Cognitive-behavioral model of natural and artificial psyche based on social-evidence and resource restriction for decision support systems

Anton Kolonin
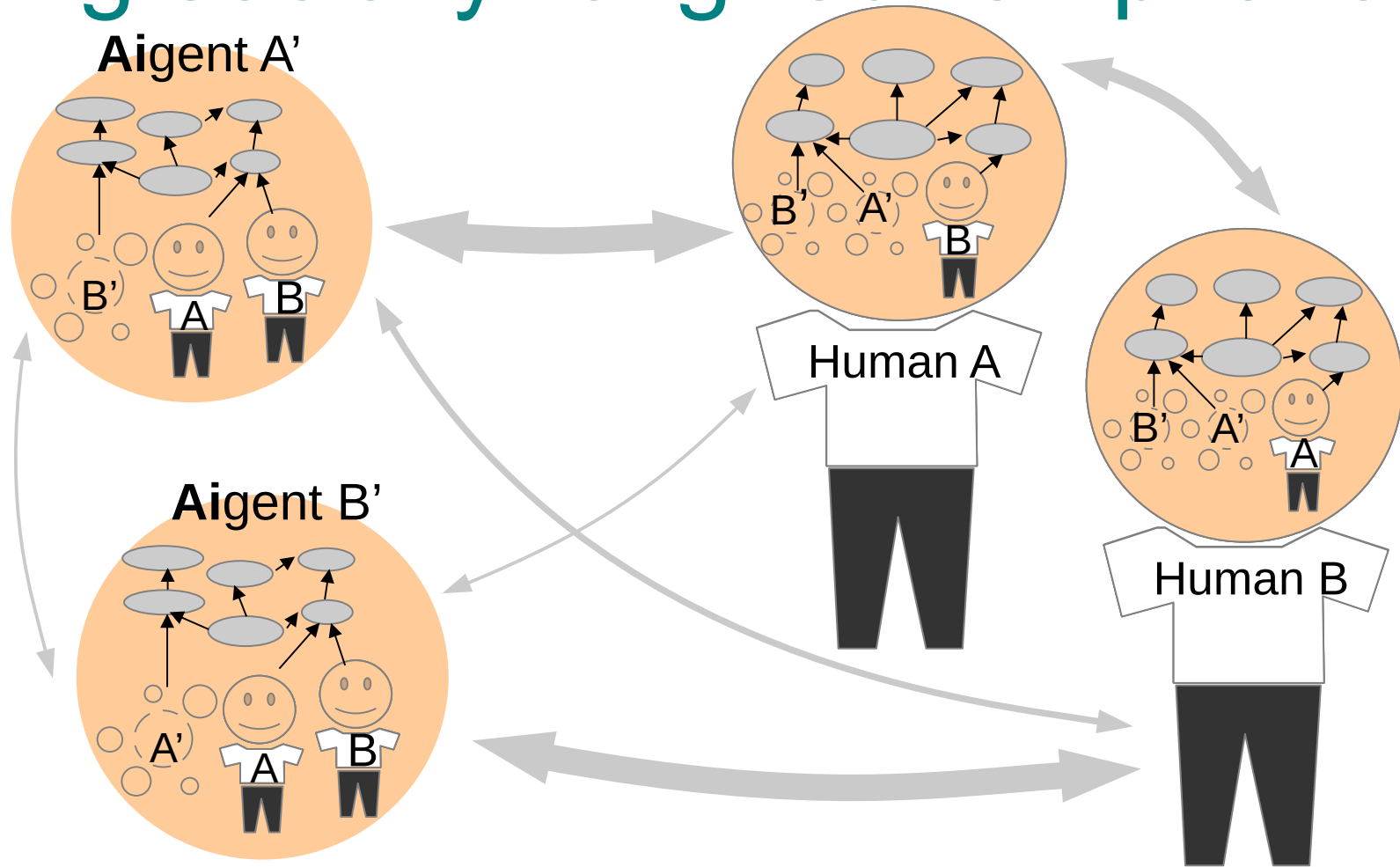
akolonin@aigents.com
Telegram: akolonin

https://www.nsu.ru

AIGENTS
https://aigents.com
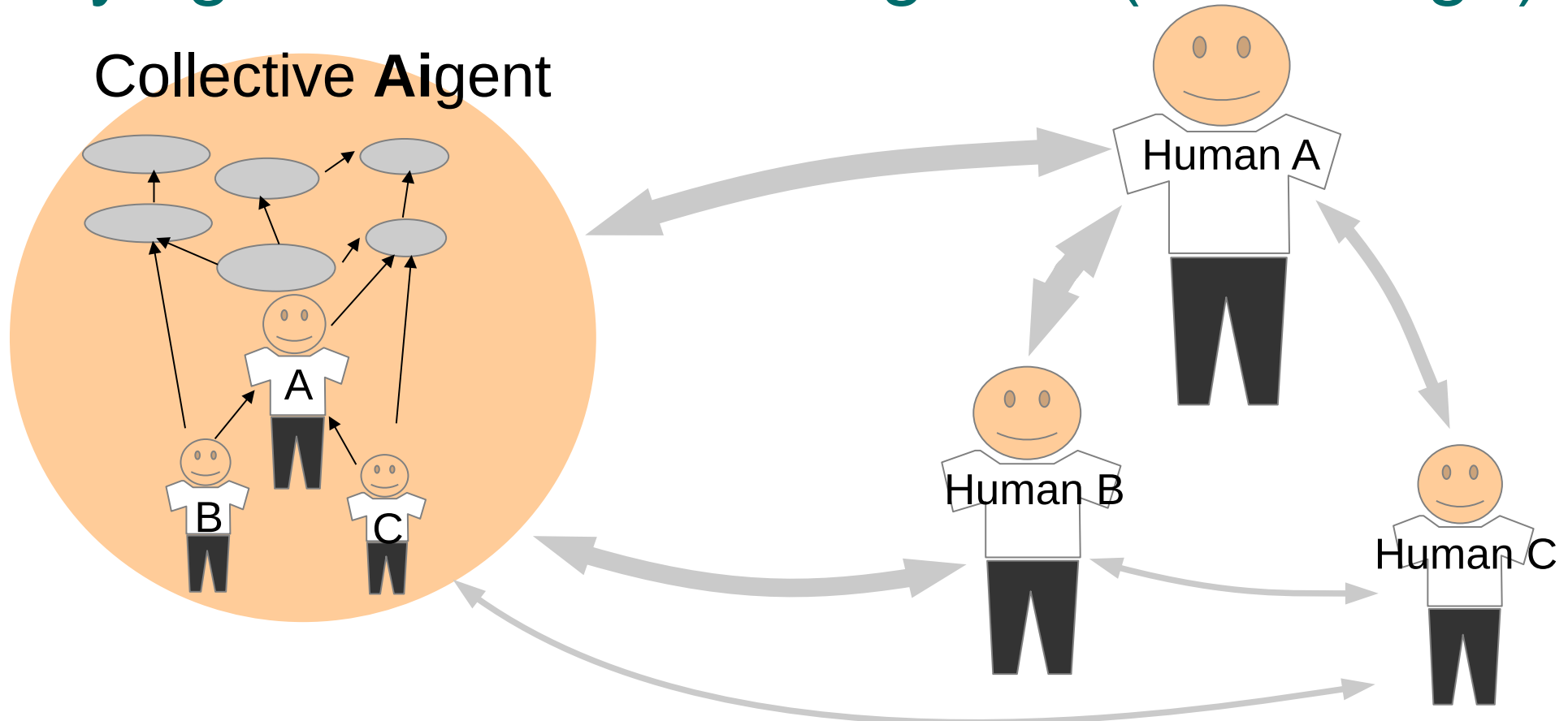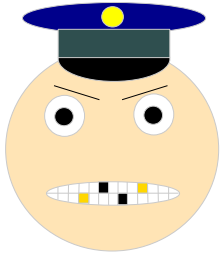
https://agirussia.org

