Augmenting Human Mind by Artificial Intelligence

Andreas Dengel @ SPRi

Agenda

- The DFKI
- Augmenting Human Mind
  - … by helping people remember
  - … by identifying situational relevance
  - … by understanding human emotions

Where is Kaiserslautern?

DFKI is a success model for research in Germany

Bremen Site

- Cyber-Physical Systems
- Robotics Innovation Center

since 11/2005

Kaiserslautern Site

- Augmented Vision
- Embedded Intelligence
- Knowledge Management
- Innovative Factory Systems

Saarbrücken Site

- Intelligent User Interfaces
- Language Technology
- Agents & Simulated Reality
- Institute for Information Systems (IWi)

since 07/1988

Berlin Office

Members of all Departments

since 05/2007

Osnabrück Office

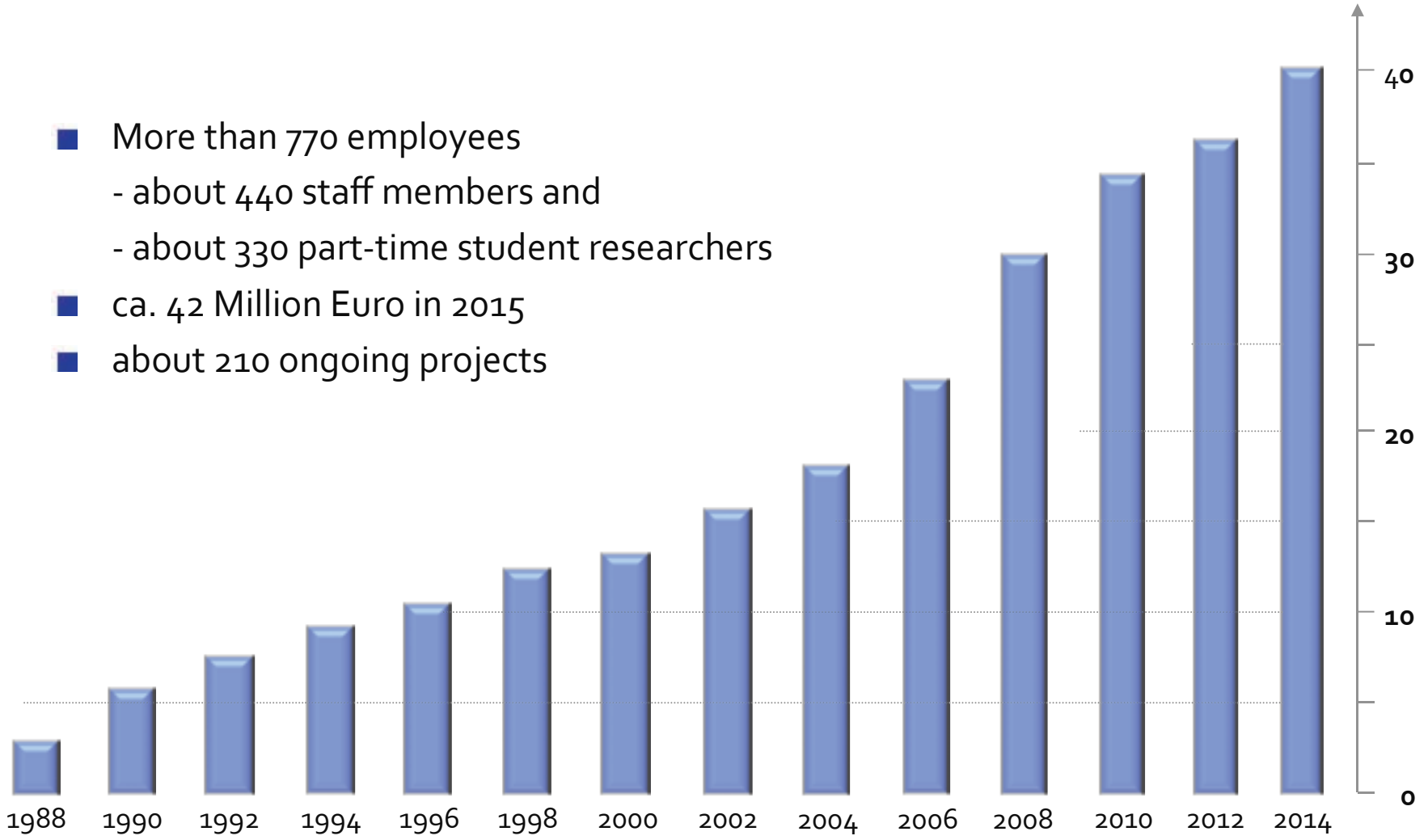
RIC members

since 03/2011

DFKI is a non-for-profit private research company with 23 international shareholders

DFKI has steadily grown since the beginning and today, is the largest AI research center in the world

- More than 770 employees
  - about 440 staff members and
  - about 330 part-time student researchers
- ca. 42 Million Euro in 2015
- about 210 ongoing projects

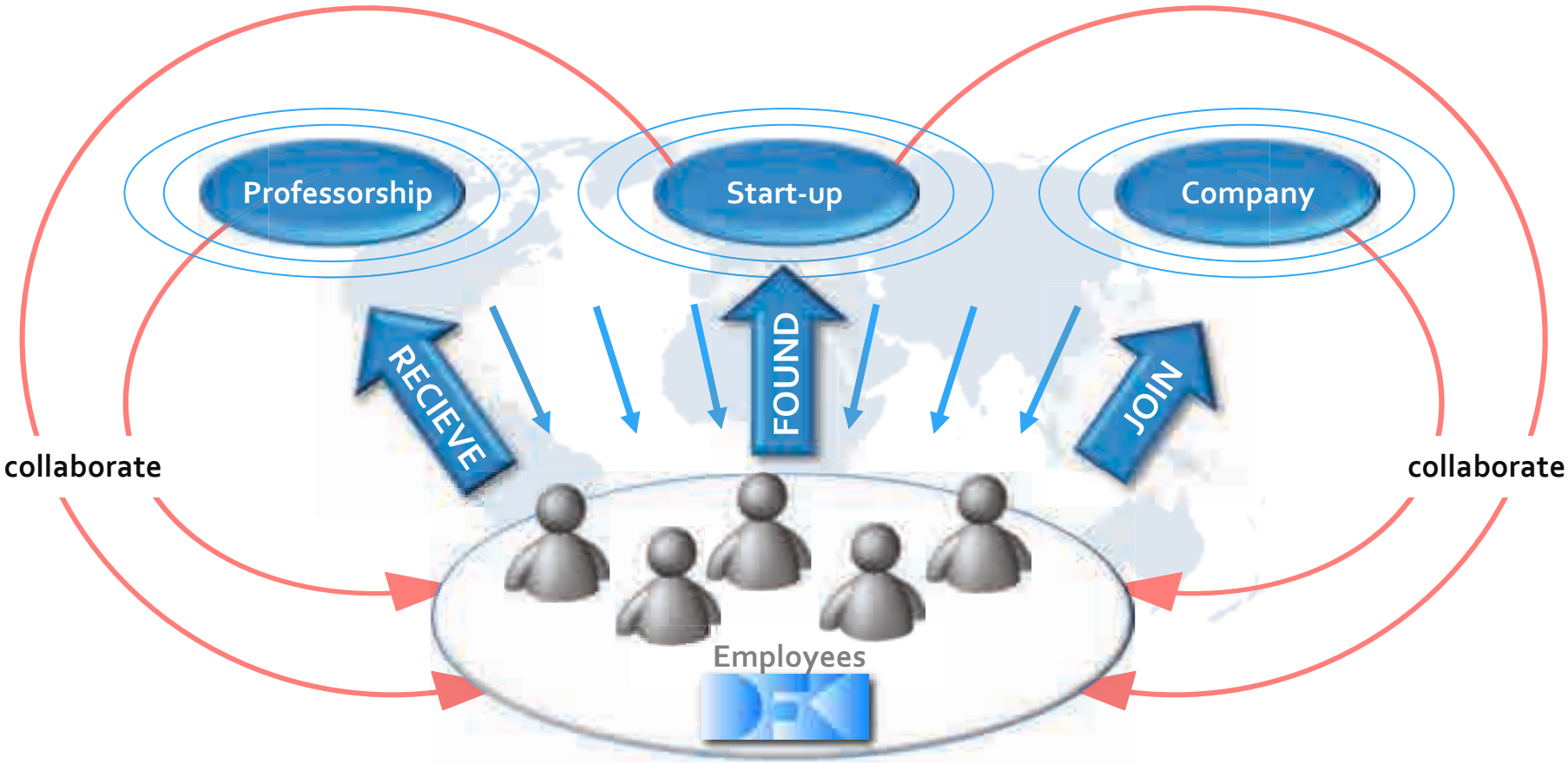


DFKI Customers (A Selection)*

* Shareholders are not specifically named!

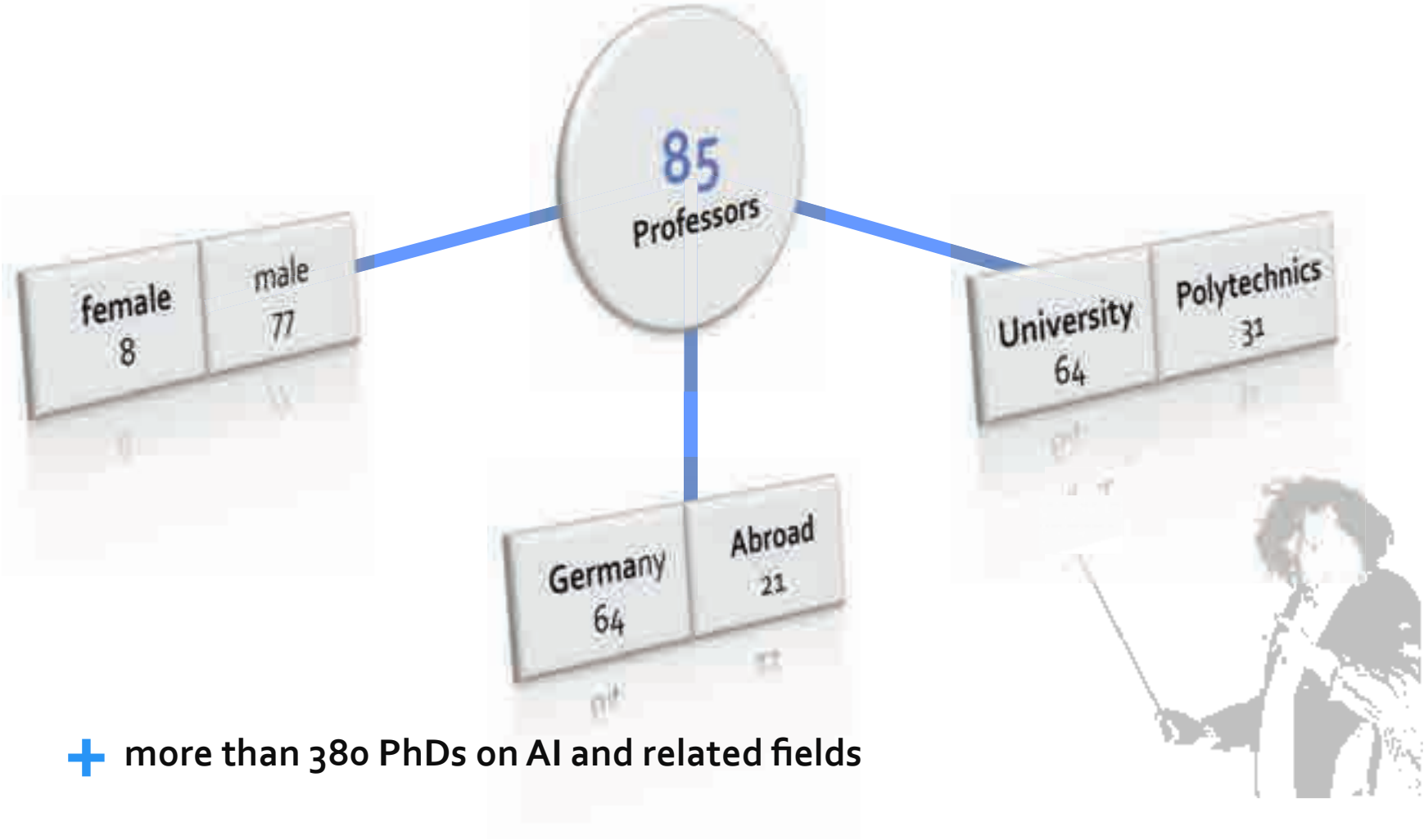
DFKI recruits worldwide: 238 foreigners from 56 countries

DFKI can build on a cycle strengthening its own network



Network of more than 4.500 Alumni!

