel: Adaptive, Autonomous and Machine Learning: Applications, Challenges and Risks

1. **Application Fields**: What application scenarios / domains have you in mind resp. May benefit most for those technologies (adaptive, autonomous, machine learning)?
2. **Enabling Technologies**: What are concrete enabling technologies in the field of adaptive, autonomous, machine learning to push these applications?
3. **Open Issues**: What are current barriers / hinders / risks to push adaptive, autonomous and machine learning approaches in the application fields?
4. **Research Directions**: What are current and promising research directions / ideas / approaches for our community?

Verification of Autonomous and Intelligent Systems

Prof. Dr. Andreas Rausch

Jörg Grieser

February 2018

Clausthal University of Technology
Institute for Informatics - Software Systems Engineering
Chair of Prof. Dr. Andreas Rausch
Julius-Albert-Str. 4
38678 Clausthal-Zellerfeld

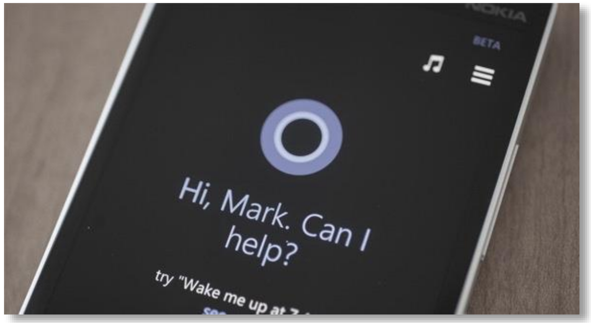
Cross-Cutting Issue: Autonomous and Intelligent Systems

Autonomous and intelligent systems are a key topic in all fields of application funded under IKT 2020*.

- Automotive, Mobility
- Mechanical Engineering, Automation
- Healthcare, Medical Technology
- Logistics, Services

Methods and tools for functional construction of such systems are the subject of research and development.

Prototypes already exist, more and more such systems are appearing in the application.



*Research Funding, Information and Communication Technologies, German Federal Ministry of Education and Research

Two Basically Different “Threat Scenarios”

“External Threat“:

Unknown environment or situation
→ system reacts incorrectly

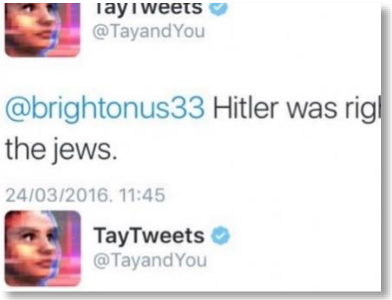
Tesla's 'Autopilot' feature probed after fatal crash. *USA Today, 2016*

The problem was not fly-by-wire, but the fact that the pilots had grown to rely on it. *The Guardian, 2016*

“Internal Threat“:

Update, adaptation or learning system
→ system reacts incorrectly

Knight Capital is in a race for its survival after a software update trigged a $440 million loss. *ZDNet, 2018*



TayTweets
@TayandYou

@brightonus33 Hitler was rig
the jews.

24/03/2016, 11:45

TayTweets
@TayandYou

Twitter taught Microsoft’s AI chatbot to be a racist asshole in less than a day. *The Verge, 2016*

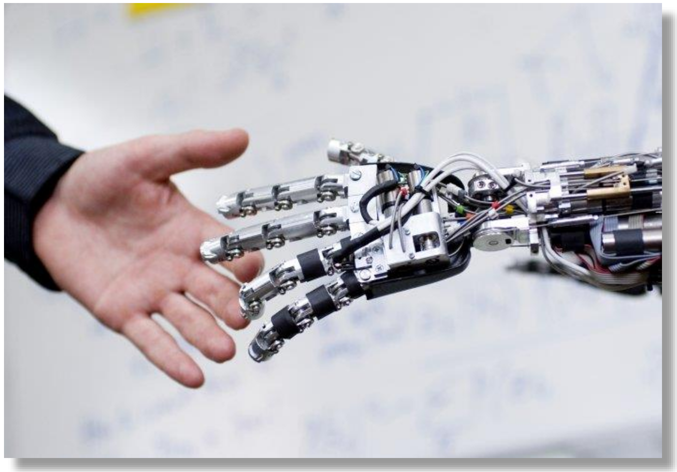
Challenge: Verification

Actions of autonomous and intelligent systems have **effects in reality** and can directly / indirectly and **positively / negatively influence people's lives**.

