

welcome!

*Jörg Schütz*

*Sven C. Andrä*

*biloom* | *group*

**andrä**  
**ag**



agenda

*Employing Machine Translation  
in Globalization Tasks*



agenda

*Part 1: changing world*

*Part 2: how MT fits in*

*Part 3: holistic solution*



knowledge worker

- creativity
- emotions
- knowledge
- experience

**Part 1:**  
**Enterprise 2.0**

Web 2.0 Globalization  
Post-Taylorism  
Internationalization  
(Machine) Translation  
Social Software New Media  
Real Time Localization  
Transaction Costs  
Collective Intelligence  
Sharable Resources  
Crowdsourcing



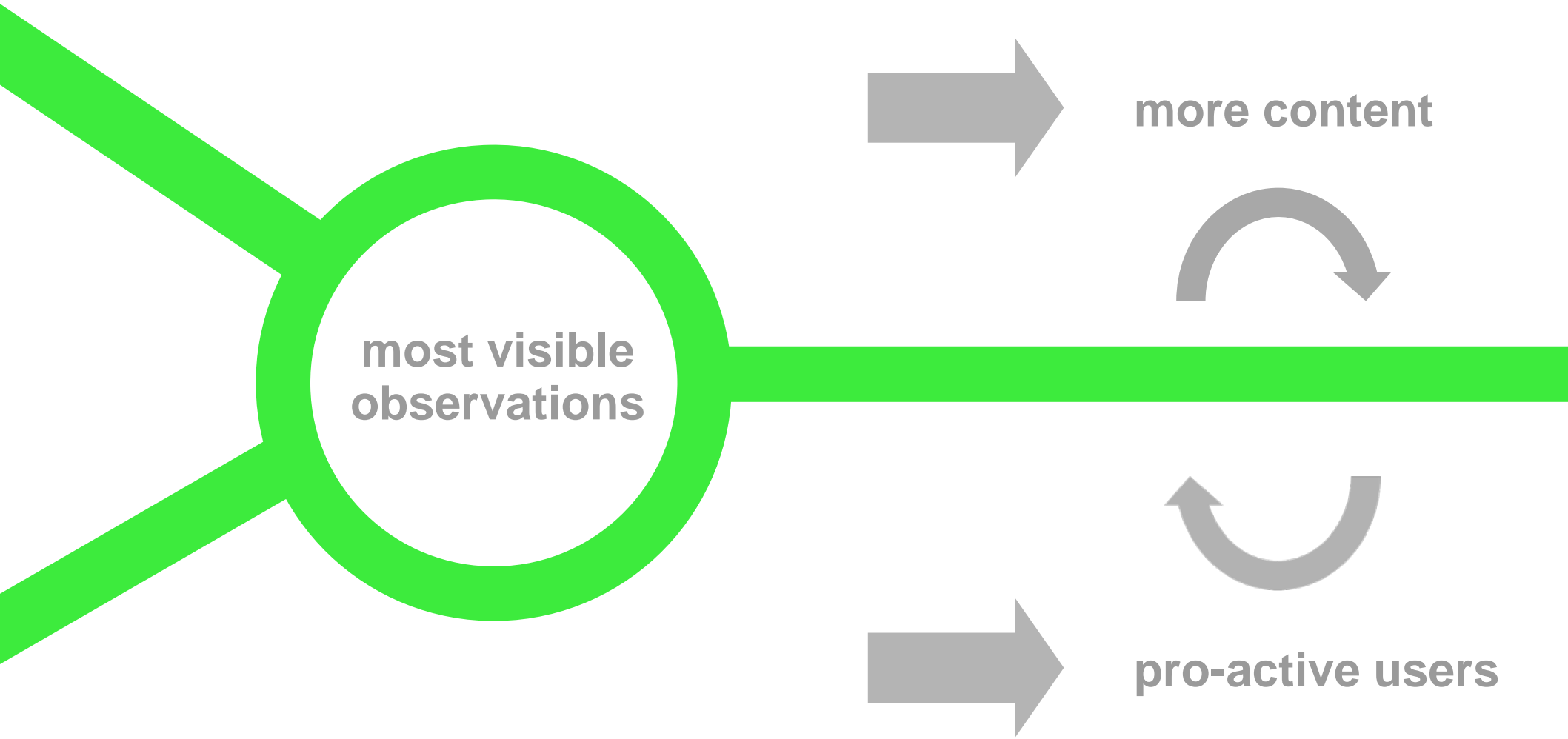
**specific views**

computer  
&  
information sciences

- *integration of social software*
- *bridge: self and alien organization*
- *shaping parameters*
- *analyses and evaluations*