DFKI is a Hot Spot for Academic Talents



+ more than 380 PhDs on AI and related fields

DFKI produced 67 start-ups and about 1,500 jobs

Agenda

- The DFKI
- Augmenting Human Mind
  - … by helping people remember
  - … by identifying situational relevance
  - … by understanding human emotions

Information entering the nervous system is under continuous editorial revision

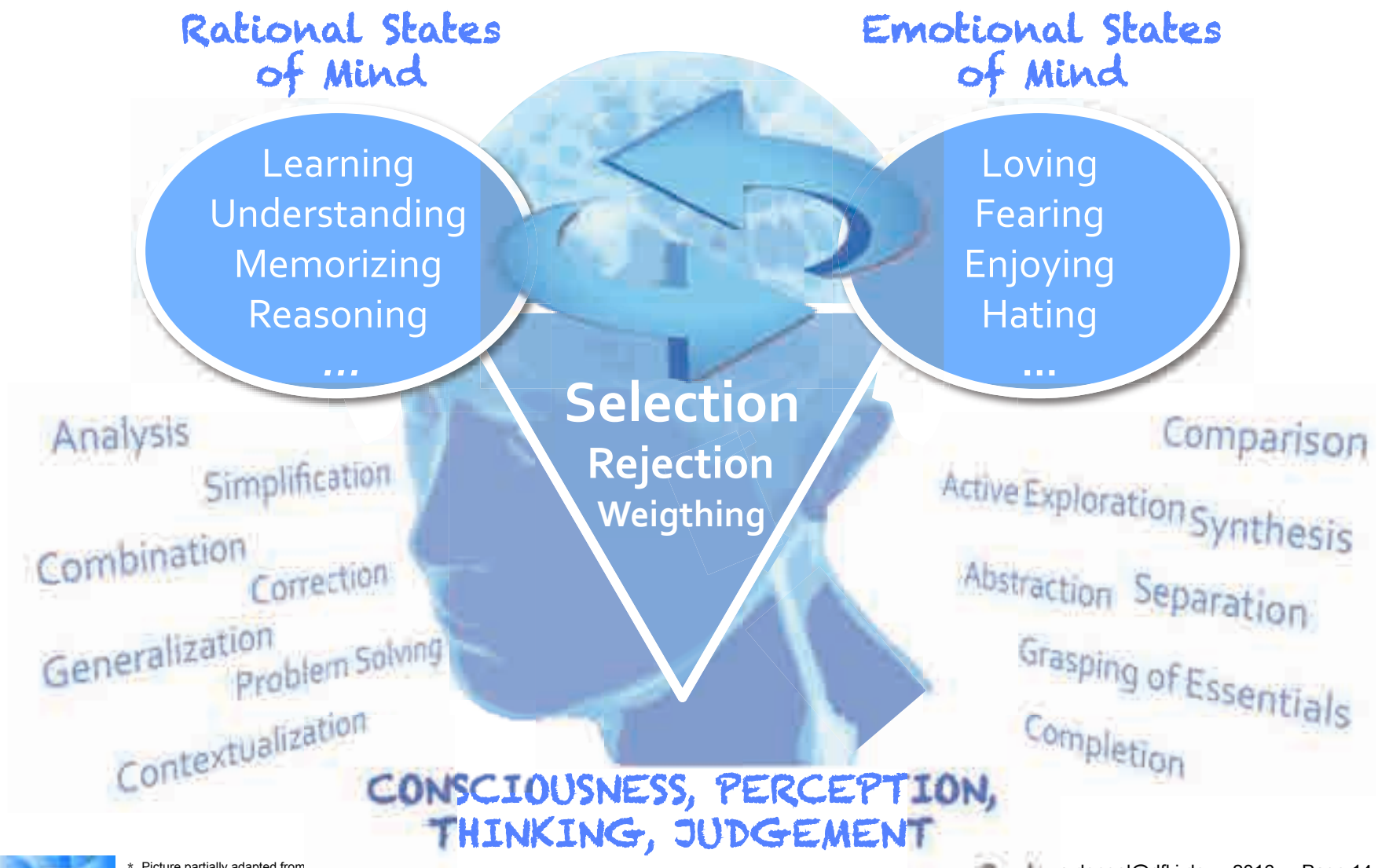
„**All varieties of perception** – indeed all varieties of thought or mental activity – **are accomplished** in the brain **by** parallel, multitrack **processes of interpretation and elaboration of sensory inputs**."*

Daniel Clement Dennett III (1942)

Fulbright and Guggenheim Fellow
Co-Director of the Center for Cognitive Studies and
Professor at Tufts University
Erasmus Prize Winner

* Dennet D.C. Consciousness Explained, Little Brown and Company Publ. (1991)

Our mind combines various rational and emotional states helping us to "make up our mind", "keep in mind", or "change our mind"



* Picture partially adapted from https://www.linkedin.com/pulse/20130211201055-175081329-12-minutes-to-create-a-mind-changing-presentation

Just recently Microsoft announced its Microsoft Graph

"In an age of information abundance, we understand people are seeking integrated experiences that help them break out of knowledge silos—ways to 'work like a network' and connect information from multiple touch points in meaningful ways."

Rob Lefferts (11/2015)

General Manager MS Office

We need systems that help us to keep up-to-date!

Source: https://fortunedotcom.files.wordpress.com/2015/11/microsoft-graph_hero-image.png?w=1024

Augmenting Human Mind means helping people to remember

The Semantic Desktop

L. Sauermann, A. Bernardi, and A. Dengel, *Overview and Outlook on the Semantic Desktop*, Proceedings ISWC, 6th International Semantic Web Conference, Galway, Ireland (Nov. 2005), pp. 1-19.

A. Dengel and B. Adrian, *Helping People Remember: Coactive Assistance for Personal Information Management on the Semantic Desktop*, in: A. Fred et al (Eds.) Proceedings IC3K2009, Revised Selected Papers, CCIS 128, Springer (2011), pp. 3-16.

Knowledge workers, like Heiko, have to remember much more than they are able to

From: heiko.b@dfki.de
Date: 02. <March 2015 09:28:11 MEZ
To: tanigalli@imu.portobello-u.it
Subject: Visit at DFKI

Dear Prof. Tanigalli,

Prof. Dengel asked me to send you the agenda for your visit on April 28th.

If you have any question, let me know.

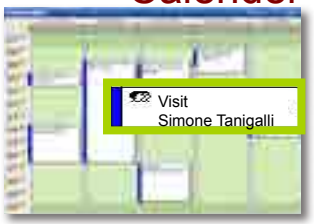
Best regards,
Heiko B.

Heiko B.
Event Management
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH
Trippstadter Straße 122
67663 Kaiserslautern
Tel.: +49 631 205 75 1999
Fax: +49 631 205 75 1990
Email: heiko.b@dfki.de

Agenda.pdf

I have to organize the visit of Prof. Tanigalli on March 28th!

Calender



Processes

Agenda

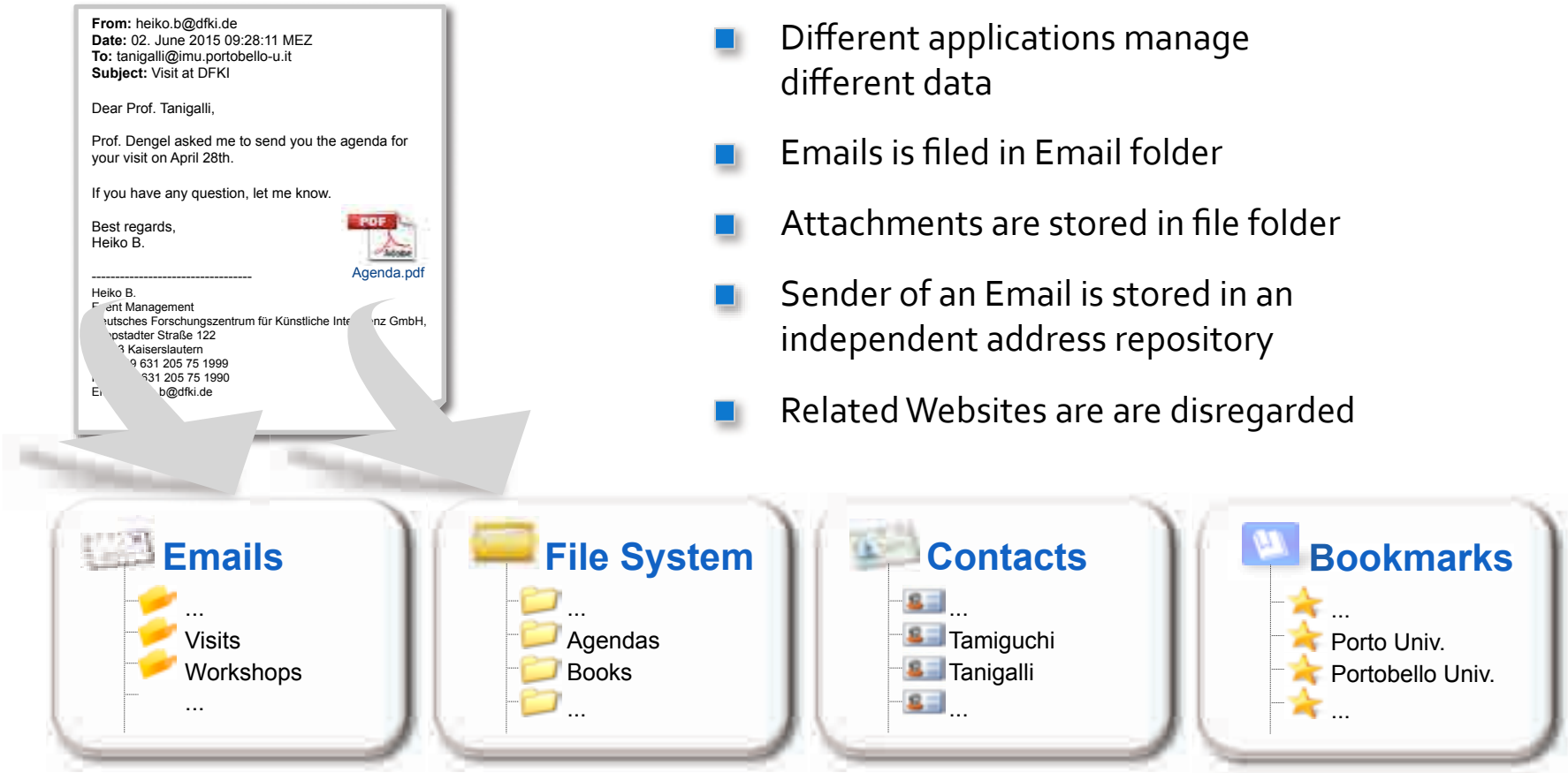
Contact data

Prof. Simone Tanigalli

Simone TANIGALLI
Professor
Image and Media Understanding Lab
Portobello University
T +39 ...
F +39-....
E tanigalli@imu.portobello-u.it

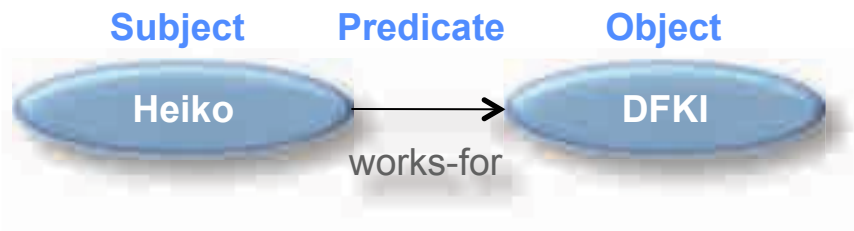
- For solving his tasks, Heiko has to consider various aspects of information from different applications

The limits of today's desktop information management lead to a cut between mental models and document contents



- Different applications manage different data
- Emails is filed in Email folder
- Attachments are stored in file folder
- Sender of an Email is stored in an independent address repository
- Related Websites are are disregarded