Building socially-"aligned" emphatic AI-s

2

Implement decision support systems (DSS) relying on collective intelligence (knowledge)
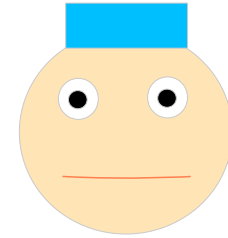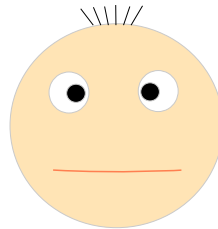

Collective **Ai**gent

Enforcing/preventing manipulation (applied "belief engineering")
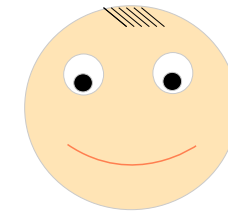

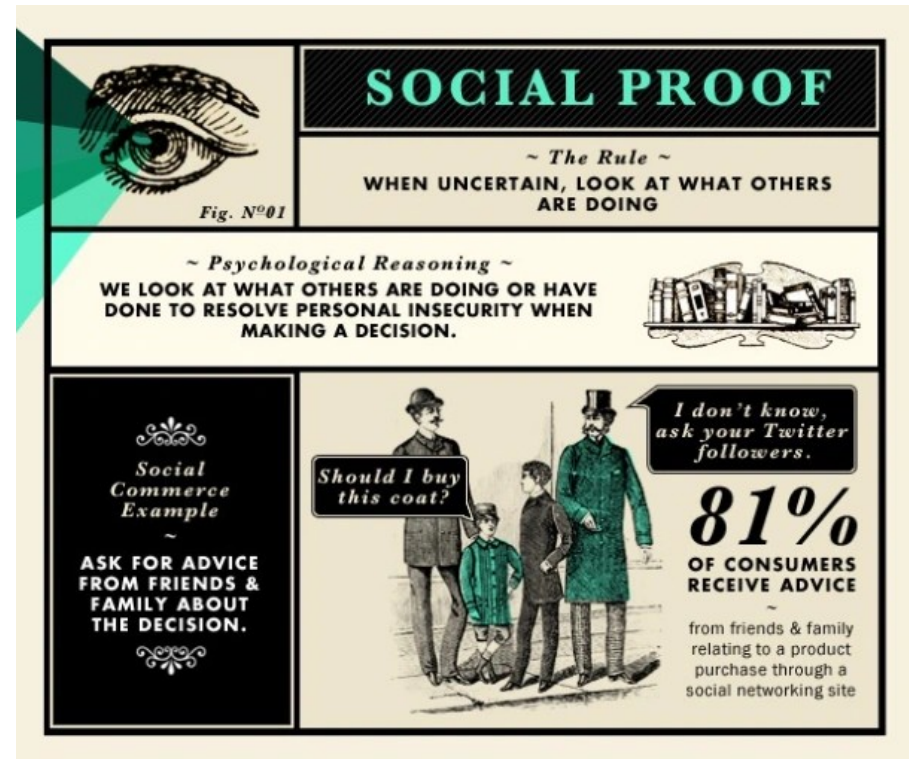PsyOps,
Social Engineering
Scam
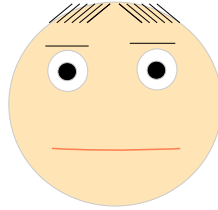
Sales,
Marketing,
Ads

Psychology,
Psychotherapy,
Psychoanalysis

4

Robert Cialdini: Social Proof

Social proof is a psychological and social phenomenon wherein people copy the actions of others in choosing how to behave in a given situation. The term was coined by Robert Cialdini in his 1984 book Influence: Science and Practice, and the concept is also known as informational social influence.

SOCIAL PROOF

Fig. Nº01

~ The Rule ~
WHEN UNCERTAIN, LOOK AT WHAT OTHERS ARE DOING

~ Psychological Reasoning ~
WE LOOK AT WHAT OTHERS ARE DOING OR HAVE DONE TO RESOLVE PERSONAL INSECURITY WHEN MAKING A DECISION.

Social Commerce Example
~
ASK FOR ADVICE FROM FRIENDS & FAMILY ABOUT THE DECISION.

Should I buy this coat?

I don't know, ask your Twitter followers.

81% OF CONSUMERS RECEIVE ADVICE

from friends & family relating to a product purchase through a social networking site

https://www.amazon.com/Influence-Practice-Robert-B-Cialdini/dp/0205609996   *1984*

Computable cognitive-behavioral model based on social evidence and restricted by resources

B.Goertzel, A.Kolonin, J.Pressing, C.Pennachin: Compassion-based artificial psyche design (1999-2000)
https://www.goertzel.org/benzine/WakingUpFrom TheEconomyOfDreams.htm

A.Kolonin: Computable cognitive model based on social evidence and restricted by resources (2015)
https://ieeexplore.ieee.org/document/7361869

Anton Kolonin, Evgenii Vityaev, Yuriy Orlov: Cognitive Architecture of Collective Intelligence Based on Social Evidence (2016)
https://www.sciencedirect.com/science/article/pii/ S1877050916317239

A.Kolonin: Resource-Constrained Social Evidence Based Cognitive Model for Empathy-Driven Artificial Intelligence (2018)
https://link.springer.com/chapter/10.1007/978-3-3 19-97676-1_10