- *motivation*
- *communities of practice*
- *person / community extension*
- *hedonism and incentives*

psychology  
&  
sociology





**general  
observations**

shrinking  
accelerating  
expanding

} **world**



“open source software” development

**added value = appreciation / esteem  
minimizing transaction costs**

**specific  
observations**

- micro enterprises
- small and medium enterprises
- large enterprises

**added value through  
optimizing transaction costs**



**challenges**

- platforms
- media
- data / information / knowledge
- languages and cultures
- copyright and ownership
- safety and security
- ...

necessary changes  
(transformations)

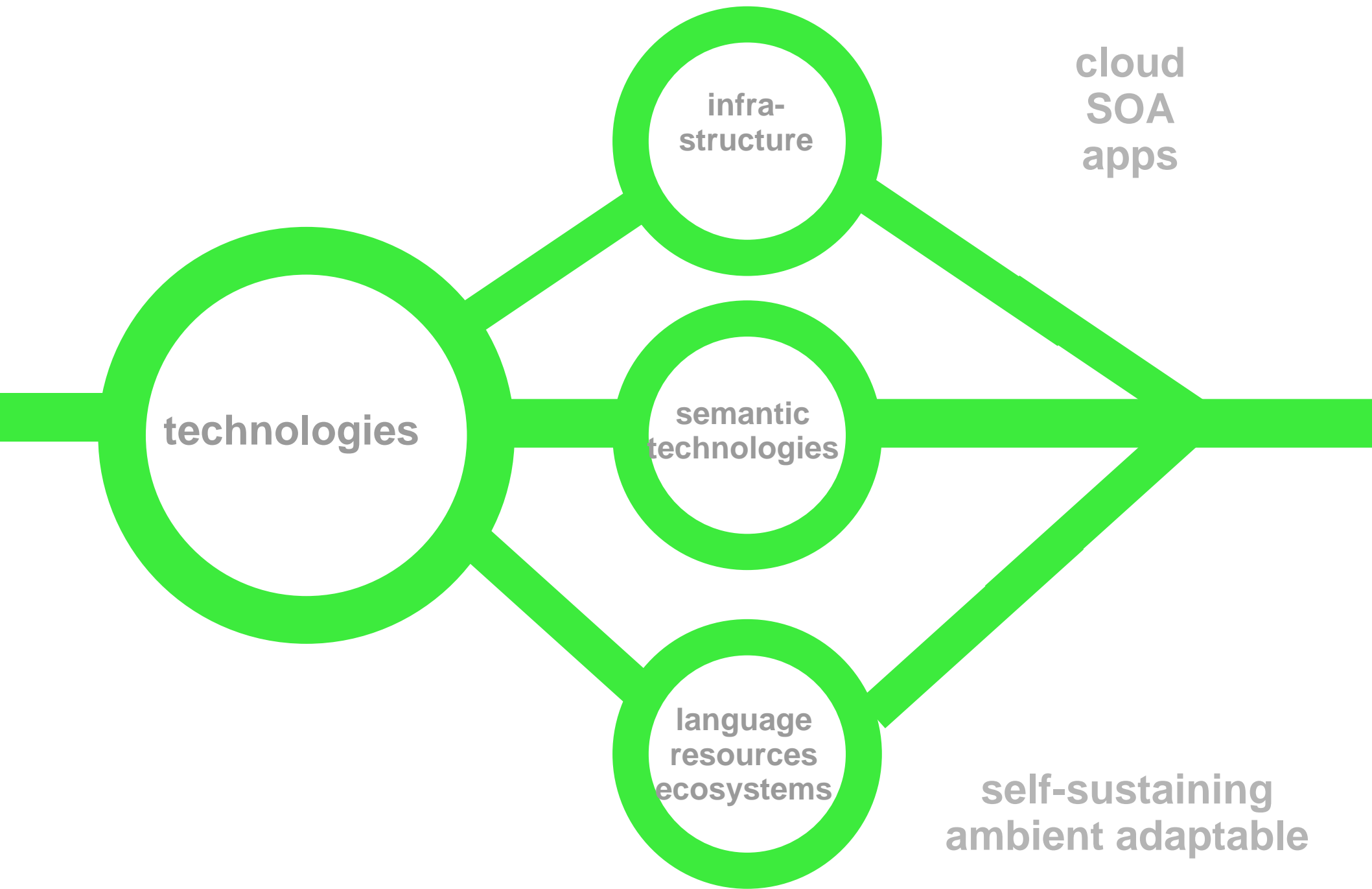
multiple  
levels

data / information / knowledge  
process / workflow  
enterprise / market  
society / politics