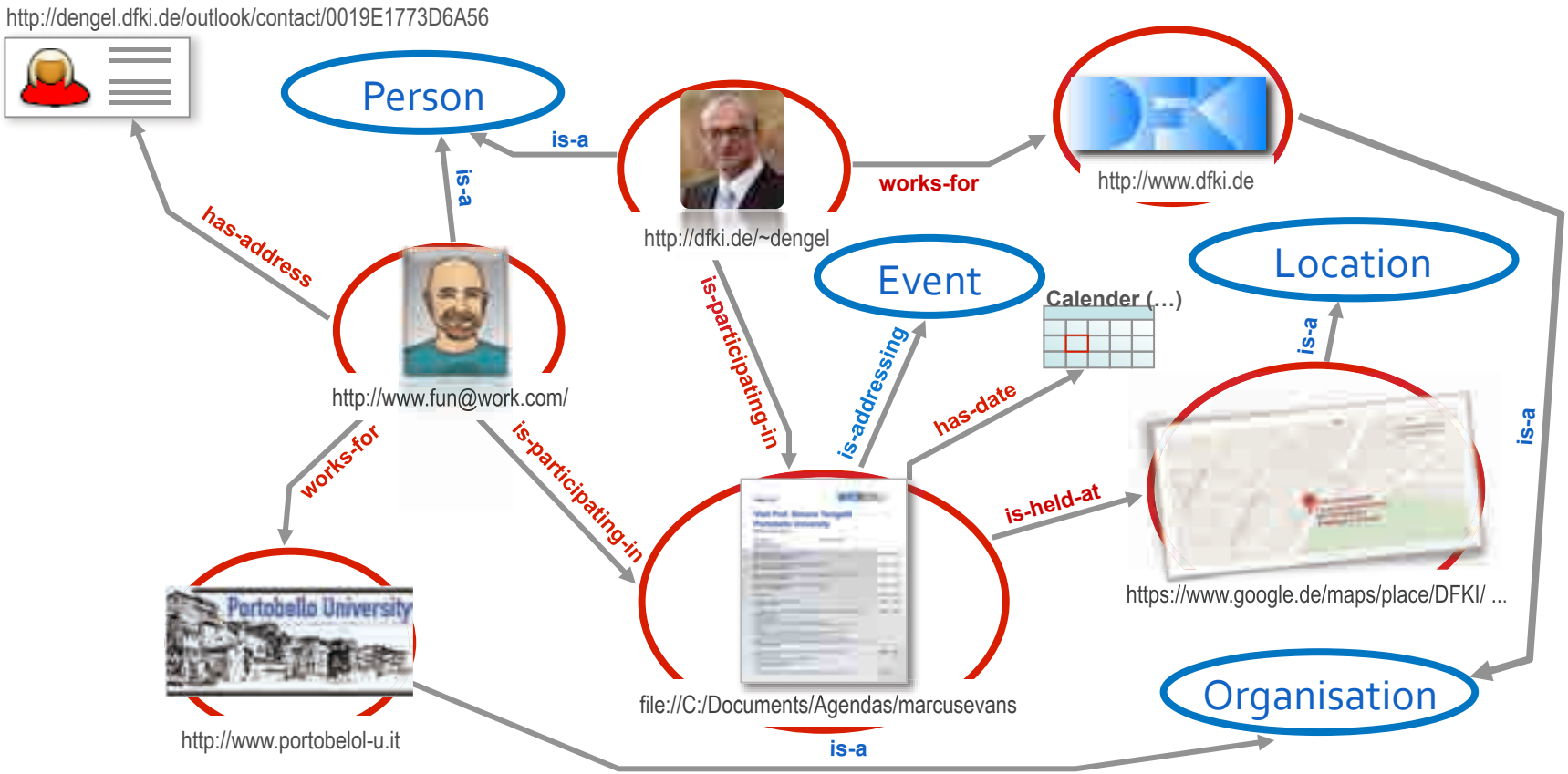
Heiko's Semantic Desktop assists him in remembering the multi-perspective aspects of his work



- **Information objects** at the workspace **are semantically interconnected**
- The Semantic Desktop makes use of **WWW-Standards** for describing meaning by simple sentences (**Subject-Predicate-Object**)
- The uniqueness of resources is guaranteed by using their URI (**Uniform Resource Identifier**)
- All **information objects** (resources) **may be categorized** according to a class such as Person, Event, Locations, Topics, Tasks or Organization

(*see next slide*)

Based on such a consideration, Heiko establishes a Personal Information Model (PIMO)

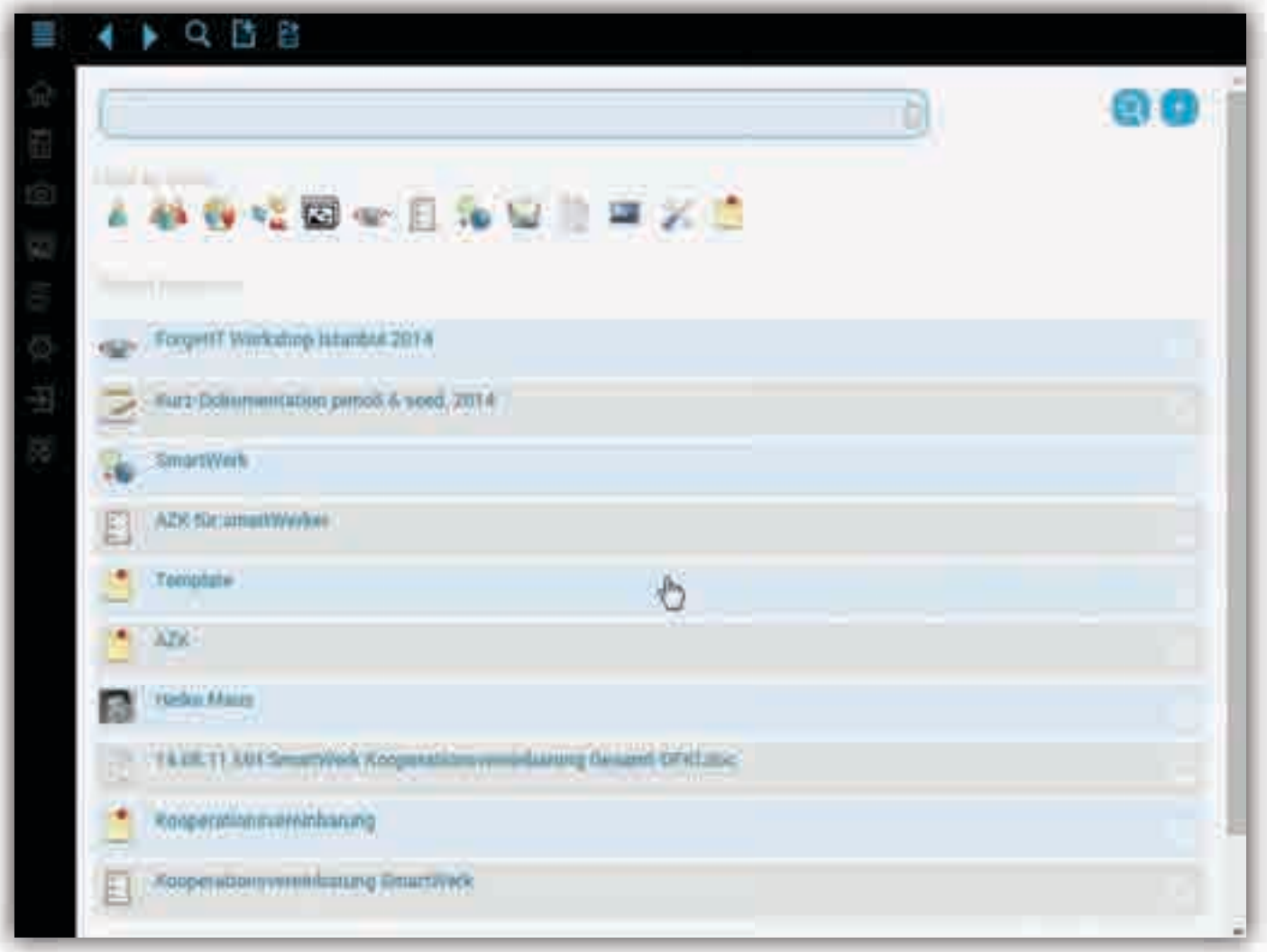


The Semantic Desktop acts as an anticipating information butler

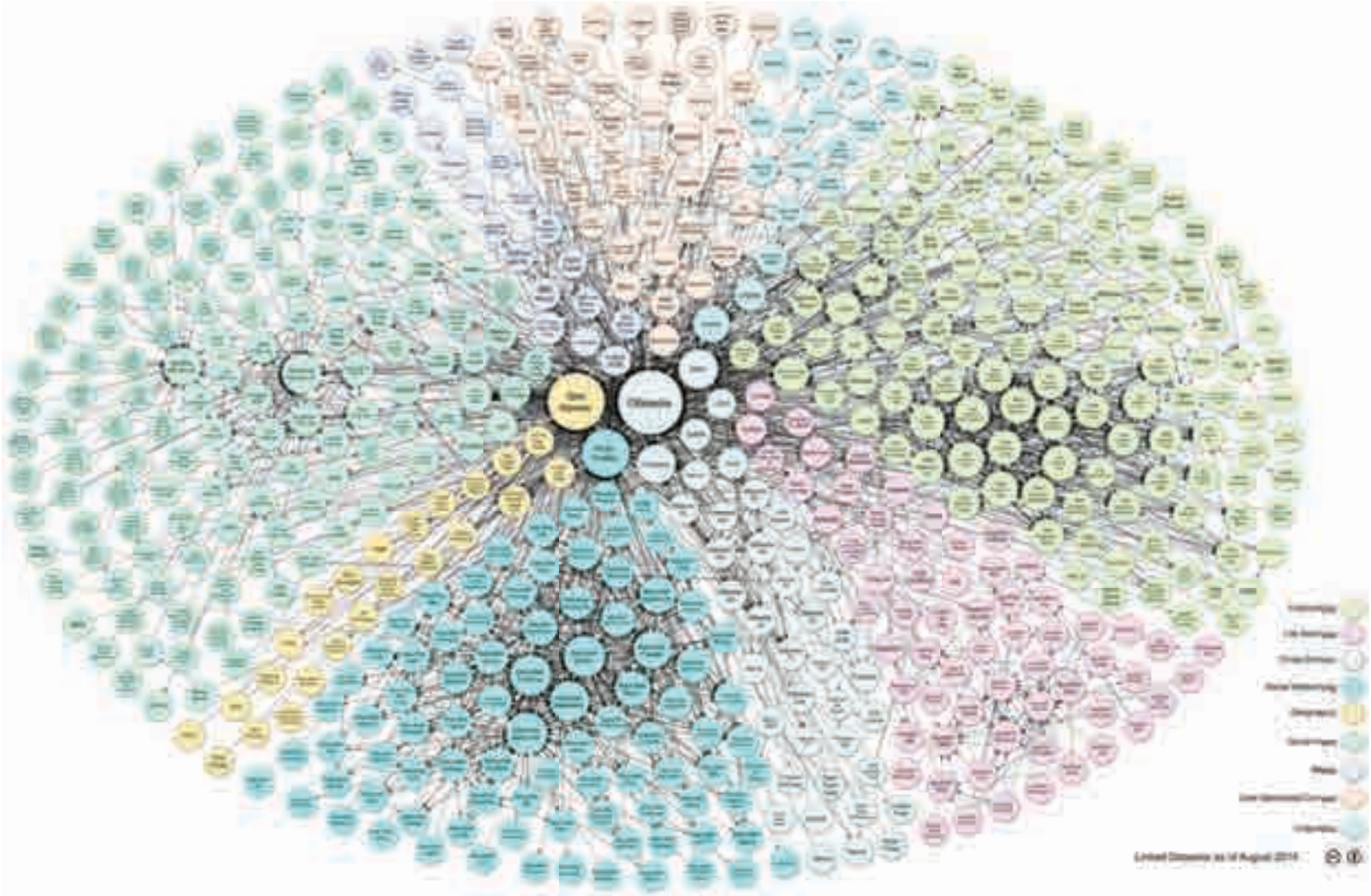
Let's go back to July 29th, 2014

- Besides his many other tasks, Heiko is the coordinator of the EU project ForgetIT
- For that reason he is planning to have a workshop in Istanbul, Turkey on July 30th, 2014
- One day before the workshop he is preparing some topics he intends to talk about
- While doing so, his PIMO is continuously employed to support him

… and is used in daily life (at least within DFKI)



A PIMO may be connected to the Linked Open Data Cloud



Augmenting Human Mind means identifying what might be relevant in a given situation