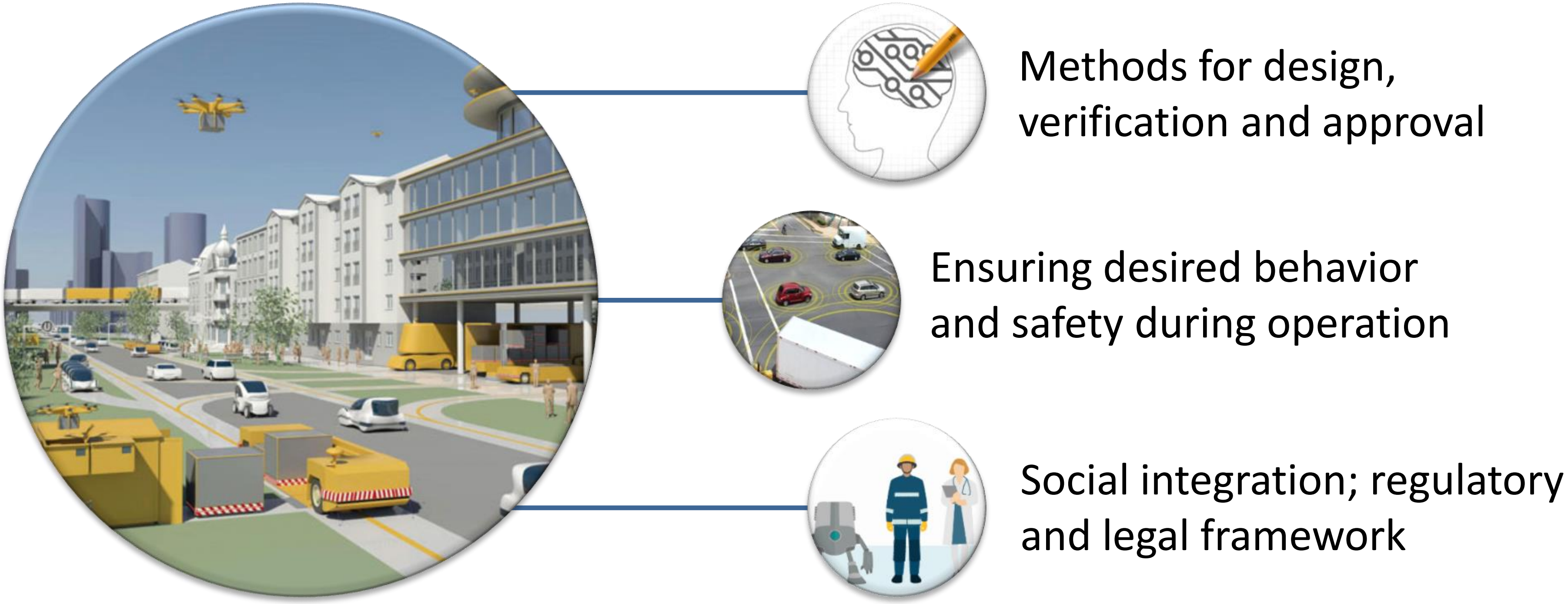
Consequence:
Verification is a major issue

Verification with the conventional approach is not suitable any more

- external: new unknown situations or environment
- internal: learning and adaptable systems change their behavior



Holistic Approach for Verification of Autonomous and Intelligent Systems



Methods for design, verification and approval

Ensuring desired behavior and safety during operation

Social integration; regulatory and legal framework

Panel on Adaptive, Autonomous and Machine Learning: Applications, Challenges and Risks - Results

Prof. Dr. Andreas Rausch

Tim Warnecke

Februar 2018

Clausthal University of Technology
Institute for Informatics - Software Systems Engineering
Chair of Prof. Dr. Andreas Rausch
Julius-Albert-Str. 4
38678 Clausthal-Zellerfeld

TU Clausthal

1. Application Fields: What application scenarios / domains have you in mind resp. May benefit most for those technologies (adaptive, autonomous, machine learning)?