Social referee

Social referee

Self-esteem
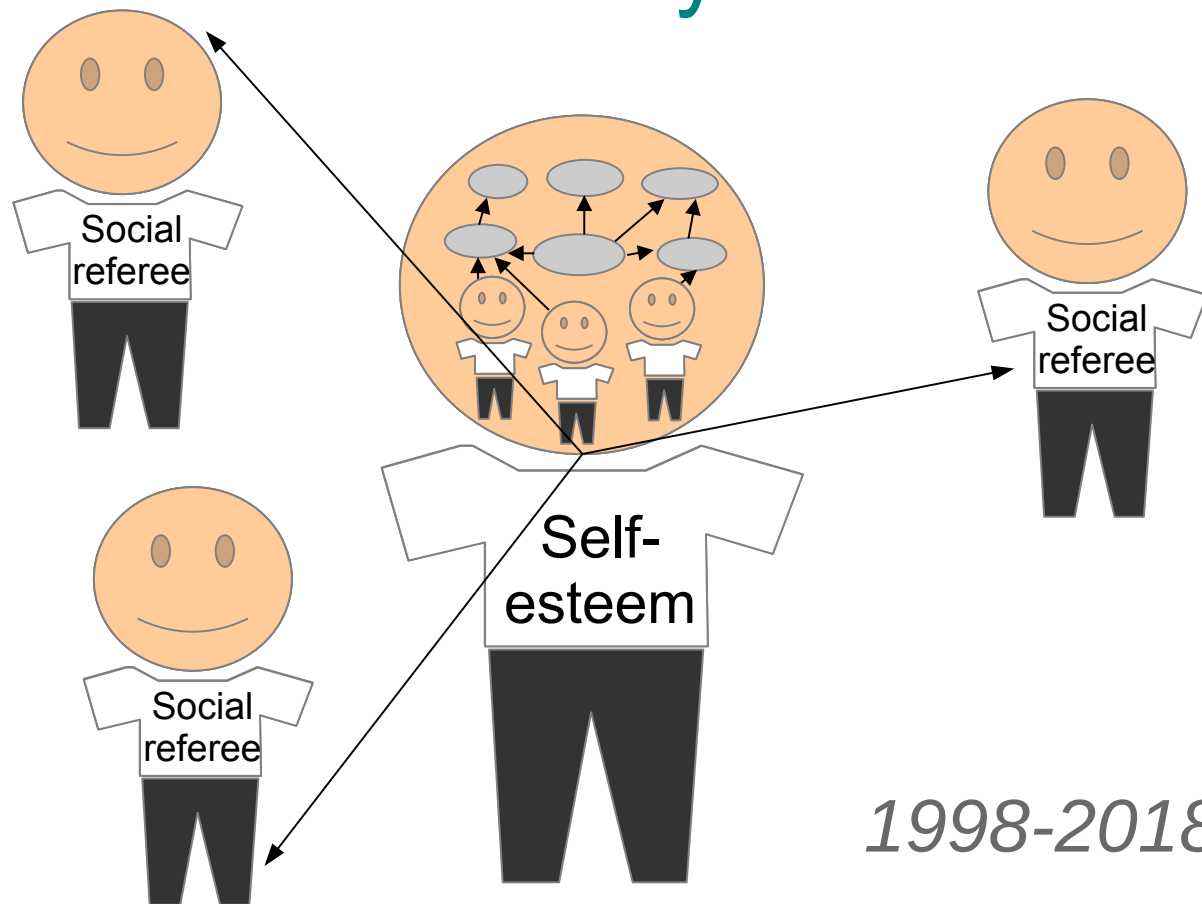
Social referee

1998-2018

Vladimir Lefebvre: Self-consciousness as social reflexivity

Москва
Когито-Центр
2017

Владимир Лефевр

ЧТО ТАКОЕ ОДУШЕВЛЕННОСТЬ?

Издание второе, исправленное и дополненное

линия. В лидере оказываются совмещенными оба начала – выполнение "рядовых" трудовых операций и специфической организующей деятельности.
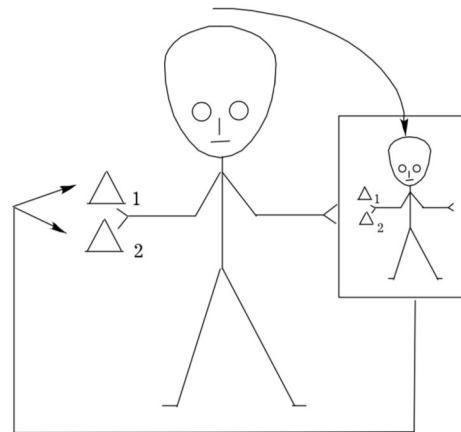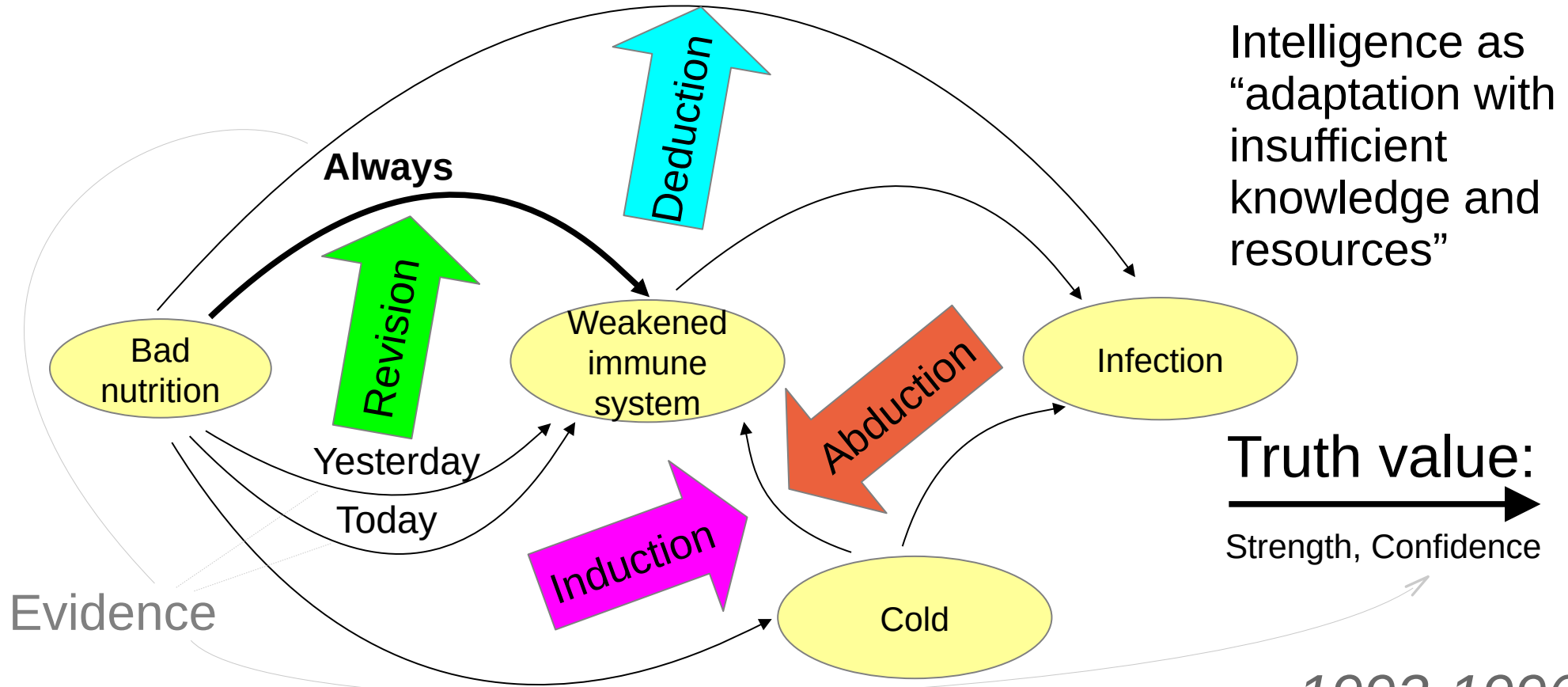