Attention Recognition

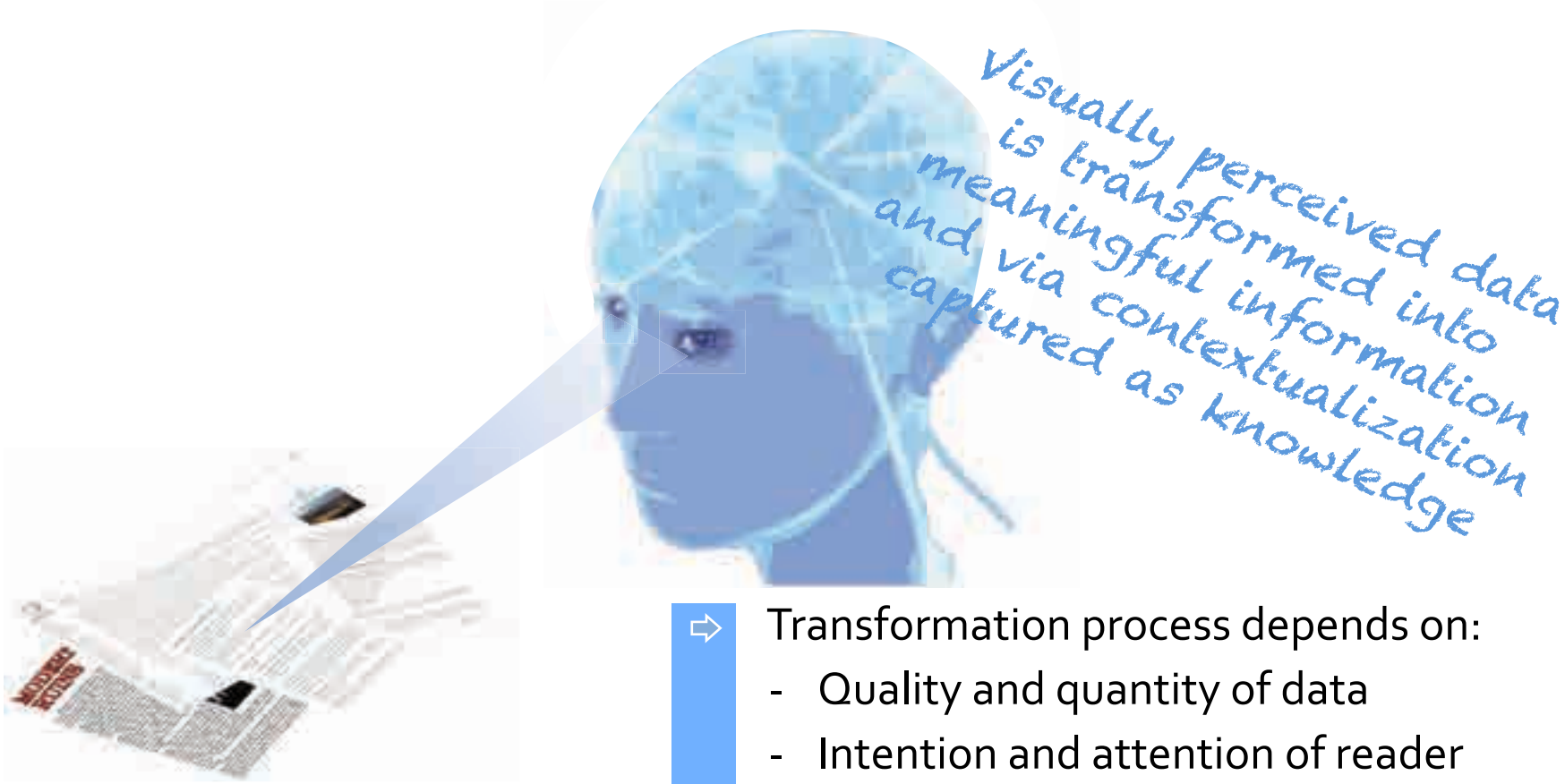
R. Biedert, G. Buscher and A. Dengel, *The EyeBook – Using Eye Tracking to Enhance the Reading Experience*, in: Informatik Spektrum 33/3, Springer Publ. (June 2010), pp. 272-281.

G. Buscher, A. Dengel, R. Biedert, and L. van Elst, *Attentive Documents: Eye Tracking as Implicit Feedback for Information Retrieval and Beyond*, ACM TiiS, Transactions on Interactive Intelligent Systems Vol. 1, No. 2, (Jan 2012).

Augmenting our mind strongly depends on intrinsic motivation and on cognitive attention to understand and learn

Cognitive attention first of all means visual attention!

Experts estimate that more than 80% of information transmitted to the brain during academic learning is transmitted visually i.e. by use of our eyes!



Transformation process depends on:

- Quality and quantity of data
- Intention and attention of reader
- Ability to establish a shared context

* Picture partially adapted from
https://www.linkedin.com/pulse/20130211201055-175081329-12-minutes-to-create-a-mind-changing-presentation

Let's start with some basics!

Our daily illusion!

Our perceptual abilities are not continuously the same!

Fixations: 100 … 1000ms (200-300 avg.)

Saccades: 30 - 80 ms

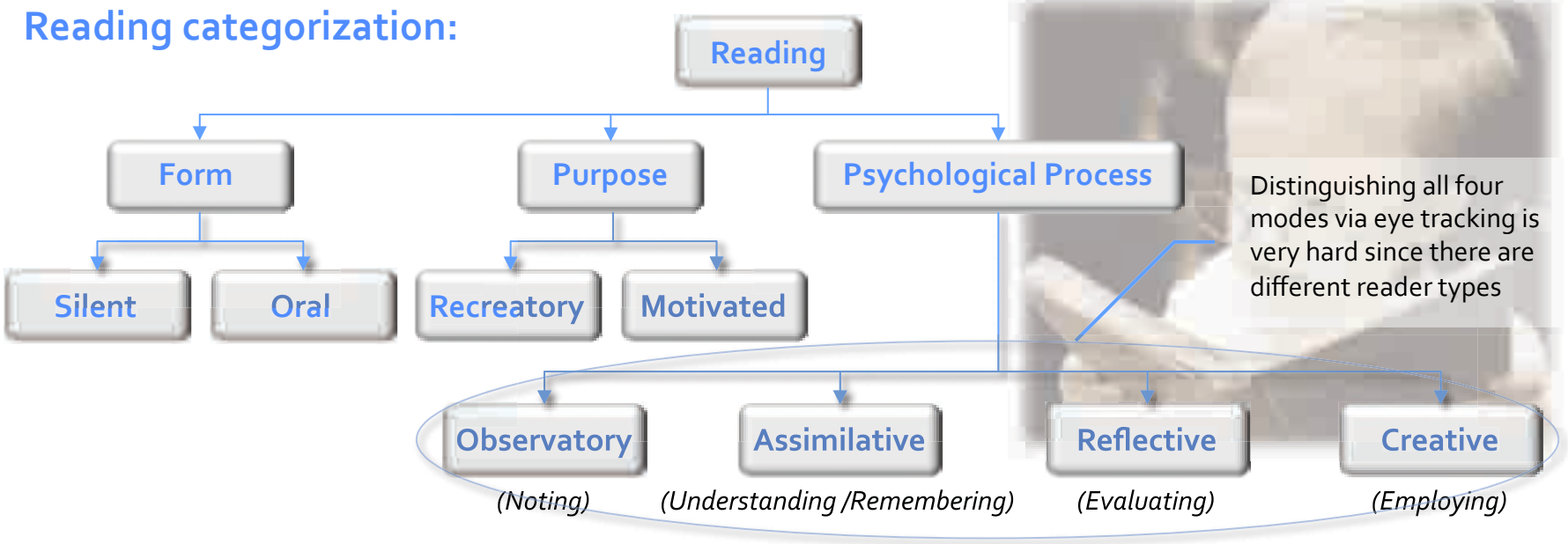
Saccadic Suppression!

Our approach was to use eye-trackers to measure visual attention and employ them for vivid interaction

- Gaze data is collected using either a remote or head-mounted eye-tracker
- A (infrared) light source is directed toward the eye
- Camera tracks the reflection of the light source along with visible ocular features
- Data is used to extrapolate the rotation of the eye and ultimately the direction of gaze
- Visual path is analyzed across an interface and transformed it into a set of pixel coordinates, i.e.
  - which features are seen, when a particular feature captures attention, how quick is the eye moving, what content is overlooked, etc.

Attention data can be employed for example to understand what is relevant to a reader

Reading categorization:



Distinguishing all four modes via eye tracking is very hard since there are different reader types

- Skimming addresses a quick movement of the eyes across the page, picking up the occasional observation or idea
  - … when we try to get an orientation

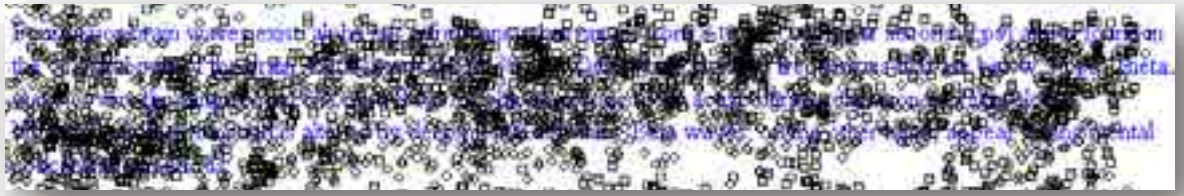
Not (that) relevant

- Reading has process every sentence, and then try make use of the salient arguments
  - … when we know we'll later profit from the material

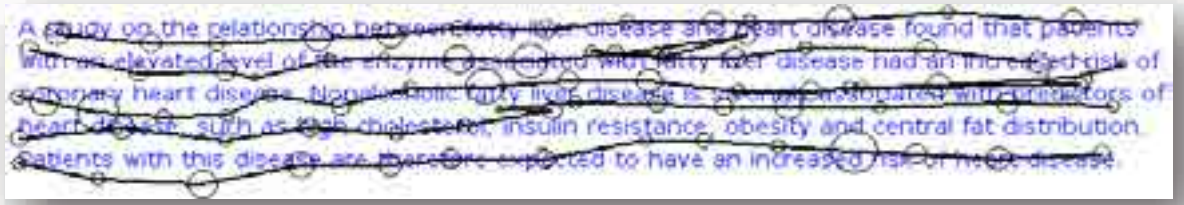
Relevant

For reading mode detection we use various filters along a processing chain to distinguish reading from skimming behavior

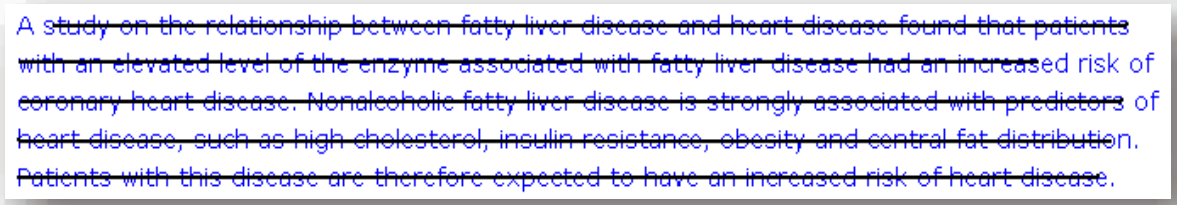
Noisy gaze data from the eye tracker



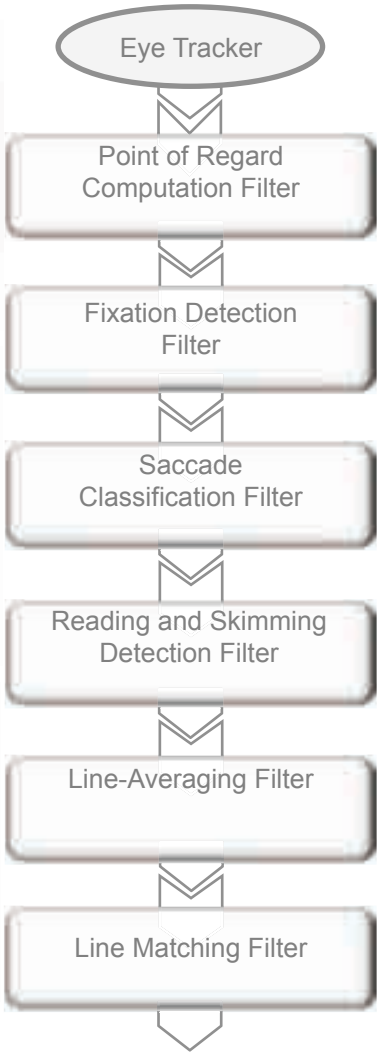
Fixation detection and saccade classification



Reading identification and saccade sequence alignment



Line-matching by mapping with line segmentation results (plus OCR)



Results can be represented by gaze-based document meta data

Store reading information as document annotations in a semantic Wiki

[Rayner 1998], the eye shows a very characteristic behavior composed of fixations and saccades. A fixation is a time of about 250ms on average when the eye is steadily gazing at one point. A saccade is a rapid, ballistic eye movement from one fixation to the next. The mean left-to-right saccade size is 7-9 letter spaces. It depends on the font size and is relatively invariant concerning the distance between the eyes and the text.