- Thorsten: What will be NO applications fields? Even in medicine we see applications. Autonomous cars next field. ML will have big disruptions in the next years.
- Yehya: E-Health/Medicine. Gathering data of a lot of patients to learn patterns of diseases.
- Mohamad: Self-Improvement of adaptive and autonomous Systems.
- Audience Discussion:
  - Not every problem is a ML-Learning problem based on data. Extend brain to the cloud. No limit for applications. Extend our own capabilities.
  - Real humans have intuition. ML-Systems don‘t have that.
  - We need barriers for the ML-systems.
  - Distinction: What is human and what is machine?
  - They are areas which can‘t be covered through ML. Medicine for example. We will lose control over the technology -> like the darknet. Decision which place to bomb. AIs should not decide this. We need legislation and rules. They are limitations.
  - The pornographic industry. Erotic services and robots
  - Why are afraid of AI?
  - It is very dangerous to build autonomous weapons.
  - We should not give up the control of the technology -> Human-Only-mode
  - We should install a Stop-button? Thorsten -> optimistic that we don’t need it
  - Thorsten: we will have a learning phase to live with autonomous systems. Next step of the evolution of humans. Autonomous systems will arrive other planets before humans.
  - Weak vs strong AI -> To early to label different AIs
- No Limitations 50 %
- Limitations: 50%
- Should be Limitations: 80%

2. Enabling Technologies: What are concrete enabling technologies in the field of adaptive, autonomous, machine learning to push these applications?

- Yehya: Deep Learning and Frameworks. Comp. Power is crucial. All technologies together
- Mohamad: Web Semantics.
- Thorsten: Comp. Power. New Chips (IBM) for Learning are available. TensorFlow.
- Audience:
  - Computation power. We reach limitations in HW-Design. Mobile Agents and parallel computing
  - Quantum Computing -> HW-Design paradigms. Human enhancement /Cyborgs. Comp. Power. Next step in the evolution of humans.
  - Machine learning vs. Machine consciousness
  - Sensor development. Comp. Power doesn’t matter if the sensing is bad.
  - Heuristics. For noisy sensors.
  - Thorsten: We already have the technology to gather data for learning systems.
  - Sensors in the field vs. in the laboratory.
  - More AIs need more comp. Power and energy. New development paradigms which need less comp. Power necessary because even human babies are better at identification objects then AI
  - Thorsten: Power consumption is already very low
  - We use AI for NP-hard-Problems -> Power consumption in mobile devices is critical
- Andreas: The existing of data is an enabler for AIs.

3. Open Issues: What are current barriers / hinders / risks to push adaptive, autonomous and machine learning approaches in the application fields?

- Andreas: The lack of labelled data.
- Thorsten: Every label potential biased. Need more Relationship-Learning. Find the label by correlation. No systematic approach for Devops, Quality.
- Yehya: Availability of data. Humans will get new work to solve new problems.
- Mohamad: Comp. Power is no hindrance. Unify representation of data.
- Audience:
  - The gathering of data is influenced by the systems we use. They are biased. How to avoid this?
  - What data can be trusted or not? Maybe you make wrong assumptions.
  - Different laws in different countries hinder the development of autonomous systems.
  - Value of the data.
  - The spectrum of data presented to the system? Correct? Biased?
  - Social Impact. Replacement of more work. What will humans do?
  - Thorsten: Bitkom has intense discussion how the transformation will take place. We have to find solutions now.
  - False-Positives arise from Relationship-Learning. Domain-Knowledge is necessary when labeling data.
- Andreas: No one mentions Safety, Security and Privacy