Рис. 9

Впервые индивидуальная деятельность оказывается *органи-зованной*. Механизм, который раньше действовал в масштабе коллектива, переходит в индивидуалльную деятельность. Лидер превращается в *саморефлексивную систему*.
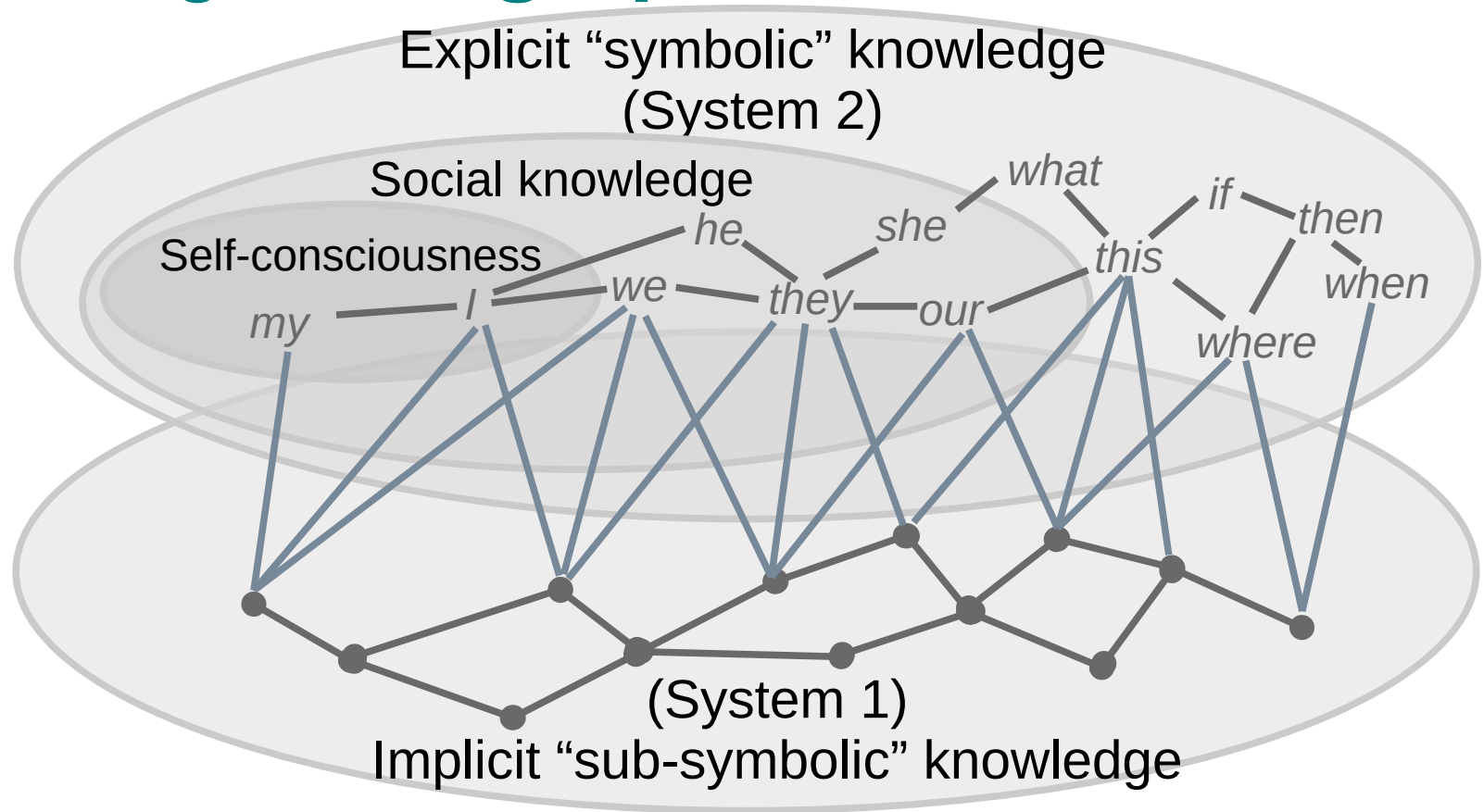
2018

https://www.litres.ru/book/vladimir-lefevr/chto-takoe-odushevlennost-66216968/

7

Pei Wang: Non-Axiomatic Reasoning System (NARS)

Intelligence as "adaptation with insufficient knowledge and resources"

Deduction

Always

Revision

Bad nutrition

Weakened immune system

Abduction

Infection

Yesterday

Today

Induction

Cold

Evidence

Truth value:

Strength, Confidence

1993-1996

Pei Wang: Non-axiomatic reasoning system: exploring the essence of intelligence, 1993-1996
https://www.researchgate.net/publication/2690739_Non-Axiomatic_Reasoning_System_-_Exploring_the_Essence_of_Intelligence

Knowledge as a **graph** with maintenance **cost**

Explicit "symbolic" knowledge
(System 2)

Social knowledge

Self-consciousness

my   I   we   they   our   he   she   what   this   if   then   when   where

(System 1)
Implicit "sub-symbolic" knowledge

https://www.amazon.com/Thinking-Fast-Slow-Daniel-Kahneman/dp/0374533555
https://amit02093.medium.com/atomspace-hyper-graph-information-retrieval-system-450cab9d751e

9

General intelligence: The ability to achieve complex goals in complex environments, using limited resources
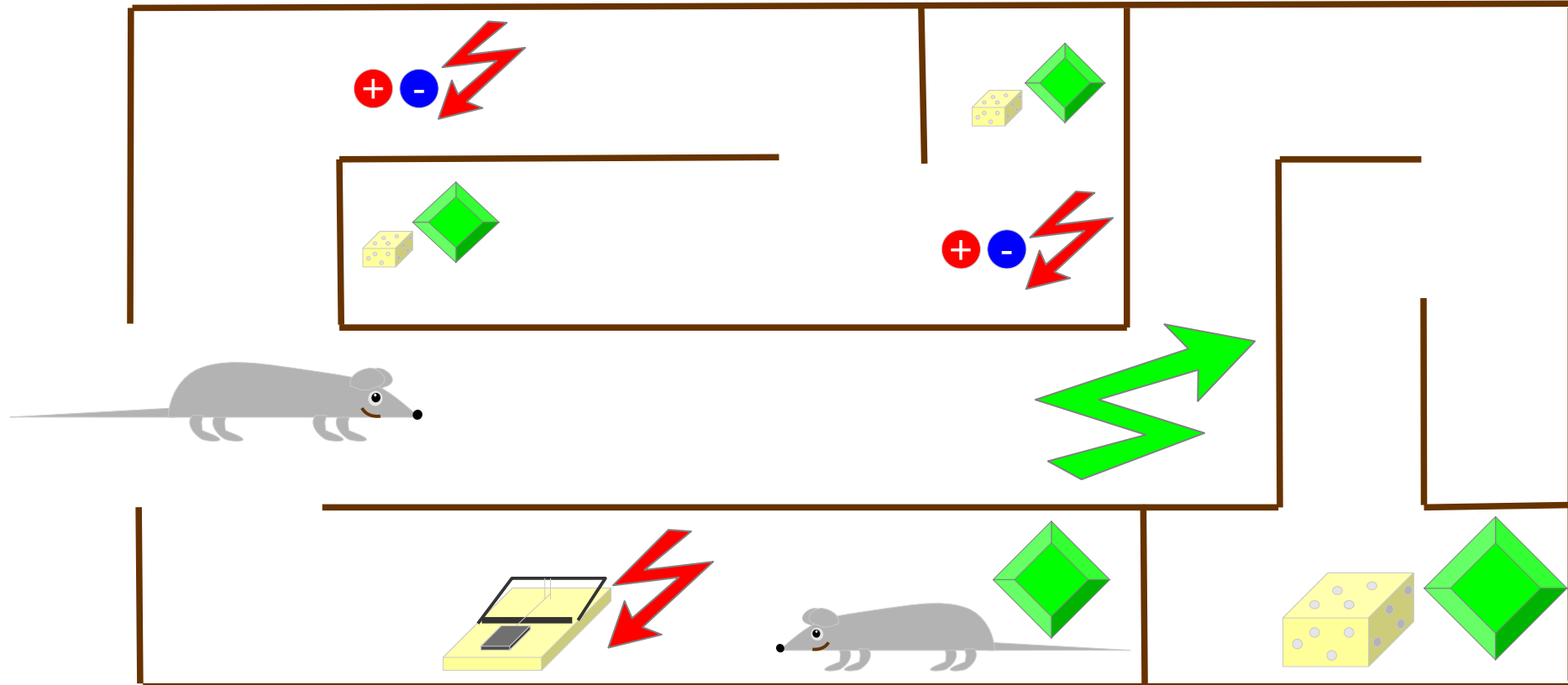