- from insufficient automation to *interoperability*
- from broken process and data streams to *continuity*
- from manual, disruptive tasks to *coherent flows*
- from nontransparent scoring and evaluation to *valuable assets*



processes



**technologies**

**infra-  
structure**

**semantic  
technologies**

**language  
resources  
ecosystems**

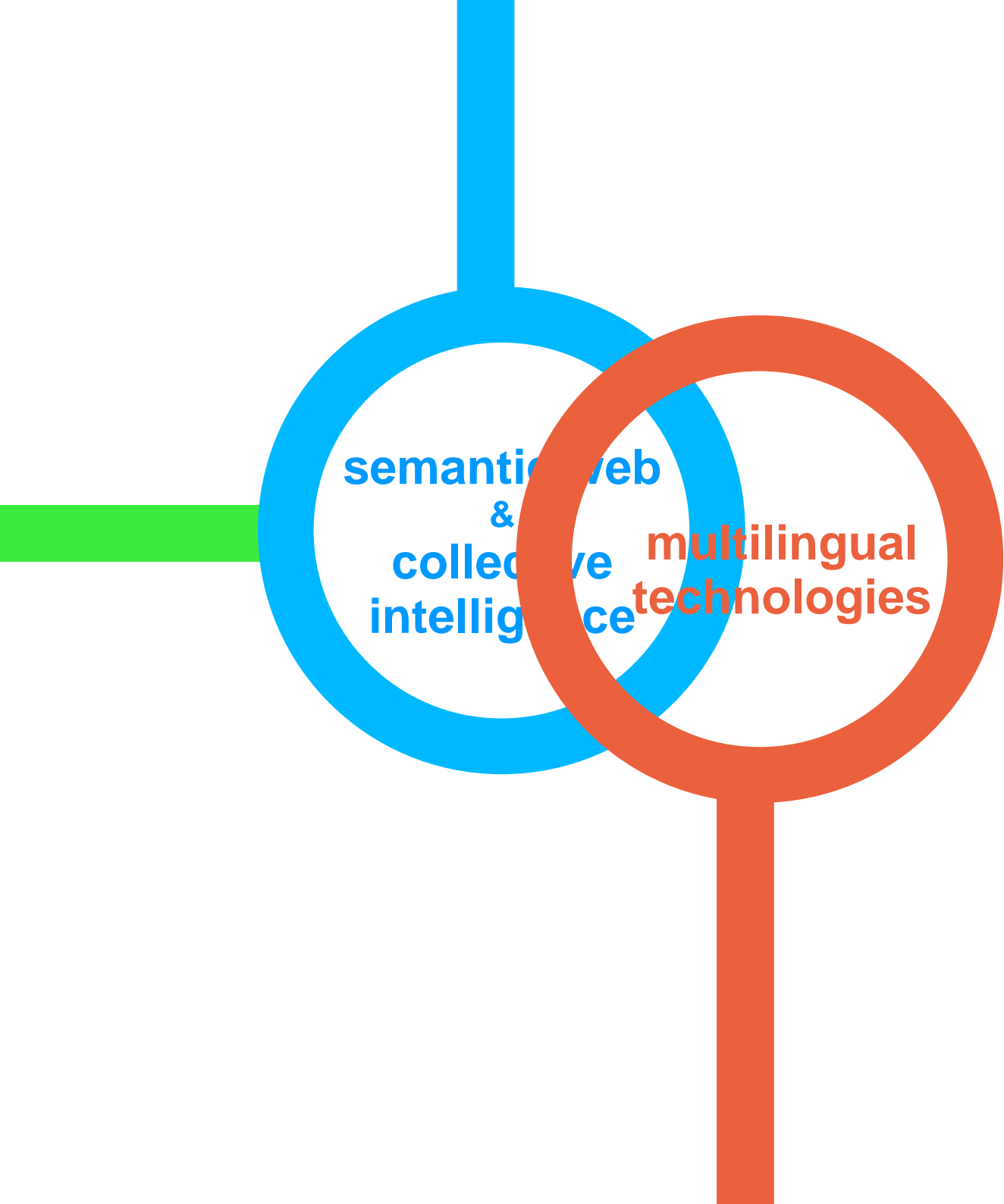
**cloud  
SOA  
apps**

**self-sustaining  
ambient adaptable**




**semantic web  
&  
collective  
intelligence**

- from markup to *self-describing data*
- from ontology to *pattern in data*
- from time-delayed to *real time*
- from objects to *information shadows*