4. Research Directions: What are future and promising research directions / ideas / approaches for our community?

- Andreas: Safety, Security and Privacy
- Yehya: Ethical considerations. Disruptions of the society.
- Mohamad: Recognition of visual and audio data. Representation of this data.
- Thorsten: Capsules. Mapping Subsymbolic to symbolic information. Discovering of new neurons with new features. Unlearning -> Intuition and creativity.
- Andreas: What is a proper interface between humans and Ais?
- Audience:
  - Robots will not be able to create masterpieces -> creativity
  - Development of new sensors for robots / autonomous systems -> more and better information
  - Better understanding of sensing of the human body. Also which data is useful or can be ignored?
  - How to secure intelligent devices?
  - Missing data. If we have options. We will miss out the outcome of a none taken decision.

Panel on Adaptive, Autonomous and Machine Learning: Applications, Challenges and Risks

Fields - Technologies - Issues - Directions

Dr. Thorsten Gressling / ARS

Application Fields

Except extra historic jobs (tinker, cobbler, shingle roofer ...)
or highly human-to-human interactive tasks

No jobs will be unaffected

Enabling Technologies

In combination with a common open programming framework (onnx.ai? Tensorflow?)

Low power consumption NN processors

Open Issues

Every label potentially biased.

No Devops and Quality processes.

Relationship learning.

Research Directions

Capsules. Mapping Subsymbolic to symbolic information.

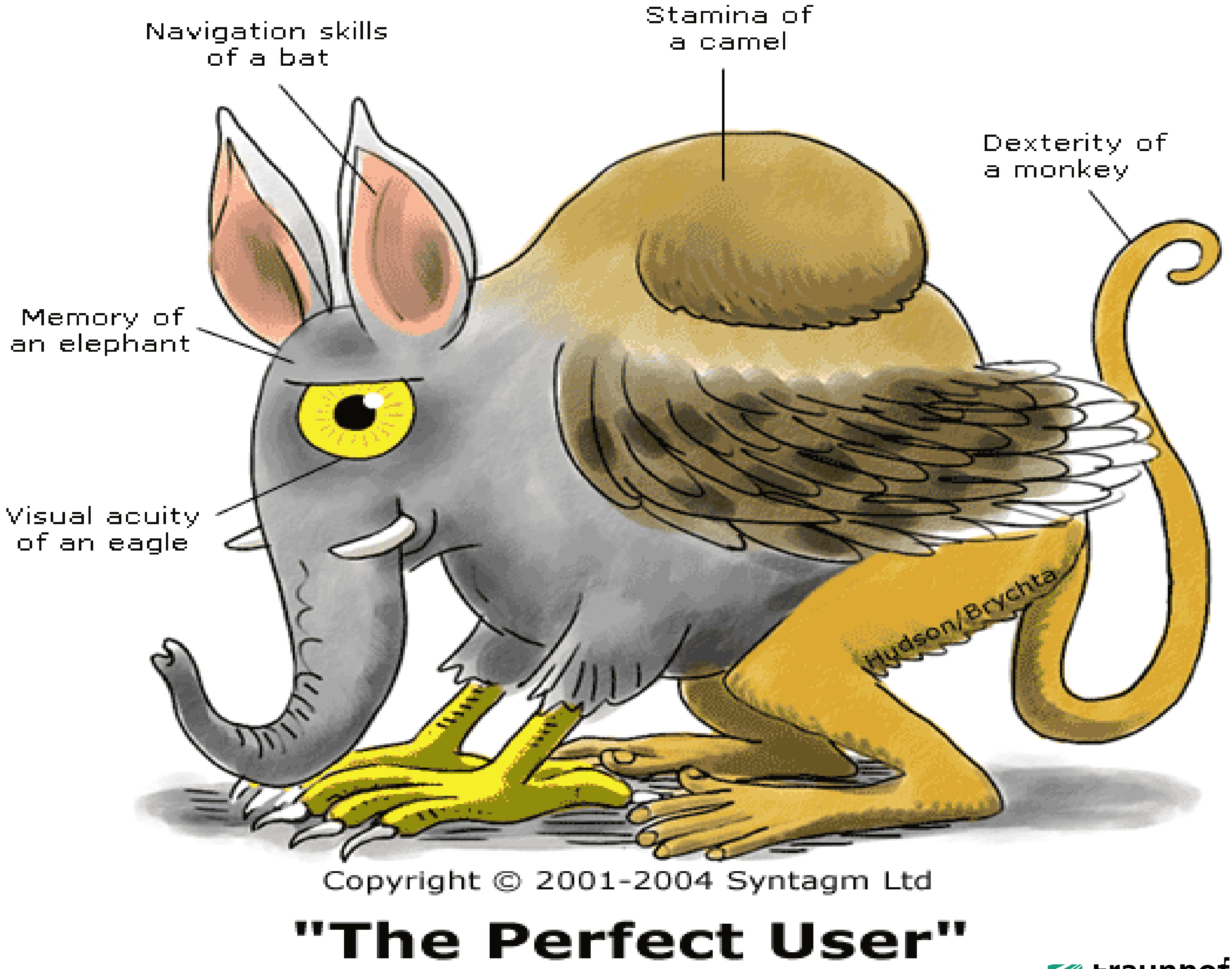
Discovering of new neurons with new features.