Annotation (Read) Delete
author: Georg
start date: 07.12.2014 10:46:08
End date: 07.12.2014 10:46:12
length: 226 chars
mean fixation duration: 217ms
mean saccade length: 9.4 chars
regression ratio: 13.9%
task: ***write report***

An enormous amount of research has been done during last one hundred years concerning eye movements while reading. When reading silently, as summed up in [Rayner 1998], the eye shows a very characteristic behavior composed of fixations and saccades. A fixation is a time of about 250ms on average when the eye is steadily gazing at one point. A saccade is a rapid, ballistic eye movement from one fixation to the next. The mean left-to-right saccade size is 7-9 letter spaces. It depends on the font size and is relatively invariant concerning the distance between the eyes and the text.

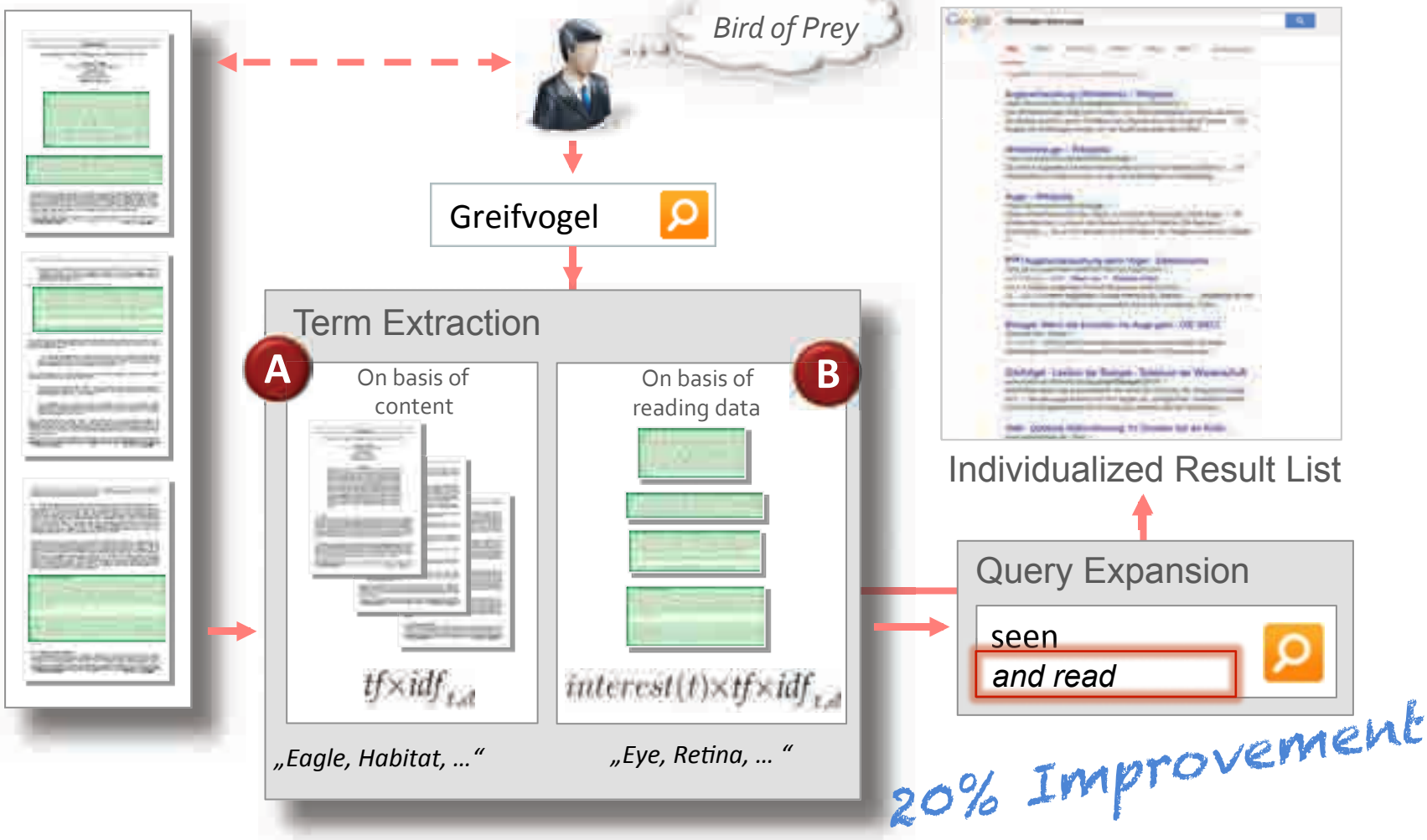
G. Buscher, A. Dengel and L. van Elst, *High Level Eye Movement Measures for Relevance Assessments of Information Items*, Proceedings CHI 2008, Florence, Italy (Apr. 2008).

R. Biedert, J. Hees, and A. Dengel, and G. Buscher, *A Robust Realtime Reading-Skimming Classifier*, Proceedings 7th Biennial Symposium on Eye Tracking Research & Applications, Santa Barbara, CA, USA (March 2012), pp. 123-130

This approach may be used for improving relevance feedback

Application Example: Relevance of Search Results (I)

Application Example: Relevancy of Search Results (II)



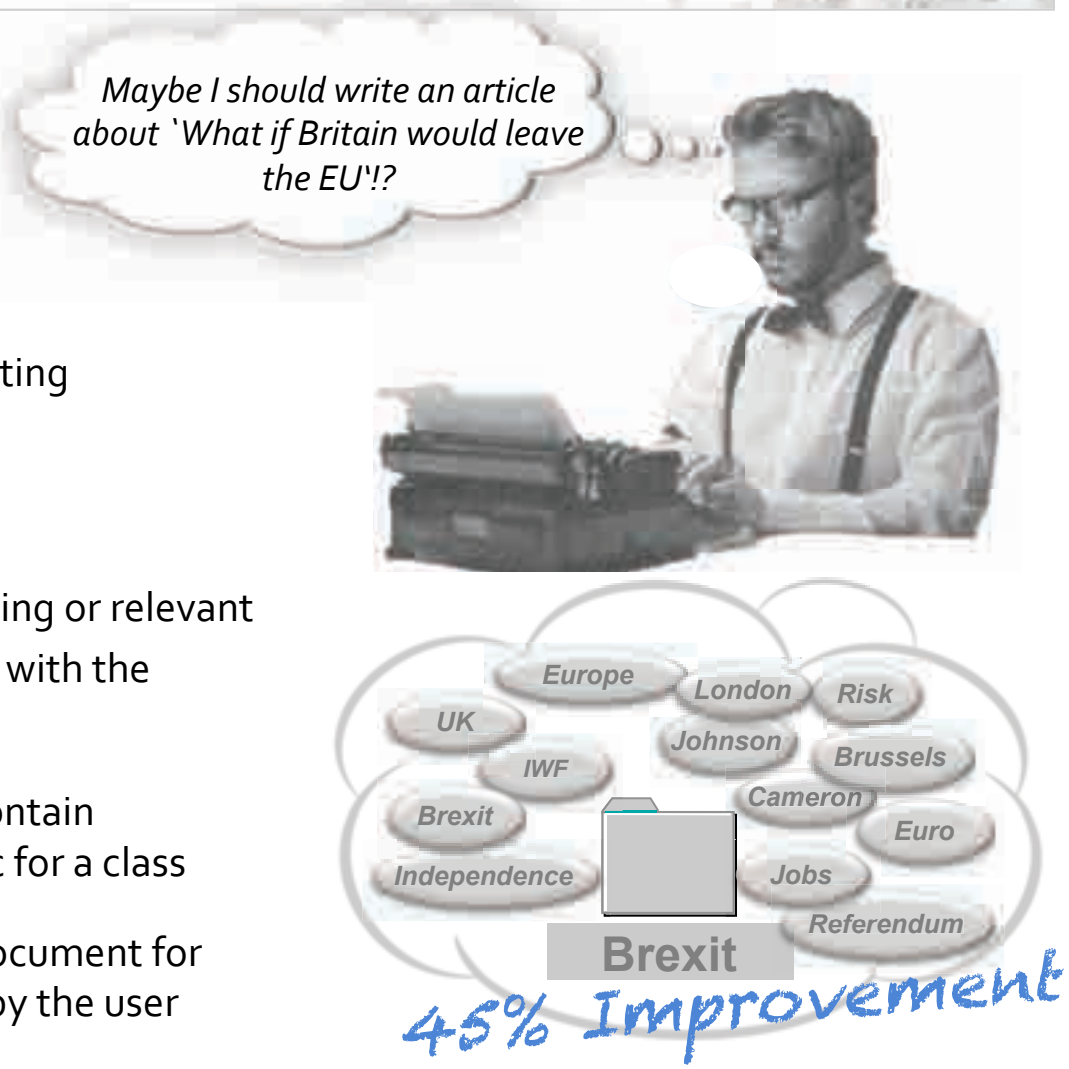
G. Buscher, A. Dengel and L. van Elst, *Query Expansion Using Gaze-Based Feedback on the Subdocument Level*, Proceedings SIGIR '08, 31st Annual Int'l ACM SIGIR Conference, Singapore, (July 2008).

… or for improving classifier learning

Application Example: Document Classification

- Classification is done by manually moving documents into a folder
- Classification is based on subjective perception of content
- For making her/his decision respecting categorization, a user...
  - ... reads some passages
  - ... skims over others
  - ... skips parts that are not interesting or relevant

  according to the her/his familiarity with the sources, her/his interest, ...
- All documents within one folder contain terminology which is characteristic for a class
- Only consider those parts of the document for classifier learning, which are read by the user



G. Buscher and A. Dengel, *Attention-Based Document Classifier Learning*, Proceedings DAS08, Int'l IAPR Workshop on Document Analysis Systems, Nara, Japan, IEEE Comp. Society Press (Sep. 2008), pp. 87-94

… but there are also new applications in infotainment

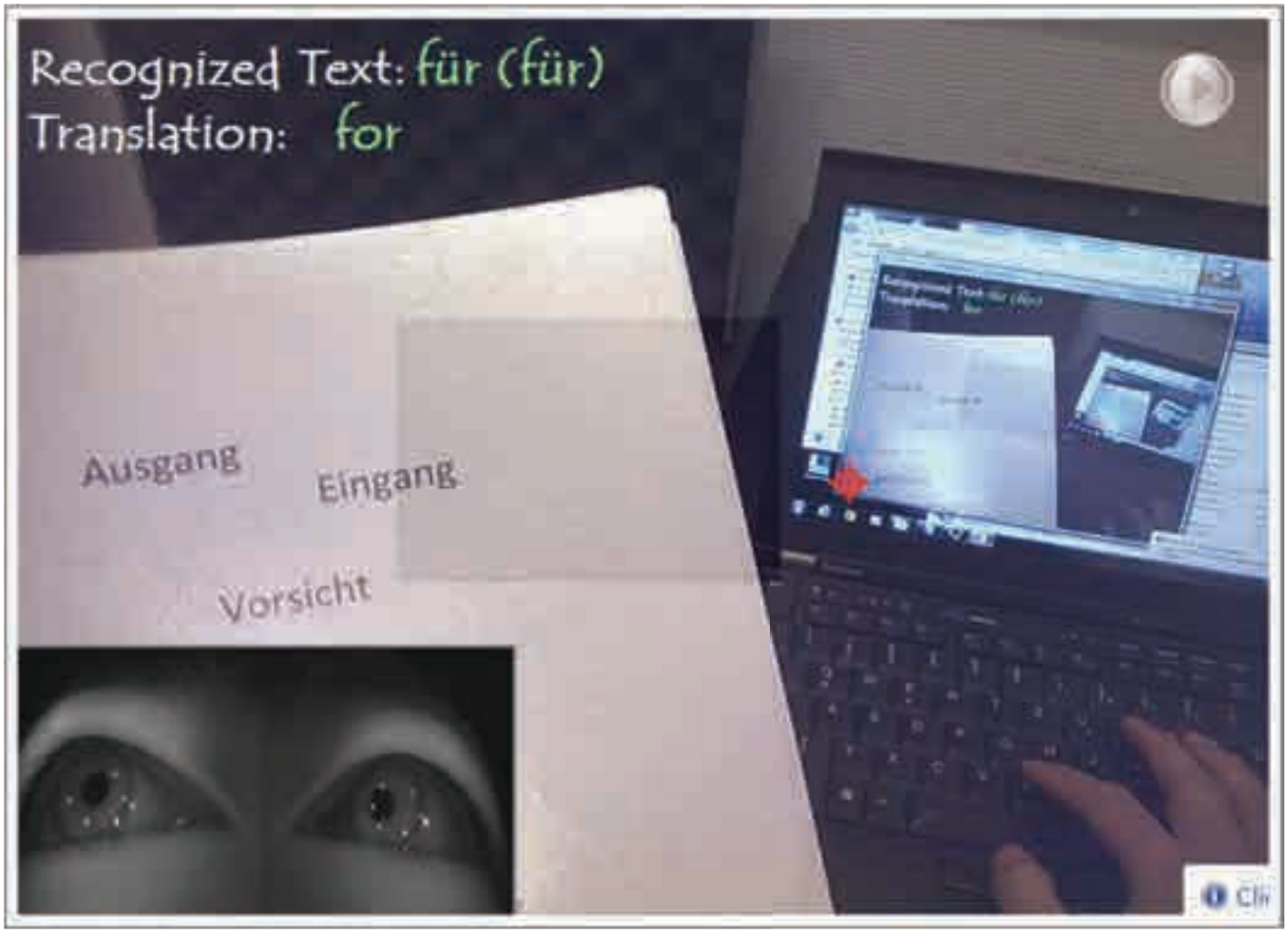
Imagine there were input devices which could allow text to know if and how it is read

- When it was introduced, Text 2.0 was an innovative interaction mode between humans and computer
- It is build on the idea that the computer knows on which text line, sentence, or word a person looks
- It supplements the text by hidden "attentive mark-ups" that are activated during reading, i.e. recognizing a specific reading mode
- Reveals new business options, .e.g. in online marketing and advertisement

R. Biedert, G. Buscher and A. Dengel, *The EyeBook – Using Eye Tracking to Enhance the Reading Experience*, in: Informatik Spektrum 33/3, Springer Publ. (June 2010), pp. 272-281 .