Intelligent being

2012
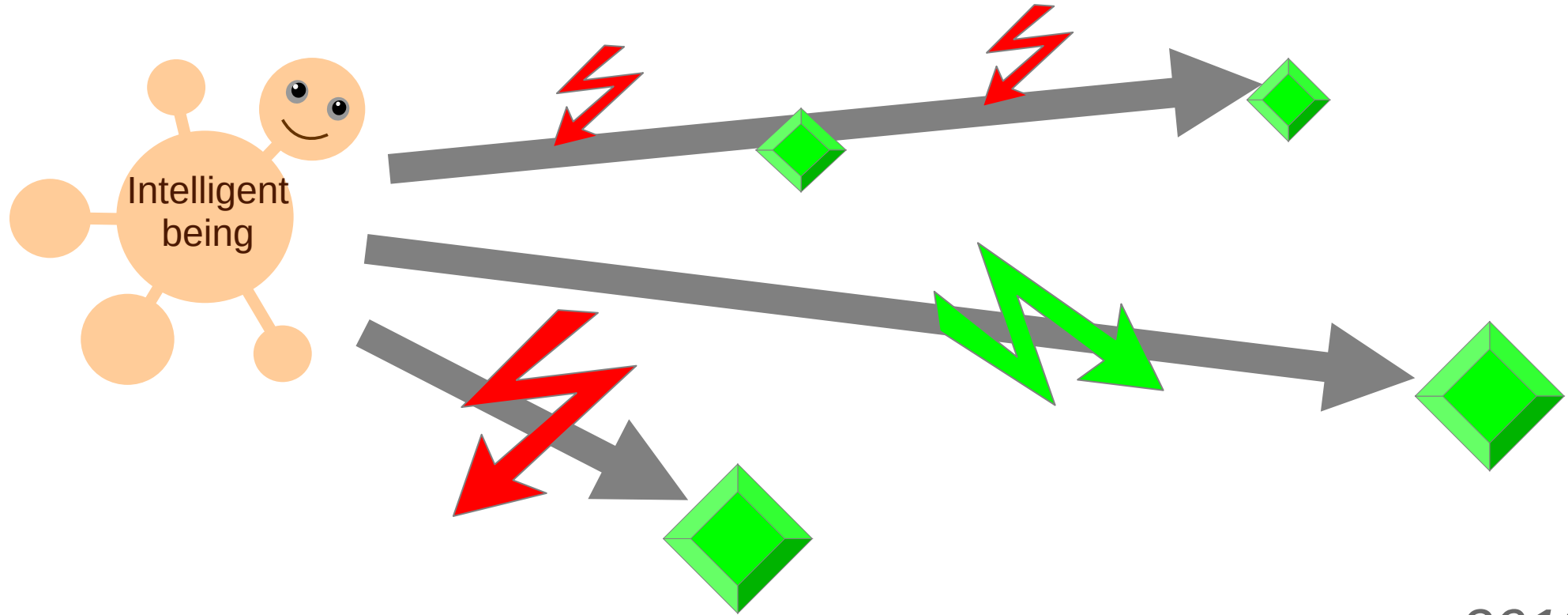
Ben Goertzel: CogPrime: An integrative architecture for embodied artificial general intelligence
https://www.researchgate.net/publication/289258669_CogPrime_An_integrative_architecture_for_embodied_artificial_general_intelligence

Biological Intelligence: Search for nutrition and reproduction opportunities given multiple physical threats and limited **temporal**, **physical** and **biological** resources

11

Artificial Intelligence (Aigents®): Search for the most valuable information found as quickly as possible given limited **temporal** and **computational** resources
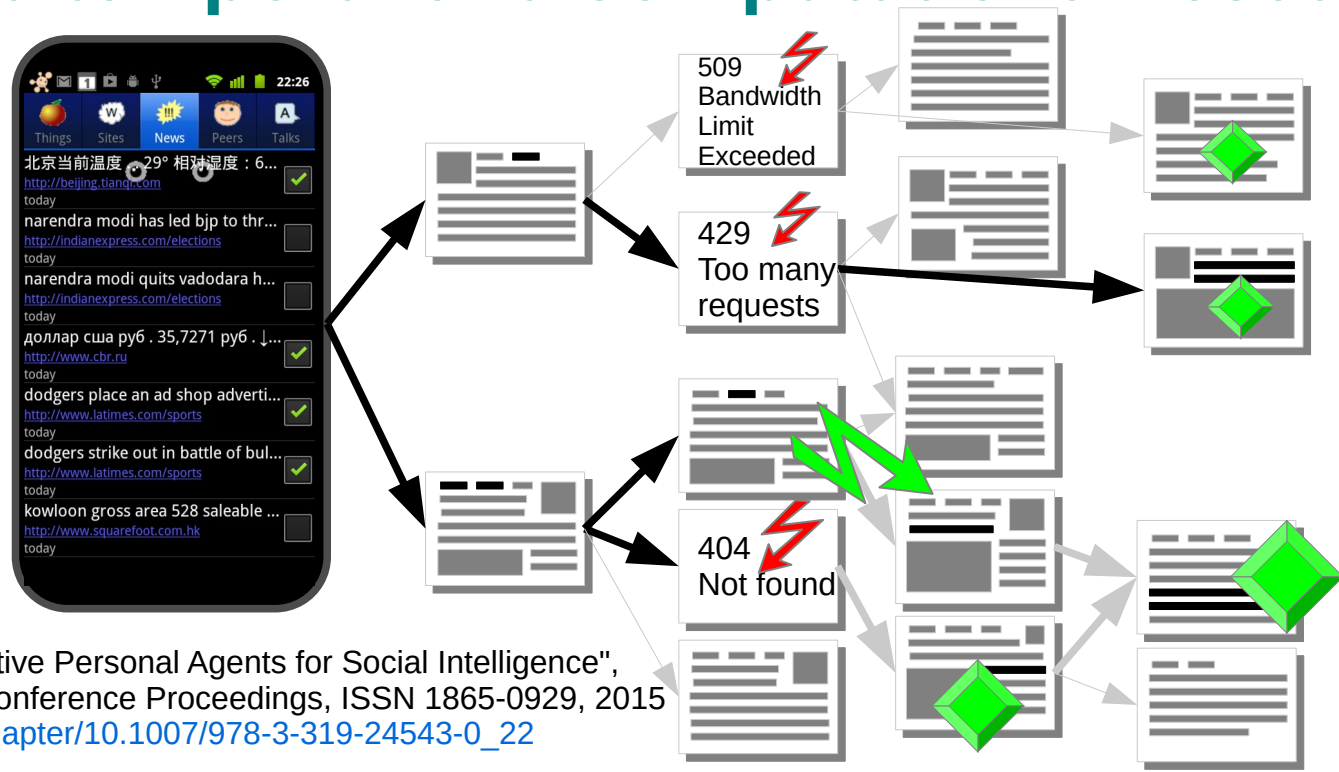
A. Kolonin, "Aigents: Adaptive Personal Agents for Social Intelligence", KESW, 6th International Conference Proceedings, ISSN 1865-0929, 2015
https://link.springer.com/chapter/10.1007/978-3-319-24543-0_22