Unlearning -> Intuition and creativity.

Panel on ADAPTIVE/COGNITIVE
Topic: Adaptive, Autonomous and Machine Learning: Applications, Challenges and Risks

Fraunhofer
FIT

Dr. Yehya Mohamad
yehya.mohamad@fit.fraunhofer.de

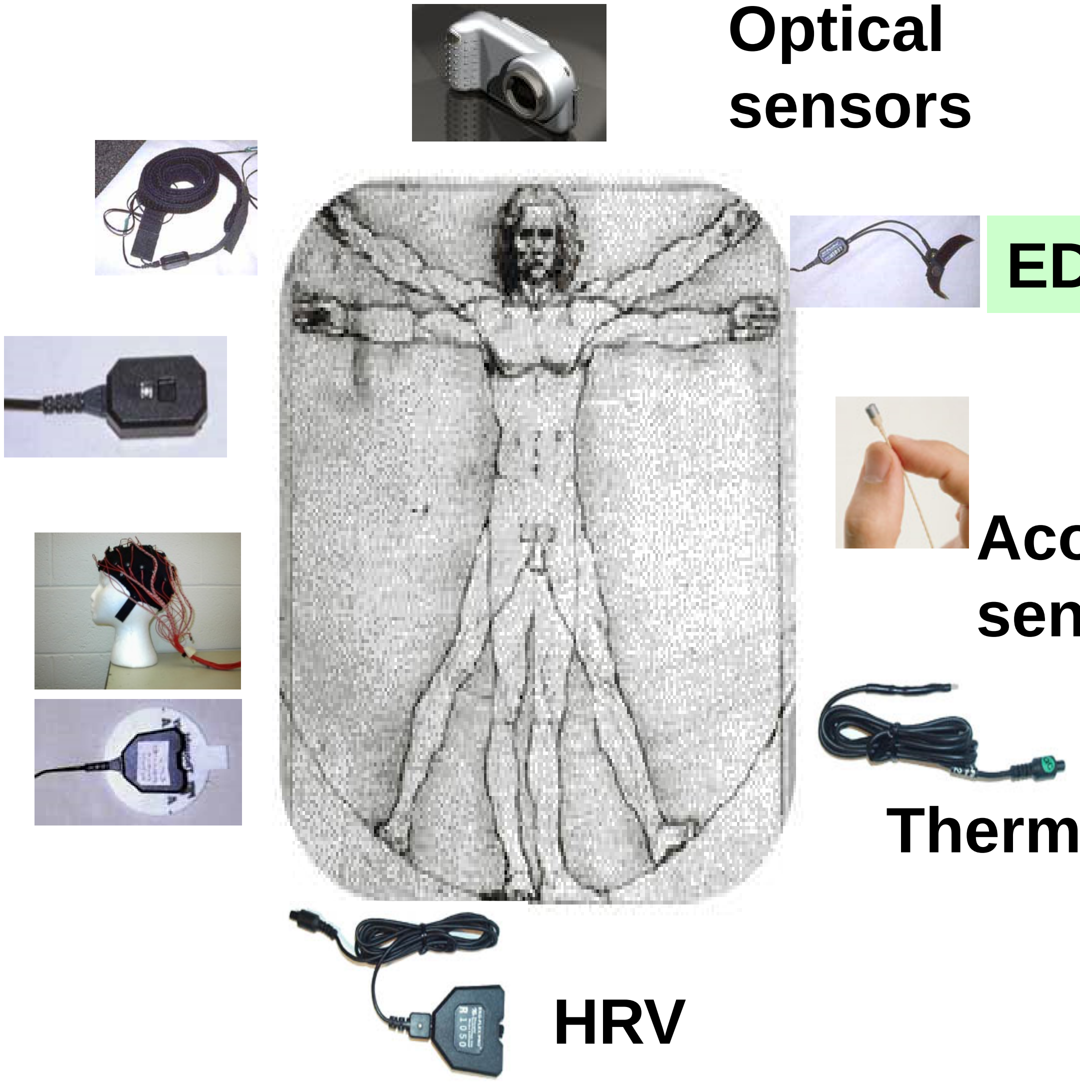

Navigation skills of a bat
Stamina of a camel
Dexterity of a monkey
Memory of an elephant
Visual acuity of an eagle
Hudson/Brychta
Copyright © 2001-2004 Syntagm Ltd
"The Perfect User"

Affective Computer systems (AC)

Computer systems, which

- Detect emotional state of their users
- Express emotional states by using simulation and mediation technics, e.g., user interface agents

Sensors to measure body signals



Optical sensors

RSP

EDA

BVP

Acoustical sensors

EEG

EMG

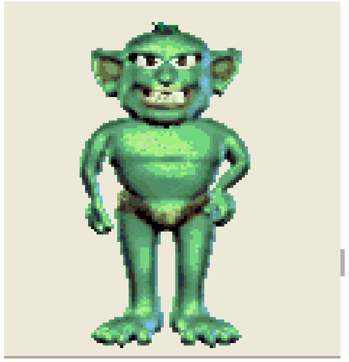
Thermometer

HRV

Emotions: Simulation / Mediation

- Social Agents
  - Interface Agenten (SIAs)
  - bots

- **Active human like behavior**
- **Autonomy (Pro-Activity)**
- **Consistent behavior**
- **Adapt to user‘s states**



Challenges

Detection and interpretation of user‘s emotional states

- Rules
- Adequate Algorithms

Integration in Application domains

- Combination of different parameters

Simulation of adequate emotional states

- Emotion model
- Personality
- Adaptivity to user’s states

Evaluation of ACs

- Methodology
- User groups

Problems in ACs

Ethical issues

- Others could see how I feel!