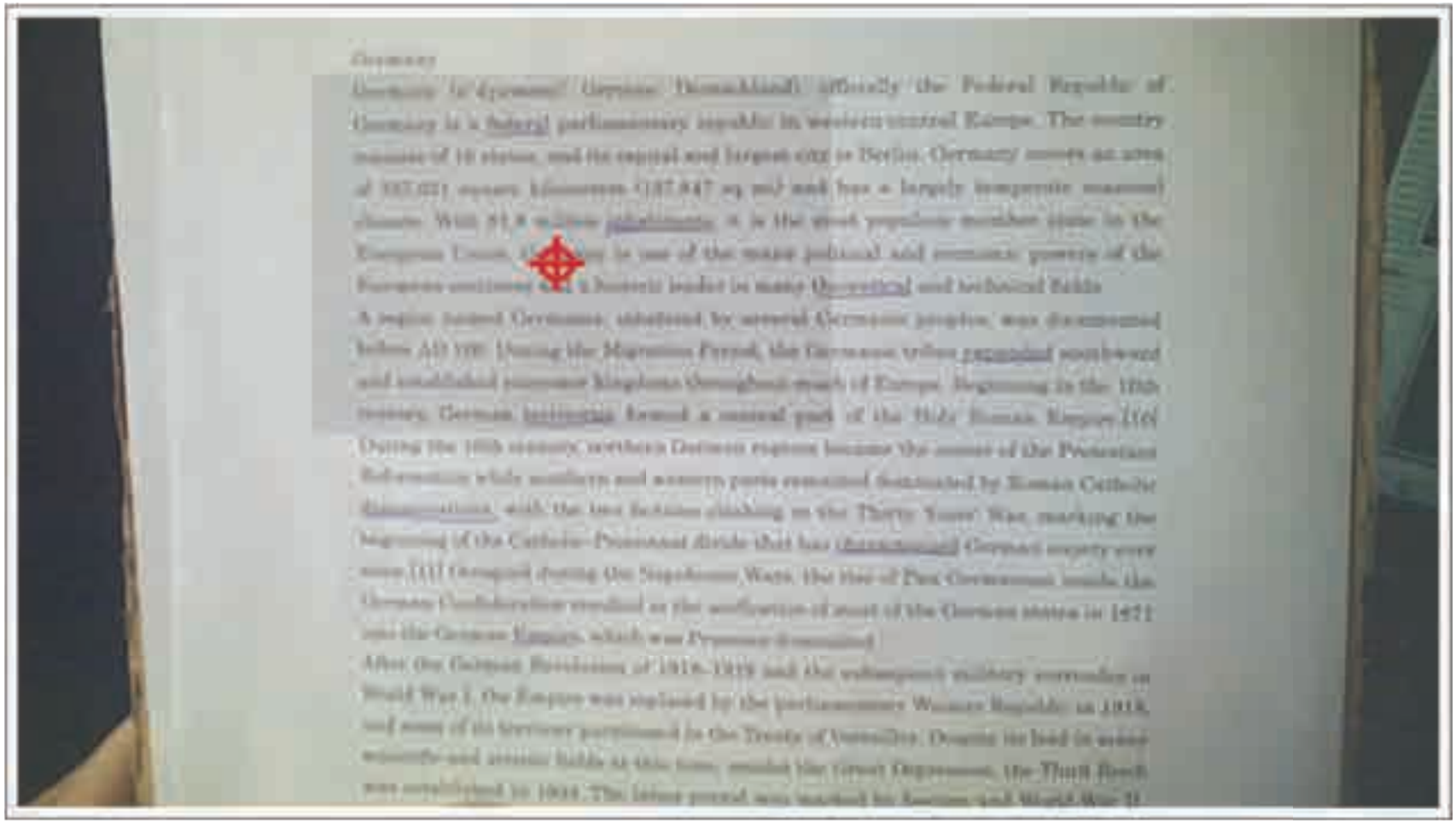
Text 2.0 was one of selected megatrends on Google's Zeitgeist conference

We employed the idea for mobile eye-trackers with head-mounted display



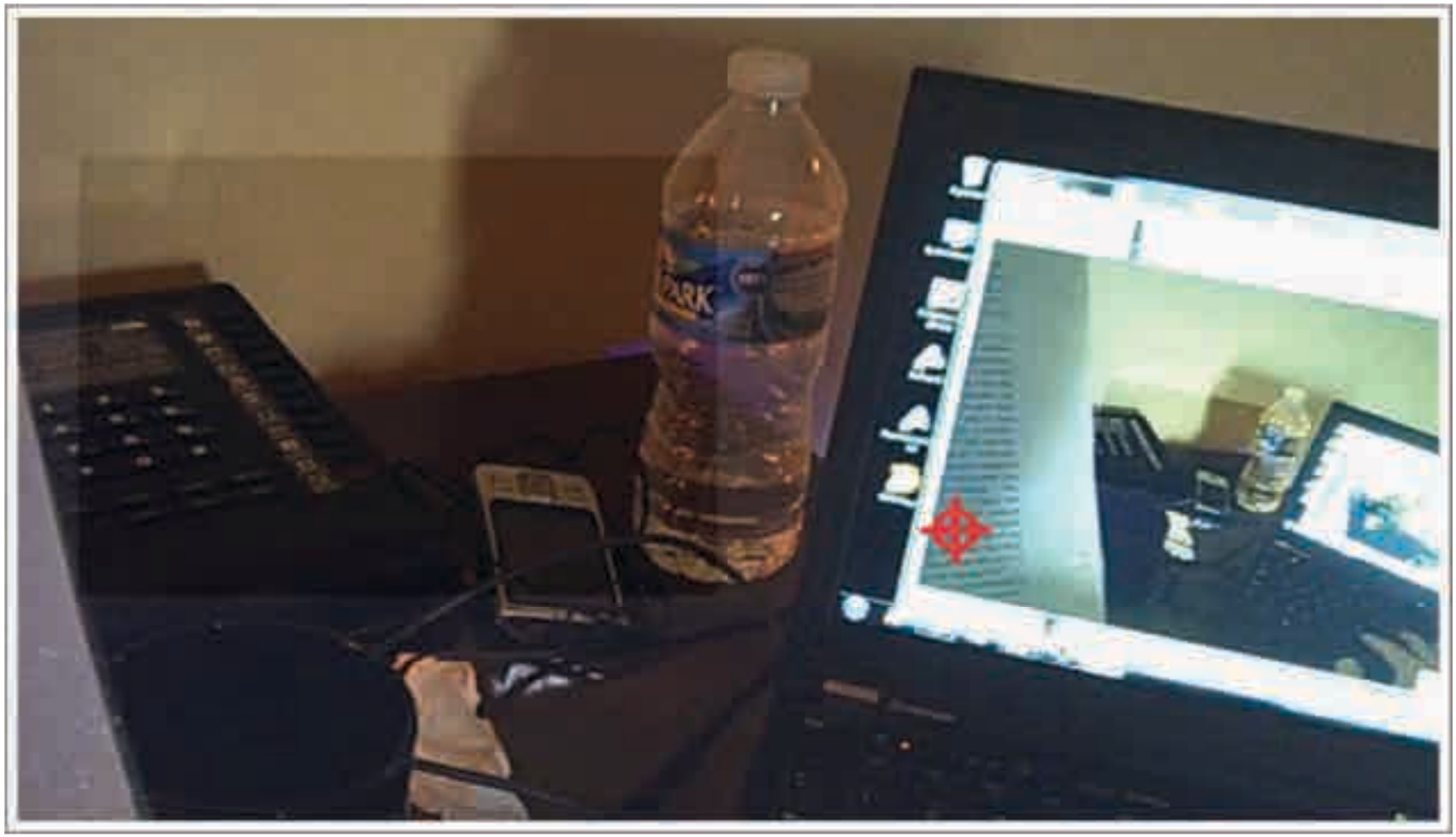
T. Toyama, D. Sonntag, A. Dengel, T. Matsuda, M. Iwamura, and K. Kise, *A Mixed Reality Head-Mounted Text Translation System Using Eye Gaze Input*, Proceedings IUI 2014, 19th Int'l Conf. on Intelligent User Interfaces, Haifa, Israel (Feb. 2014), pp. 329-334.

We employed the idea for mobile eye-trackers with head-mounted display



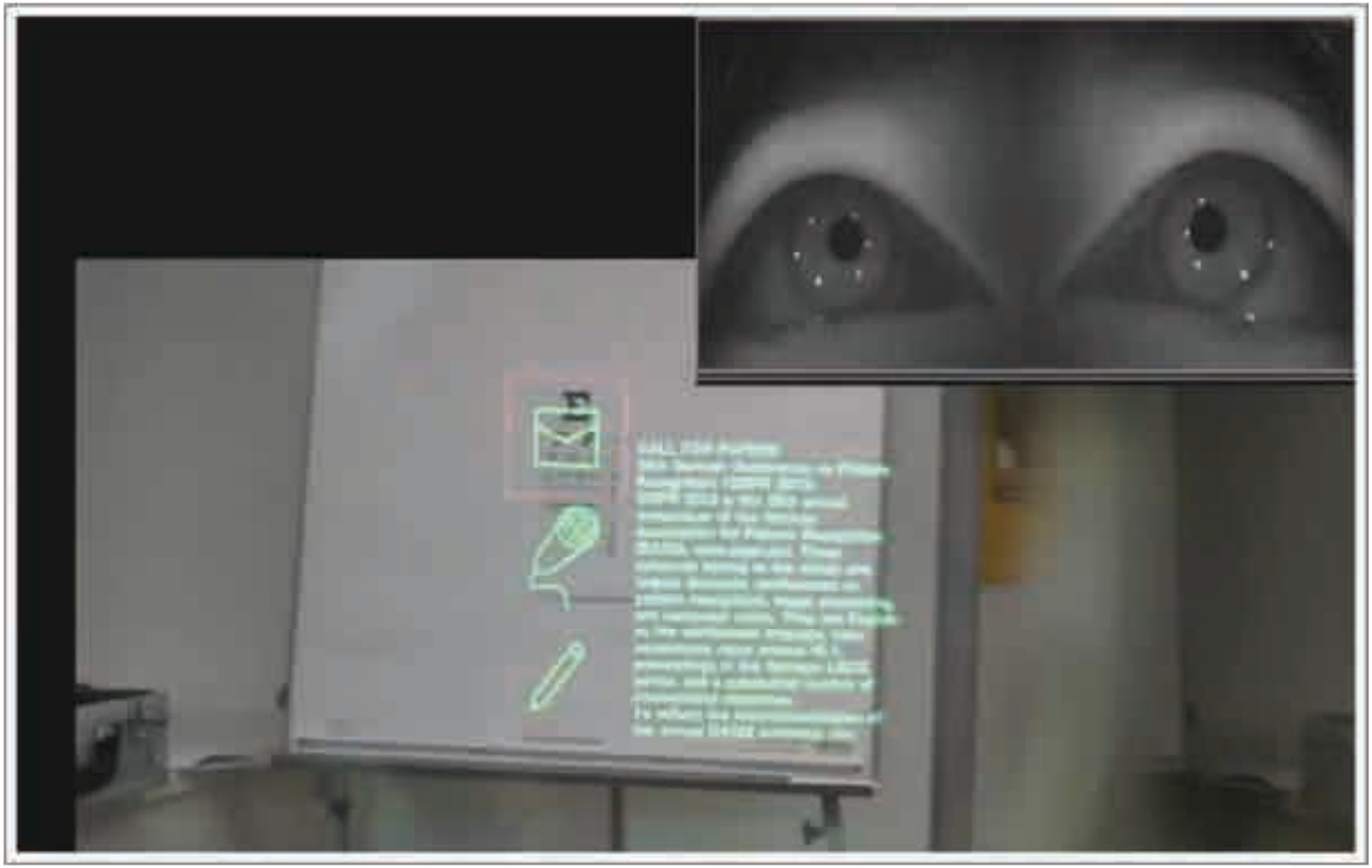
T. Toyama, D. Sonntag, A. Dengel, T. Matsuda, M. Iwamura, and K. Kise, *A Mixed Reality Head-Mounted Text Translation System Using Eye Gaze Input*, Proceedings IUI 2014, 19th Int'l Conf. on Intelligent User Interfaces, Haifa, Israel (Feb. 2014), pp. 329-334.