A.Kolonin, "Adaptive experiential learning for business intelligence agents", CSGB Symposium Proceedings, 2016
https://ieeexplore.ieee.org/document/7587679
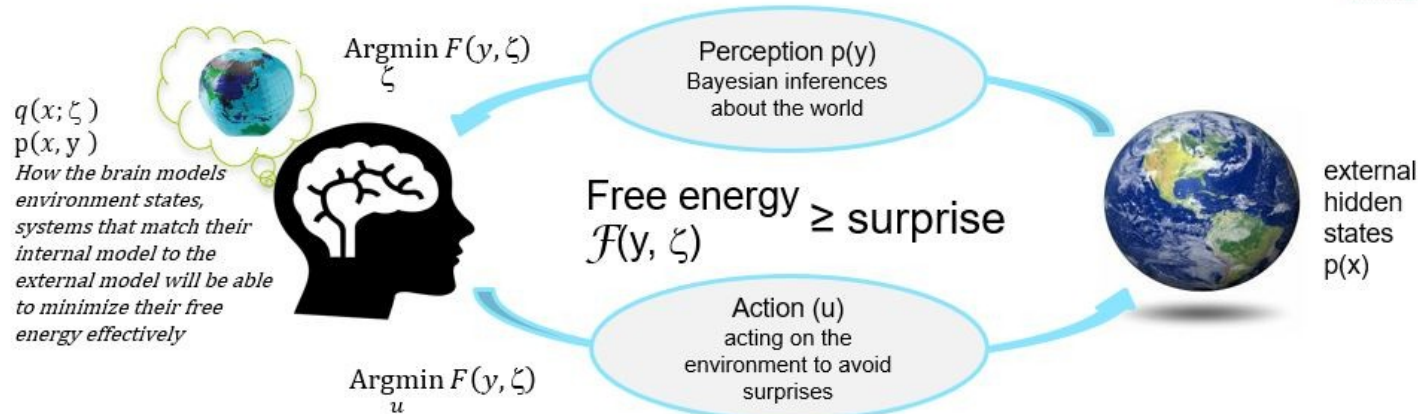
2015-2016

12

Karl Friston: Free Energy Principle

Minimize free energy

¨in summary, (i) agents resist a natural tendency to disorder by minimising a free-energy bound on surprise; (ii) this entails acting on the environment to avoid surprises, which (iii) rests on making Bayesian inferences about the world.¨

Prof. Karl Friston

$q(x;\zeta)$
$p(x,y)$
How the brain models environment states, systems that match their internal model to the external model will be able to minimize their free energy effectively

$\underset{\zeta}{\text{Argmin}}\, F(y,\zeta)$

Perception p(y)
Bayesian inferences about the world

$\text{Free energy } \mathcal{F}(y,\zeta) \geq \text{surprise}$

external hidden states p(x)

Action (u)
acting on the environment to avoid surprises

$\underset{u}{\text{Argmin}}\, F(y,\zeta)$
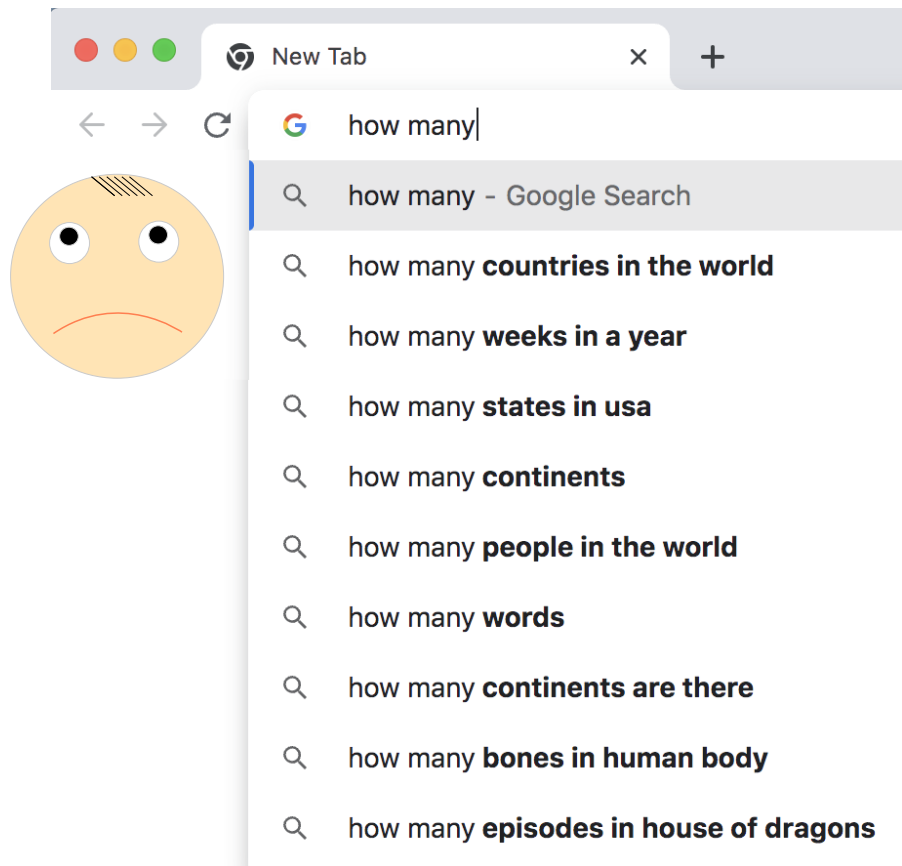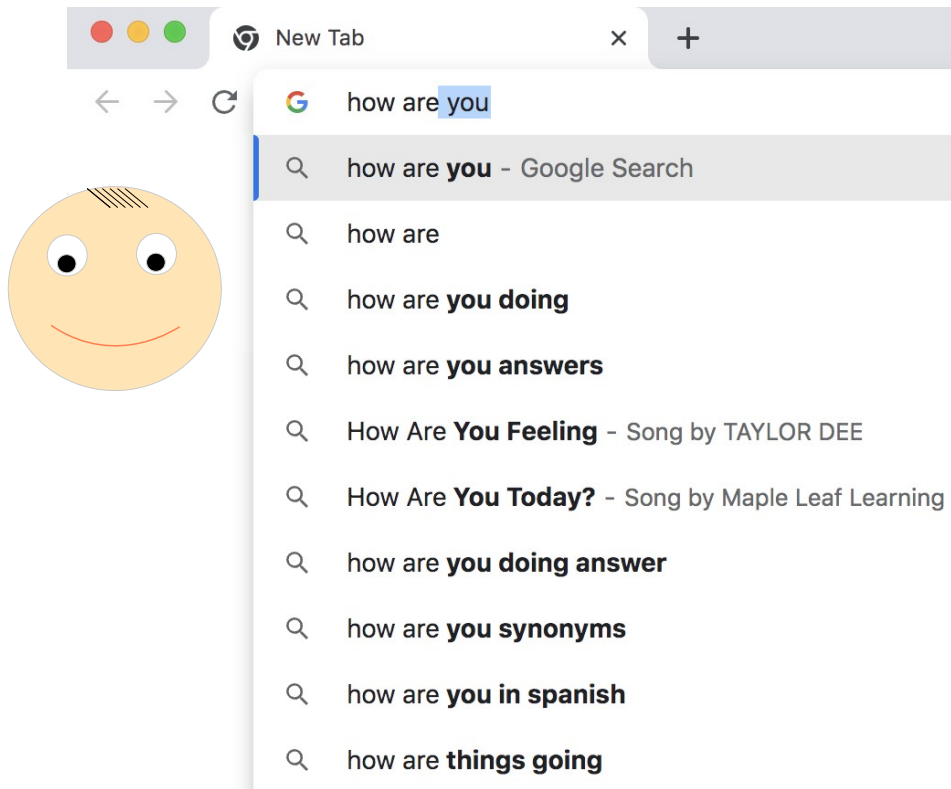
Thomas Parr, Giovanni Pezzulo, Karl J. Friston:
Active Inference: The Free Energy Principle in Mind, Brain, and Behavior, 2022
https://direct.mit.edu/books/oa-monograph/5299/Active-InferenceThe-Free-Energy-Principle-in-Mind
Slide content source: https://www.kaggle.com/code/charel/learn-by-example-active-inference-in-the-brain-2