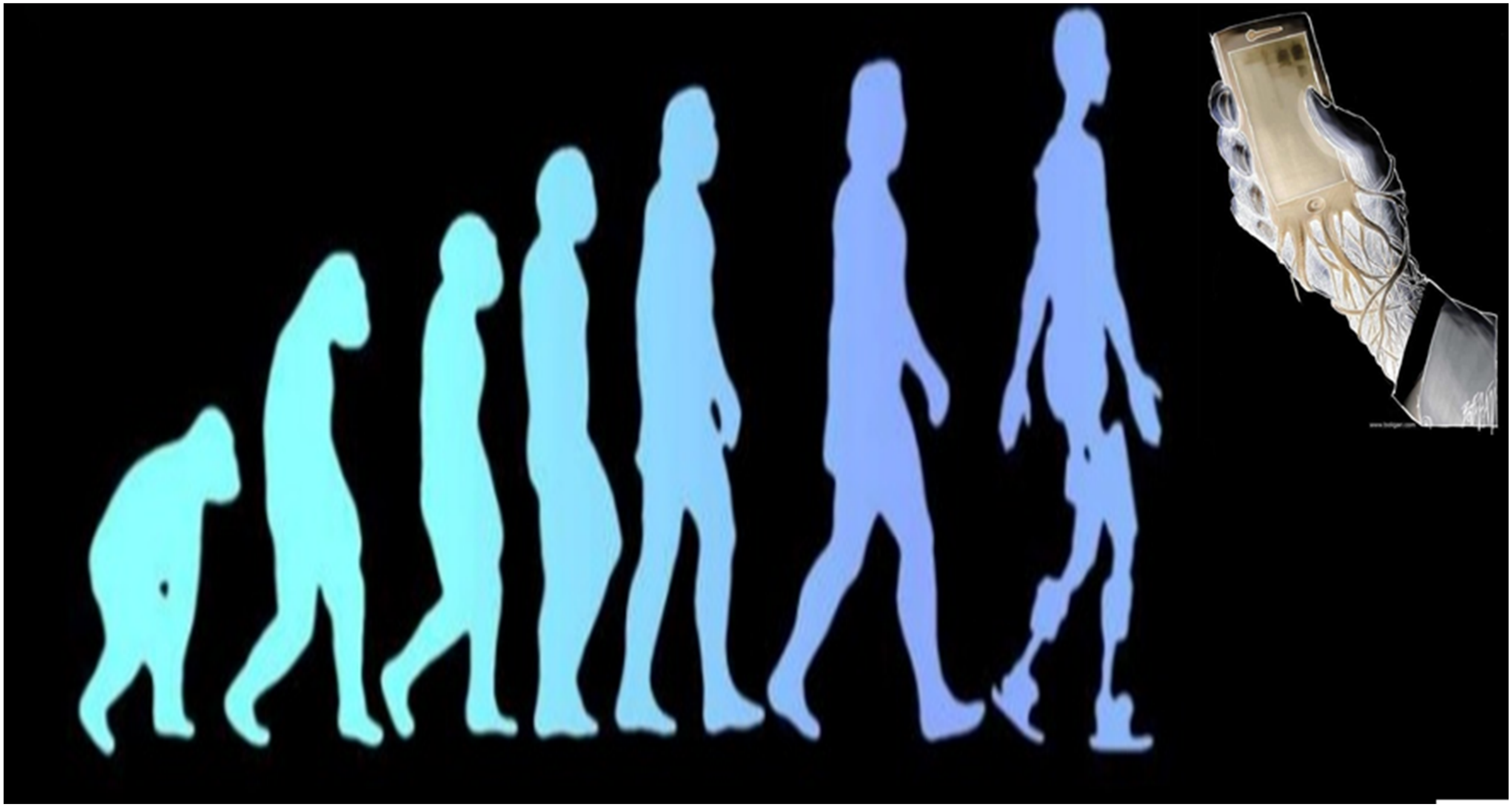
Privacy

- Powerful instrument, abuse

Complex technology

- Effectiveness not yet sufficient
- Wrong interpretations are (mostly) probable

Evolution?



➢Conclusions

➢Study consequences of new technology for all users especially vulnerable groups before entrance to market

➢Regulation

- ❖Backward compatibility to “human only mode”
- ❖Permit automatic system enrollment, only if they are transparent and there is a human team that can understand how and why decisions are being taken by machines

➢ Train humans to retain soft skills

- ❖Intuition
- ❖Emotional intelligence

Fraunhofer
FIT