We employed the idea for mobile eye-trackers with head-mounted display



T. Toyama, A. Dengel, W. Suzuki, and K. Kise, *User Attention Oriented Augmented Reality on Documents Using a See-through HMD and a Wearable Eye Tracker*, submitted to ISMAR 2013, Symp. on Mixed and Augmented Reality, Adelaide, Australia (Oct. 2013), pp. 299-300.

We employed the idea for mobile eye-trackers with head-mounted display

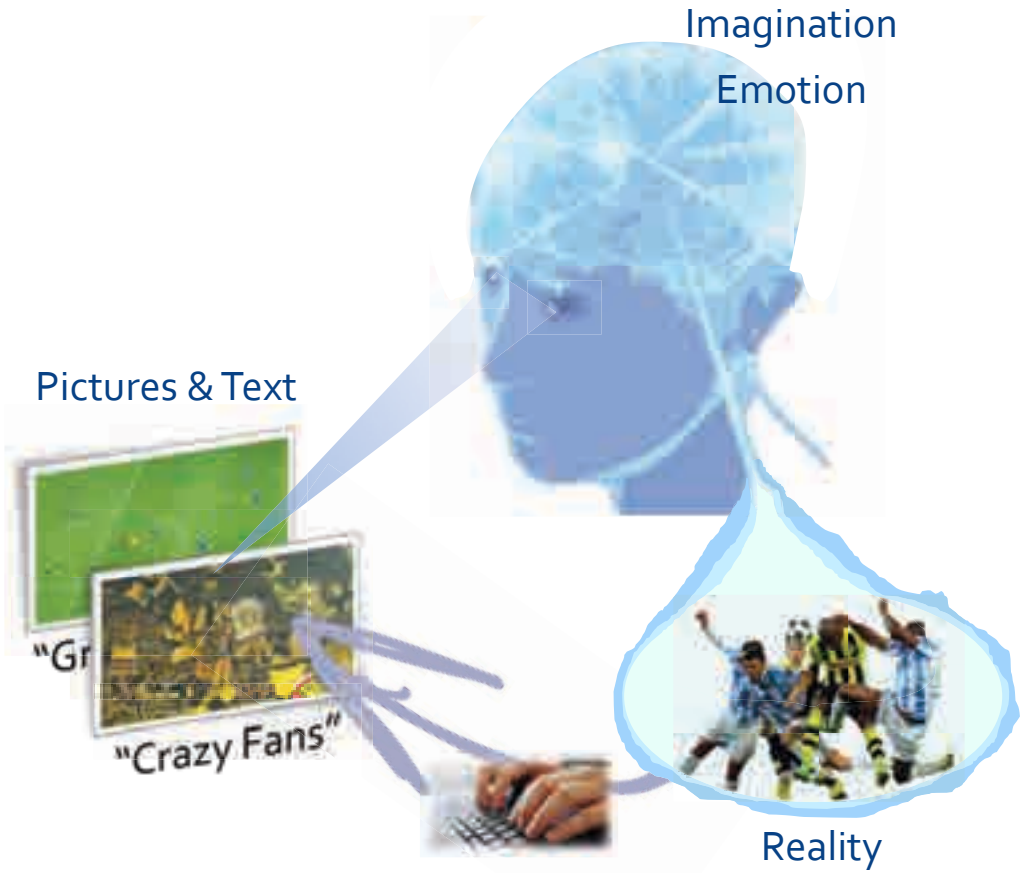


Augmenting Human Mind means understanding human's emotions

Visual Sentiment Recognition

* Picture adapted from http://blog.thepitagroup.com/wp-content/uploads/2013/03/sentiment-faces-analysis.jpg

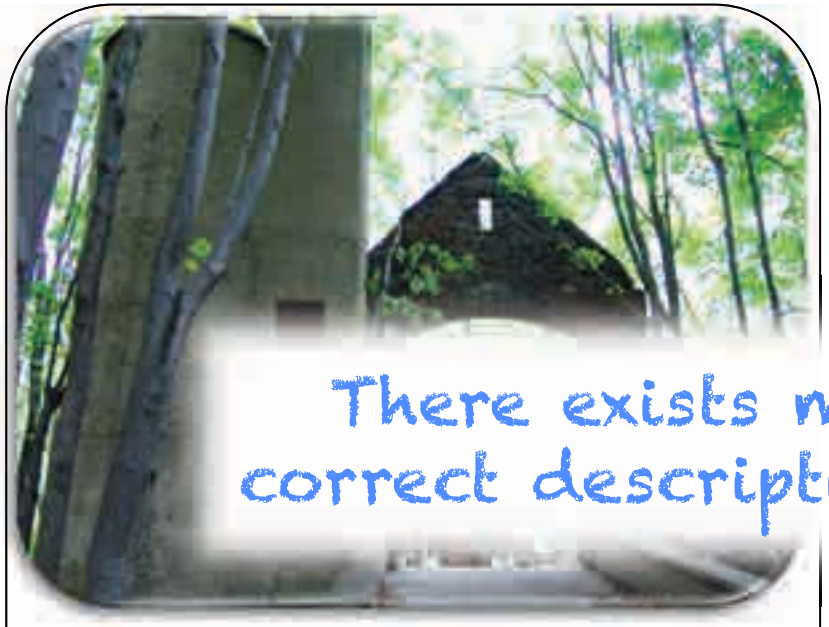
Relationship between a human, the real world, and a picture/ text can be described via the Semiotic Triangle



- Our environment consists of items, facts and events that are „real" and determine our lives („what is going on")
- To capture our thoughts, we take pictures and label them with text („what is couched or explicate")
- Looking at the picture and reading the text put contents together and create very individual imaginations („what I have in mind")

* Picture partially adapted from https://www.linkedin.com/pulse/20130211201055-175081329-12-minutes-to-create-a-mind-changing-presentation

What are these pictures showing?



There exists more than one correct description of an image

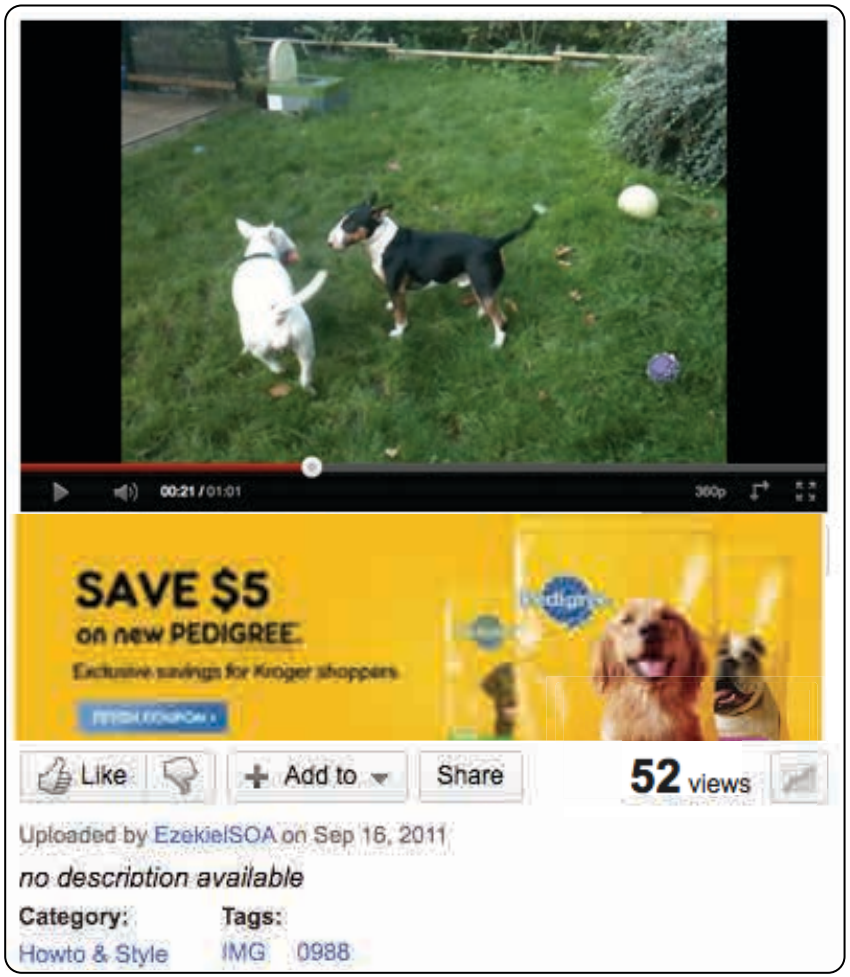
flickr user tag "abandoned building"

flickr user tag "stormy landscape"