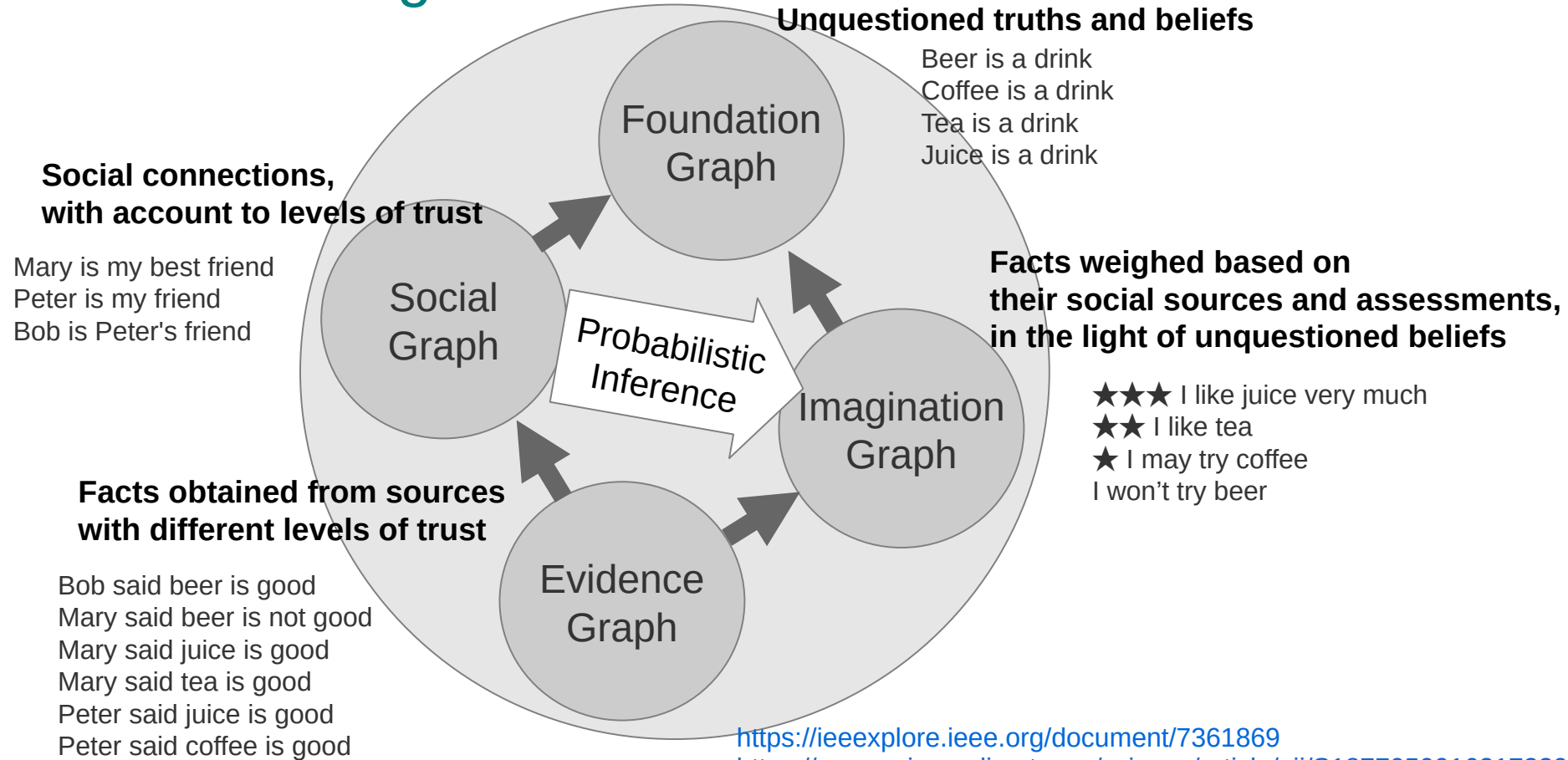*2006-2022*

Maximizing probability vs minimizing uncertainty

https://aclanthology.org/2022.emnlp-main.239/
https://arxiv.org/pdf/2303.02427.pdf

2022

14

Social evidence-based resource-constrained cognitive-behavioral model

Unquestioned truths and beliefs

Beer is a drink
Coffee is a drink
Tea is a drink
Juice is a drink

**Social connections,
with account to levels of trust**

Mary is my best friend
Peter is my friend
Bob is Peter's friend

Foundation Graph

Social Graph

Probabilistic Inference

Imagination Graph

**Facts weighed based on
their social sources and assessments,
in the light of unquestioned beliefs**

★★★ I like juice very much
★★ I like tea
★ I may try coffee
I won't try beer

**Facts obtained from sources
with different levels of trust**

Bob said beer is good
Mary said beer is not good
Mary said juice is good
Mary said tea is good
Peter said juice is good
Peter said coffee is good

Evidence Graph

2015-2018

https://ieeexplore.ieee.org/document/7361869
https://www.sciencedirect.com/science/article/pii/S1877050916317239
https://link.springer.com/chapter/10.1007/978-3-319-97676-1_10

Perception & Action driven by Believed Social Evidence

i – person of consideration

j – concept (action) of consideration (exposed evidence)

l – person's belief item (of foundation graph of size *L* - personal preference base)

k – person's correspondent (of social graph of size *K* - social reference base)

B_{il} – person's *i* mental attachment to *l* (personal preference)

S_{ik} – person's *i* social bind to *k* (social reference)

recursion

$$R^B \qquad\qquad R^S$$

$$P_{ij} = \Sigma_{l=1,L} (E^B_{ijl} * B_{il}) * \Sigma_{k=1,K} (E^S_{ijk} * S_{ik}) \quad \textit{(what we think)}$$

E^B_{ik} – concept (action) *j* agreement or compatibility with *l* in mind of *i* (believed evidence)

E^S_{ik} – concept (action) *j* expression or confirmation by *k* in view of *i* (social evidence)

P_{ij} – concept (action) *j* power for *i* (personal evidence)

$$C_{ij} = L + K + R^B + R^S \qquad \textit{(why we don't think about that)}$$

C_{ij} – concept (action) cost *j* cost for *i* (personal cost)