semantic web  
&  
collective  
intelligence

multilingual  
technologies



**Part 2:**  
**MT**  
**as a »gadget«**

how MT fits in:  
*the MT biased use  
case*





**challenges**

- ubiquity of I&C media across domains and cultures
- anytime available information in multiple languages
- new forms of data quality and data integrity

- data stream translation
- data and service interoperability
- information and knowledge sharing

necessary changes  
(transformations)

views  
(levels)

services  
feedback cycle  
self-learning



**services**

- cloud-based
- accessible
- multiple formats
- standards



**feedback  
cycles**

- open
- responsive
- adaptable
- interoperable



**self-learning**

- stimuli sensitive
- context aware
- emergent
- stigmergent
- self-sustaining



MT in  
Enterprise 2.0

- barrier-free
- network-native
- participating
- real time



- from enterprises to *transcultural communities* with micromanagement, agility and responsibility
- from numbers to *values* with streams and continuity
- from platforms to *social networks*
- from links to *relations*

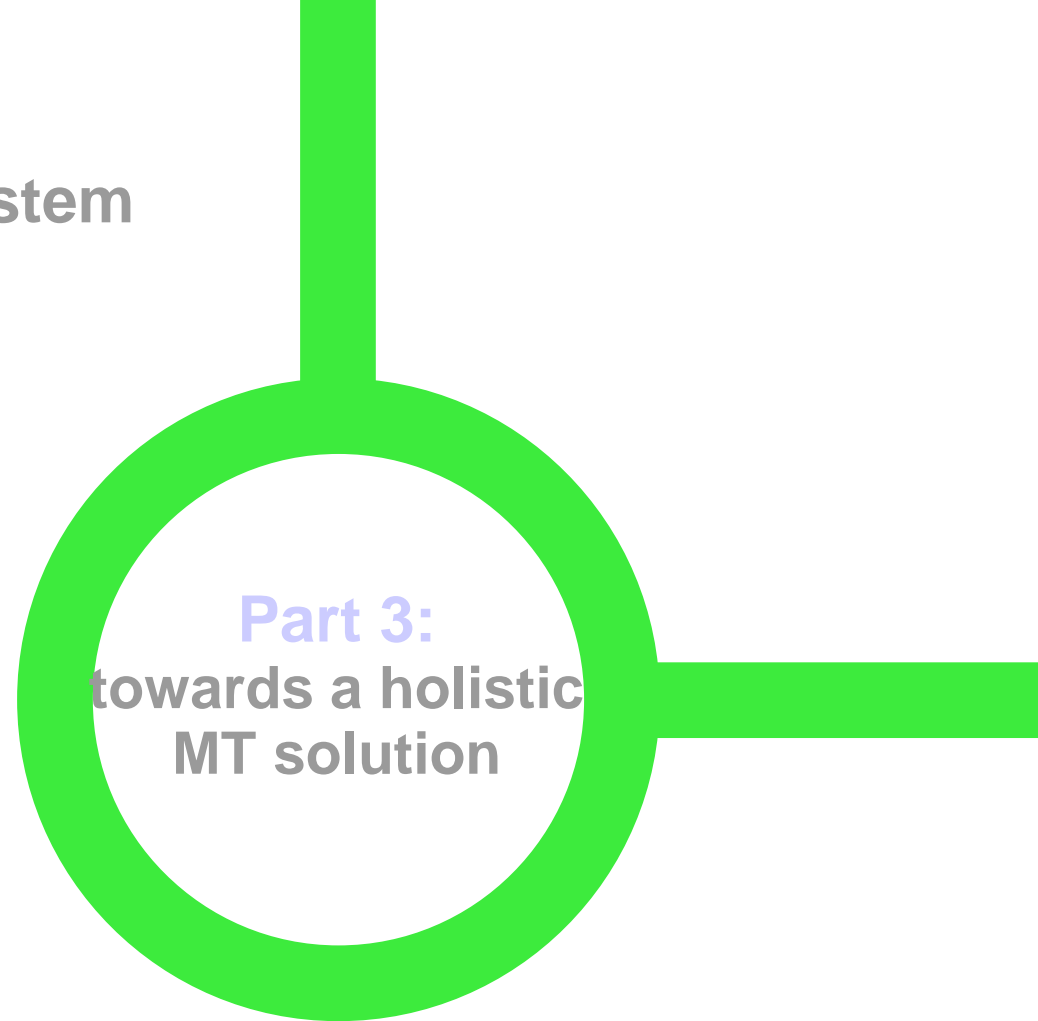
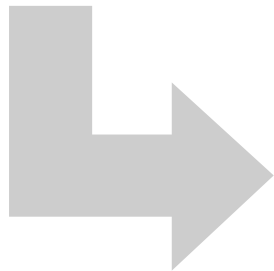
*“next practices”*  
vs.  
*“good or best practices”*