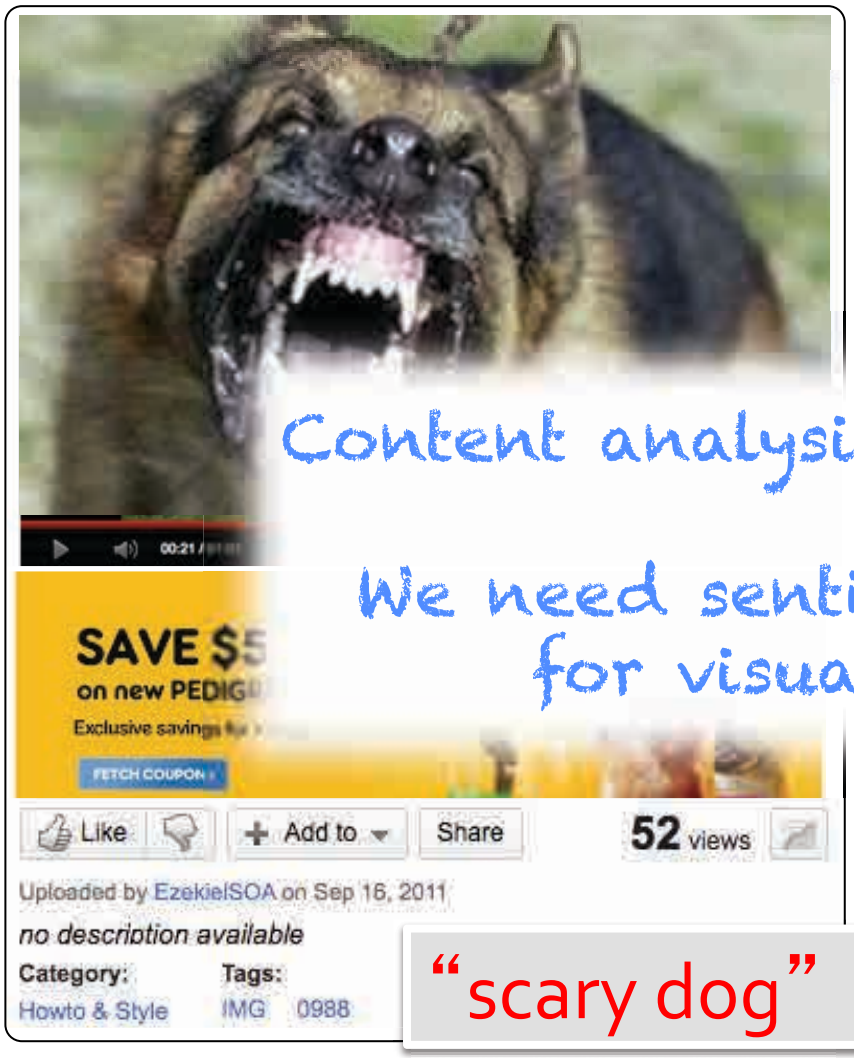
Social Media – the idea – the outcome

Social Media – the idea – the outcome

Ads placement – the idea – the outcome

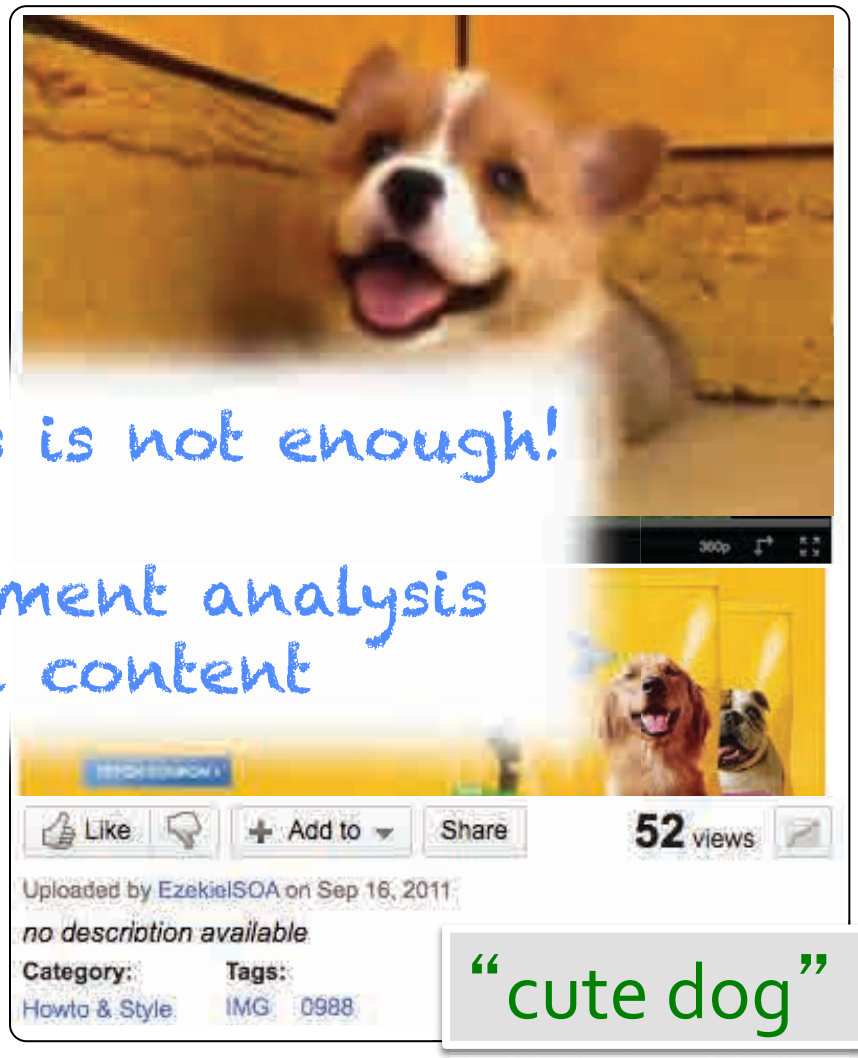


Ads placement – the idea – the outcome



Content analysis is not enough!

We need sentiment analysis for visual content

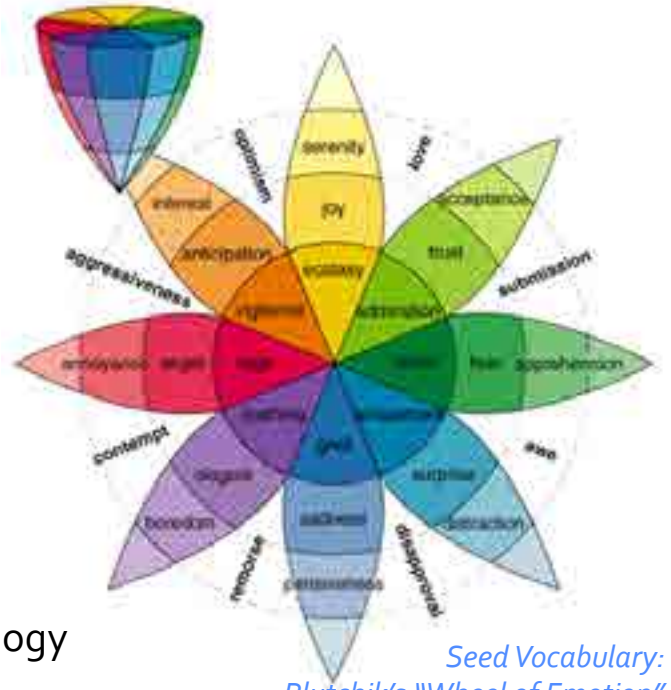


"scary dog"

"cute dog"

We implemented a visual sentiment analysis approach based on various findings

- Predict sentiment by "understanding" visual content
  - Utilize advanced **Deep Learning** approaches
- Introduce **Visual Sentiment Ontology** & SentiBank
  - Concepts are linked to emotion
  - Reflect strong sentiment
  - Are frequent on Flickr & Youtube
  - Reasonable detection performance
- Employ primary **bipolar emotions model** from psychology



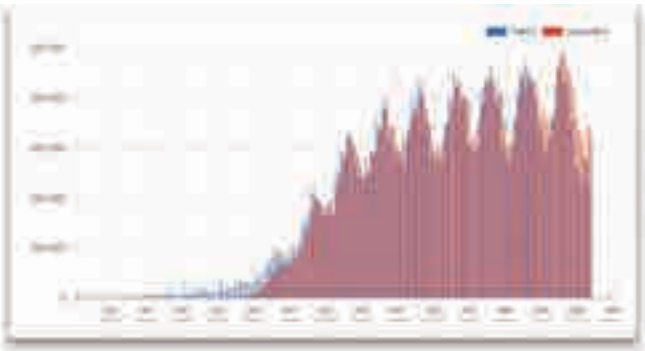
Seed Vocabulary: Plutchik's "Wheel of Emotion"

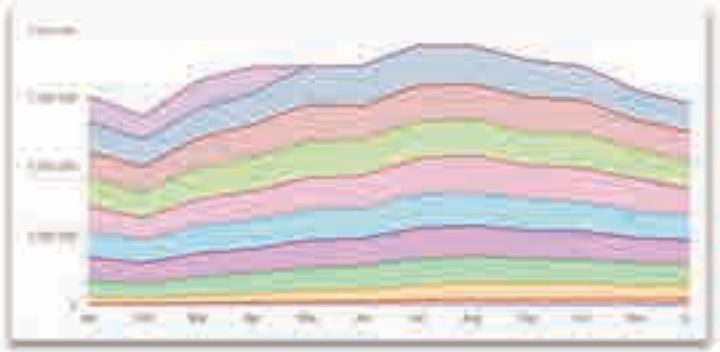
For training our network we employed the largest dataset in computer vision and multimedia research

Yahoo Flickr Creative Common 100 Million (YFCC100m)

YAHOO! flickr

- 100 Millions images and videos
  - ~80 Millions with GEO data
  - ~500 Million tags
  - ~225 Million words in titles
  - ~680 Million words in descriptions





S. Kalkowski, C. Schulze, D. Borth, and A. Dengel, *Real-Time Analysis and Visualization oft he YFCC100m Dataset,* Proceedings Workshop on Multimedia Commons, ACM Multimedia 2015, Brisbane, Australia (Oct. 2015), pp. 25-30.

Factorized Neural Nets allows classifies Adjective-Noun-Pairs and simulates of the subjective perception of humans

Deep CNNs (Deep Learning)

1. little church
2. ancient house
3. damaged church
4. ancient bridge
5. ancient church

Deep CNNs (Deep Learning)

1. stormy mountain
2. stormy coast
3. stormy waves
4. misty hills
5. stormy clouds

https://deepsentibank.appspot.com

Mapping Images to Sentiment Adjective Noun Pairs with Factorized Neural Nets, IEEE CVPR, 2016 (submitted)

We further extended DeepSentiBank towards a full CNN

Deep CNNs (Deep Learning)

1. stormy mountain
2. stormy coast
3. stormy waves
4. misty hills
5. stormy clouds

… and some more examples

S. Mozaffari , M.A. Naser, S. Bukhari, D. Borth, and A. Dengel, *What makes a Beautiful Landscape beautiful: Adjective Noun Pairs Attention by Eye-Tracking and Gaze Analysis has been updated by Syed Saqib Bukhari,* Proceedings Workshop on Affect and Sentiment in Multimedia (ASM), ACM Multimedia 2015, Brisbane, Australia (Oct. 2015), pp. 51-56.


Home :: Ontology :: Adjective Noun Pairs :: Downloads :: About
Visual Sentiment Ontology
Ontology :: Treemap Visualization
SENTIMENT
NEGATIVE
NEUTRAL
POSITIVE
JOY
HAPPY
BEAUTIFUL
TRADITIONAL
LITTLE
COLD
PRETTY
DRY
BIRTHDAY
FAMILY
FACE
CHRISTMAS
LANDSCAPE
SUNSET
SKY
CLOUDS
NATURE
SMILE
DANCE
GIRL
ICE
WINTER
SNOW
GREAT
CLOUDY
WET
YOUNG
WARM
FRESH
LADY
HALLOWEEN
MOTHER
DOG
CAT
BABY
WEDDING
KIDS
LOVELY
VIEW
SEXY
PEACEFUL
GOLDEN
MISTY
WILD
STUNNING
FANCY
FAVORITE
YUMMY
SMOOTH
SWEET
MUDDY
DRAGON
POND
SUNRISE
BEACH
FLOWER
SUN
BAY
STAR
FLOWERS
CALM
SCENIC
FALLING
RAINY
SERENE
MAGICAL
NATURAL
CLEAR
BEAUTY
WATER
LAKE
SUNNY
COLORFUL
HEALTHY
MISTY WOODS
Images: 8,040
Accuracy: 100%
http://www.sentibank.org
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
Digital Video and Multimedia Lab
Multimedia Analysis and Data Mining

Visual Sentiment Ontology

Adjective Noun Phrases :: Sorted List

Search:

Adjectives: beautiful | Nouns: all | Emotions: all | filter detectable | 1 of 2 [total: 22 ANPs]

Name	Sentiment	# Images	Accuracy	Sample Visualization	Top Emotions
beautiful flower	1.77	2,380,000	80.46		2. serenity 3. joy 4. interest 8. anticipation
beautiful girl	1.67	1,110,000	84.12		2. sadness 3. joy 4. serenity 5. interest 6. pensiveness
beautiful sunset	1.71	946,000	88.1		2. serenity 3. interest 4. joy 6. pensiveness
beautiful sky	2.0	703,000	85.0		2. serenity 3. interest 4. joy 6. pensiveness

yfcc100m Browser

Browse

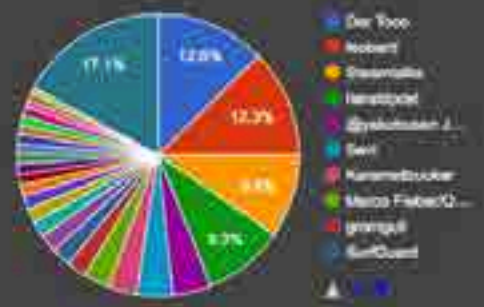
YFCC100M Image Browser ::

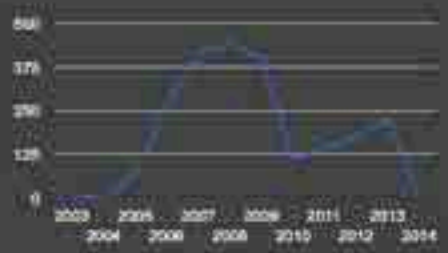
kaiserslautern

Find

Download Results as csv

kaiserslautern(+) germany(+) fußball(+) betzenberg(+) fck(+) fritz-walter-stadion(+) deutschland(+) soccer(+) 1. fc kaiserslautern(+) rote teufel(+) football(+) stadion(+) dfki(+) bundesliga(+) lautern(+) deutsches forschungszentrum für künstliche intelligenz(+) land der ideen(+) 365 orte der ideen(+) 1. fck(+) 2008(+) k-town(+) dfl(+) pfalz(+) europa(+) sport(+) westkurve(+) 2. liga(+) 2013(+) japanese(+) garden(+) leinespiel(+) day(+) betzy(+) familie(+) dcw(+) 2. bundesliga(+) 2007(+) summer(+) japan(+) stadium(+) auswärtssieg(+) family(+) may(+) mai(+) sommer(+) army(+) 2006(+) 365(+) ununterschieden(+) vacation(+)





http://www.yfcc100m.org

Thank you, … questions?

Address:
Prof. Andreas Dengel
DFKI GmbH
Trippstadter Straße 122
67608 Kaiserslautern
email: andreas.dengel@dfki.de
http://www.dfki.de/~dengel