general impact:  
real time  
» globalization «

## inspired by biological immune system

- highly distributed
- feature extracting
- adapting
- self-organizing
- learning
- memorizing



- web / net services
- terminology biased
- pattern recognition
- context accumulation
- learning capability
- evaluate function





**terminological  
data**

- hidden in enterprise documents, databases and communication streams
- searchable but often not findable
- linked but not semantically related

## hearing impairment

amblyacousia  
sociacusis

## terms

deafness

hardness of hearing

tready bondage

Aufgrund seiner **Schwerhörigkeit** war es nicht möglich, sich mit dem Patienten zu verständigen.

Für Patienten, die neben ihrem Husten über ein kratziges Gefühl im Hals klagen, sind Hustensaft und Lutschpastillen geeignet.

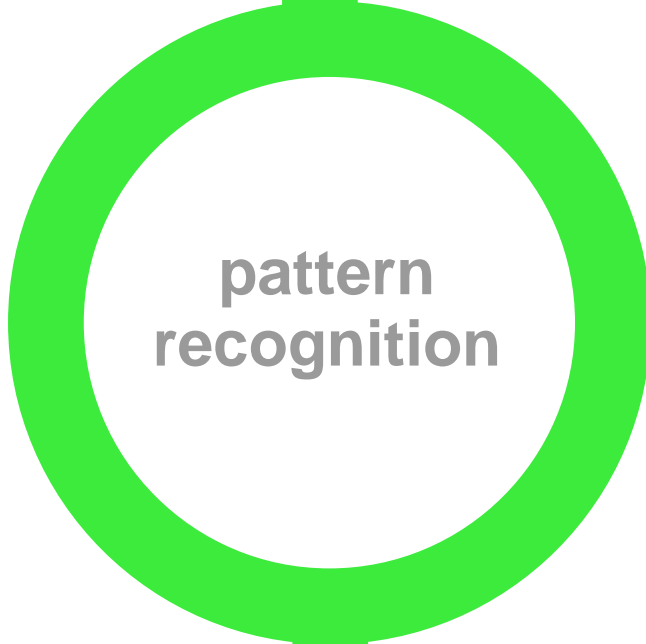
Ambroxol, Lidocain oder Benzocain wirken lokalanästhetisch, Cetylpyridiniumchlorid und Dequaliniumchlorid sowie Chlorhexidin und Hexetidin desinfizierend.

**term density** – terminology scoring  
and filtering

## numbers

$10^6$	one million	one million	eine Million
$10^9$	one <b>billion</b>	one thousand million	eine Milliarde
$10^{12}$	one <b>trillion</b>	one <b>billion</b>	eine <b>Billion</b>
$10^{15}$	one <b>quadrillion</b>	one thousand billion	eine Billiarde
$10^{18}$	one quintillion	one <b>trillion</b>	eine <b>Trillion</b>
$10^{21}$	one sextillion	one thousand trillion	eine Trilliarde
$10^{24}$	one septillion	one <b>quadrillion</b>	eine <b>Quadrillion</b>

**term selection** – terminology ranking



*learning phases:*

- grow
- adapt
- shrink

pattern translation ranking

a picture **of** Monnet  
a picture **from** Monnet  
a picture **by** Monnet  
a picture **owned by** Monnet



*ein Bild **von** Monnet*

translation context ranking

context  
accumulation

We don't like *involved discussions*  
with the *people involved*.

*learning phases:*

- grow
- adapt
- shrink

- differing contexts
- divergent translations
- cultural (mis-)matches



**evaluation**

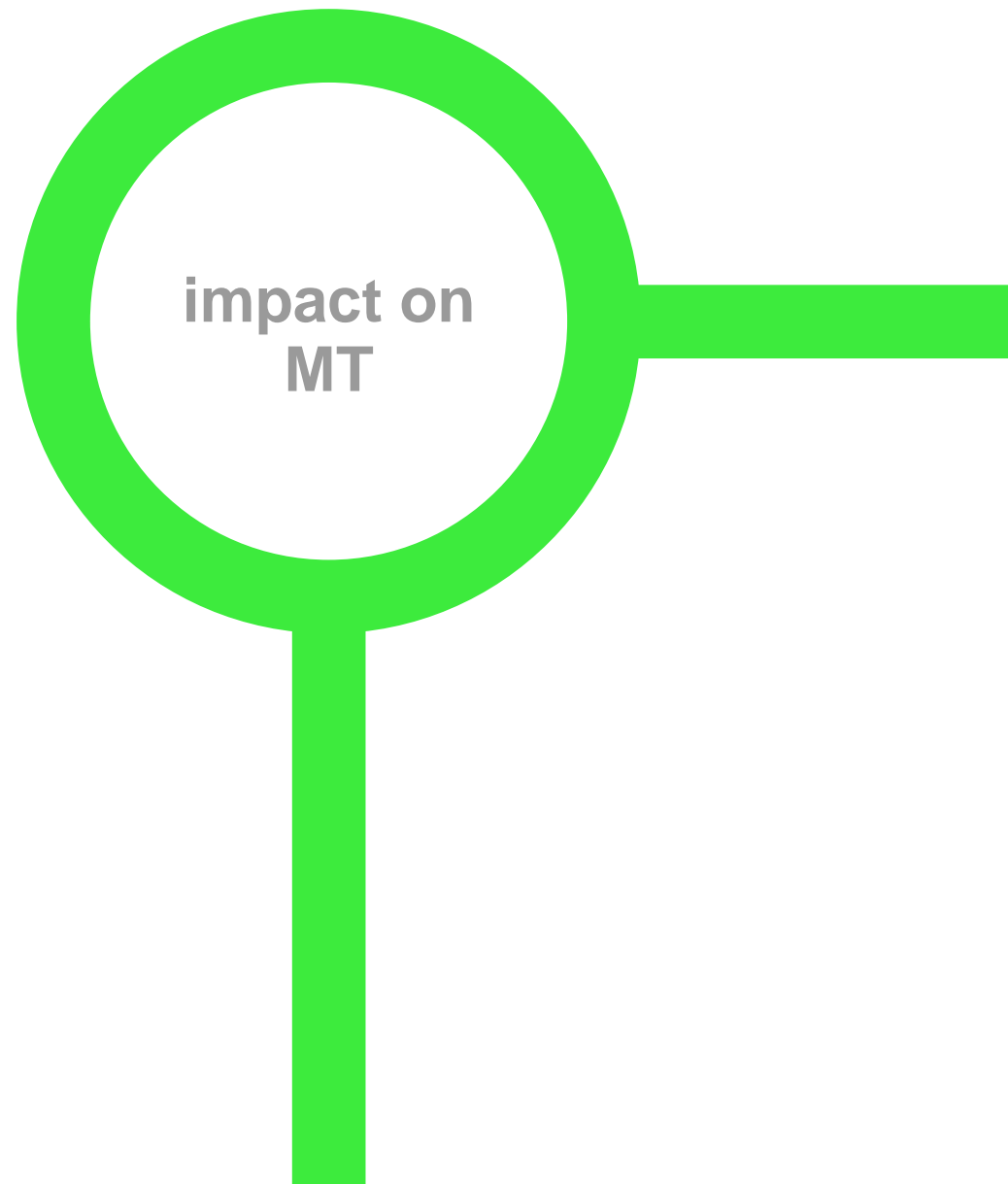
translation process evaluation with  
key performance indicators

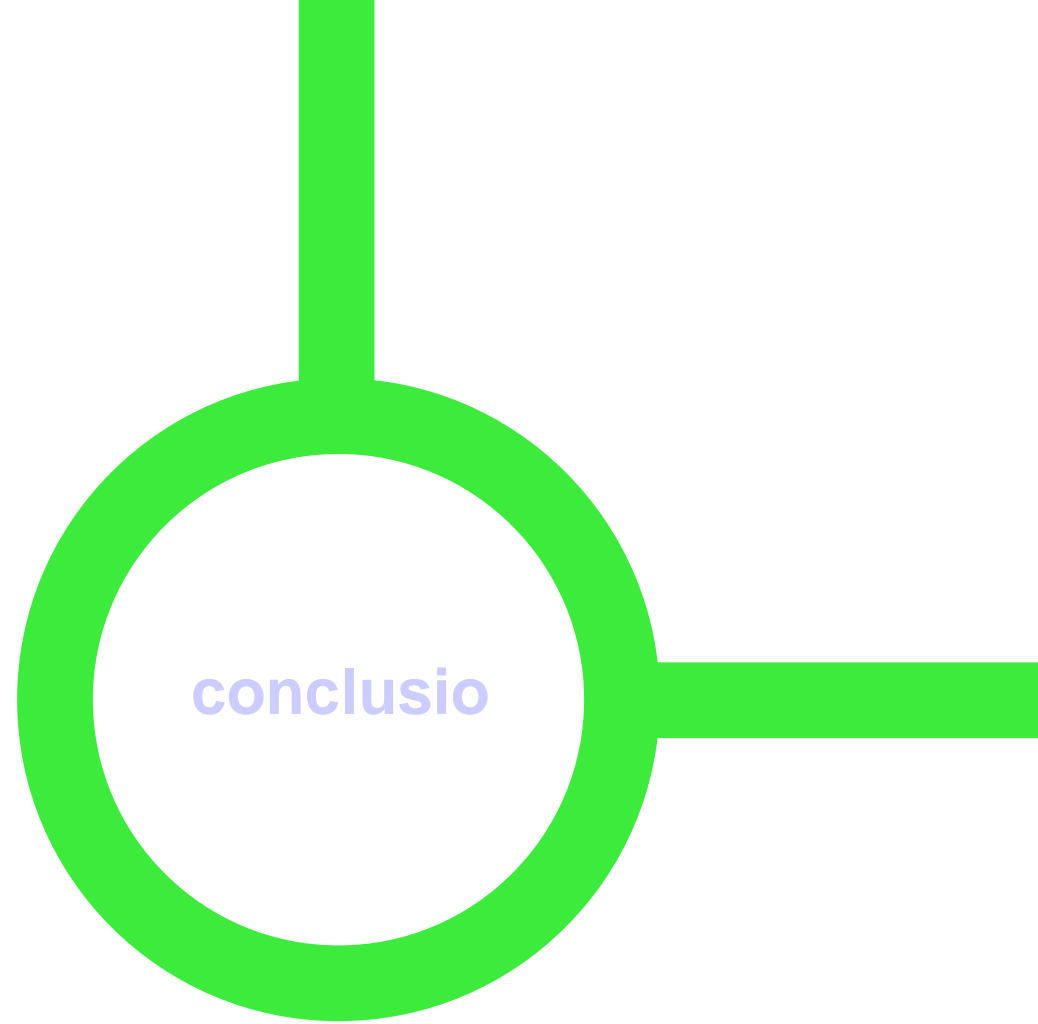
- BLEU
- variants
- ...

} standards

- term filtering, ranking and scoring
- pattern and context ranking
- discrimination between streams  
and documents

- multiple MT engines
- artificial immune system
- different scoring and ranking algorithms
- learning and memorizing
- feedback interfaces and hooks to permit learning and adapting to changing contexts within specific environments





conclusio





**MT visions**

*learning* ✓

*self-describing* ✓

*continuous streams* ✓

*feedback* ✓

*adaptable* ✓

*micromanagement* ✓

*semantic technologies* ✓

*collective intelligence* ✓

*“... the spontaneous perception of connections  
and meaningfulness in unrelated things.”*

*—William Gibson*



**thank you!**

**Prof. Dr. Jörg Schütz**  
**bioloom group**  
**joerg@bioloom.de**  
**+49 170 801 9982**

**Sven C. Andrä**  
**Andrä AG**  
**sca@andrae-ag.com**  
**+1 408 538 5362**