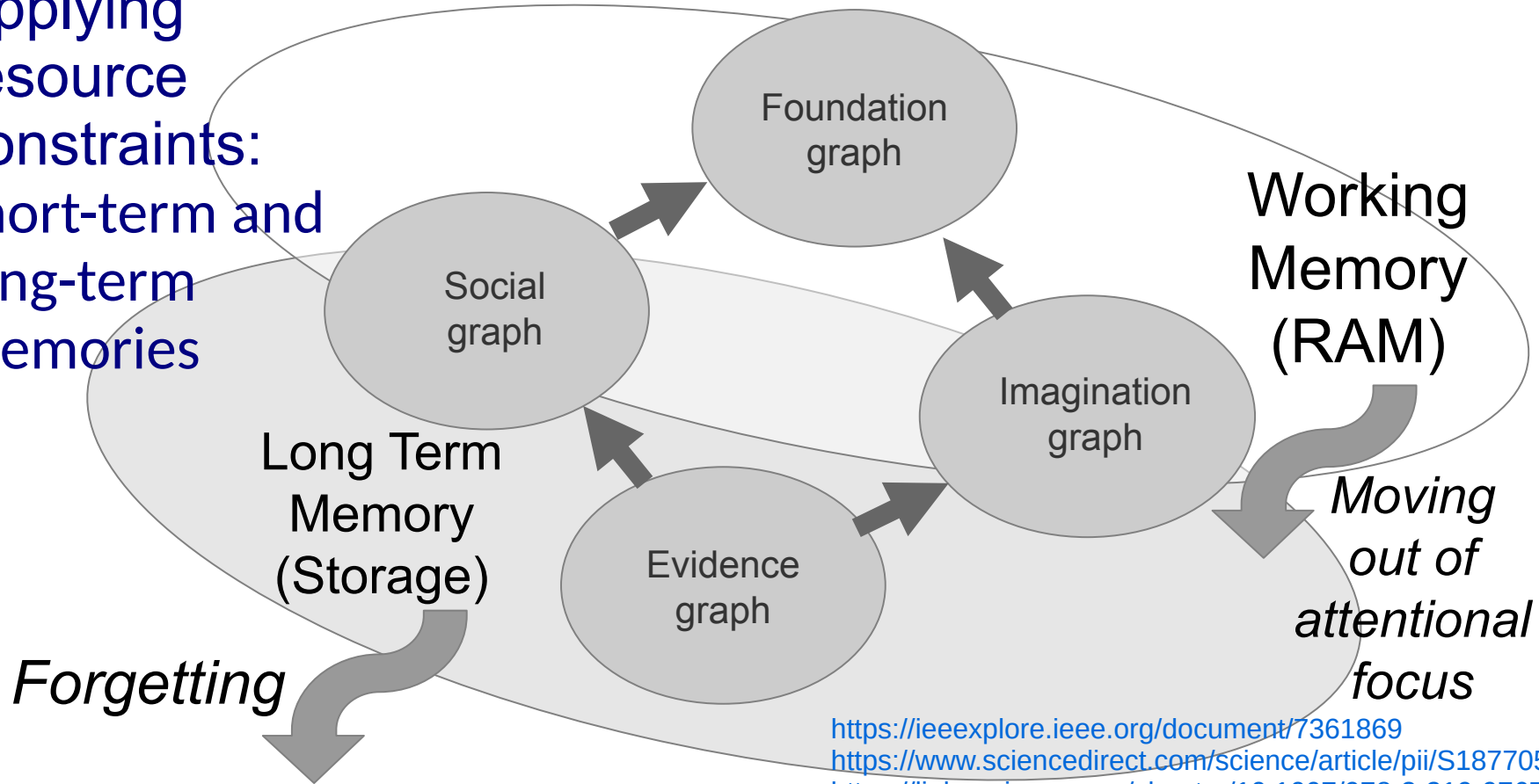# Social evidence-based resource-constrained cognitive-behavioral model

Unquestioned truths and beliefs
B_{il} – person's i mental attachment to l
l – person's belief item
(of foundation graph of size L)

**Social connections,
with account to levels of trust**

S_{ik} – person's i social bind to k
k – person's correspondent
(of social graph of size K)

$$P^B_{ij} = \Sigma_{l=1,L}(E^B_{ijl} * B_{il})$$

$$C_{ij}$$

**Facts weighed based on
their social sources and assessments,
in the light of unquestioned beliefs**

$$P^S_{ij} = \Sigma_{k=1,K}(E^S_{ijk} * S_{ik})$$

$$P_{ij} = \Sigma_{l=1,L}(E^B_{ijl} * B_{il}) * \Sigma_{k=1,K}(E^S_{ijk} * S_{ik})$$

**Facts obtained from sources
with different levels of trust**
E_{ij} - exposed evidence

https://ieeexplore.ieee.org/document/7361869
https://www.sciencedirect.com/science/article/pii/S1877050916317239
https://link.springer.com/chapter/10.1007/978-3-319-97676-1_10

Social evidence-based resource-constrained cognitive-behavioral model

Applying resource constraints: short-term and long-term memories

Working Memory (RAM)

Foundation graph

Social graph

Long Term Memory (Storage)

Imagination graph

Evidence graph

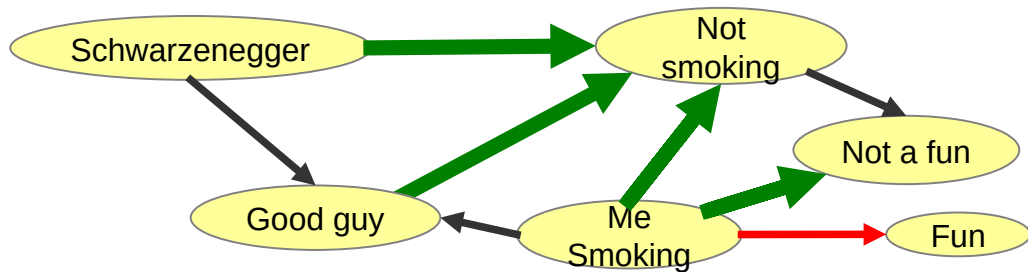*Moving out of attentional focus*

Forgetting

2015-2018

Basic technique of applied "belief engineering"

"Redundant temporal evidence"
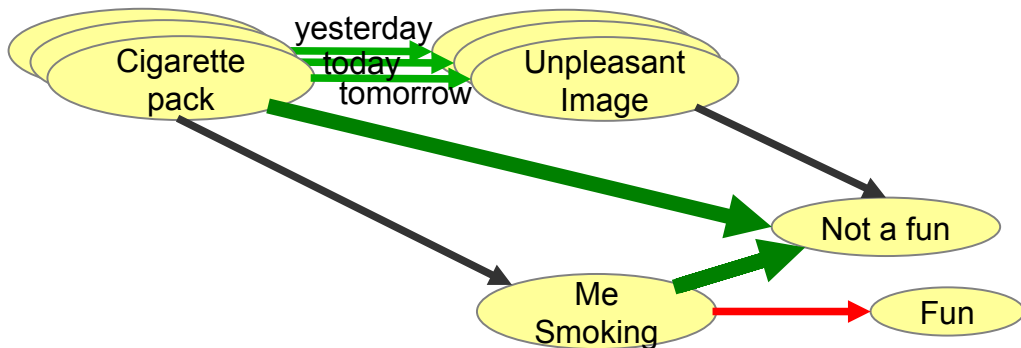
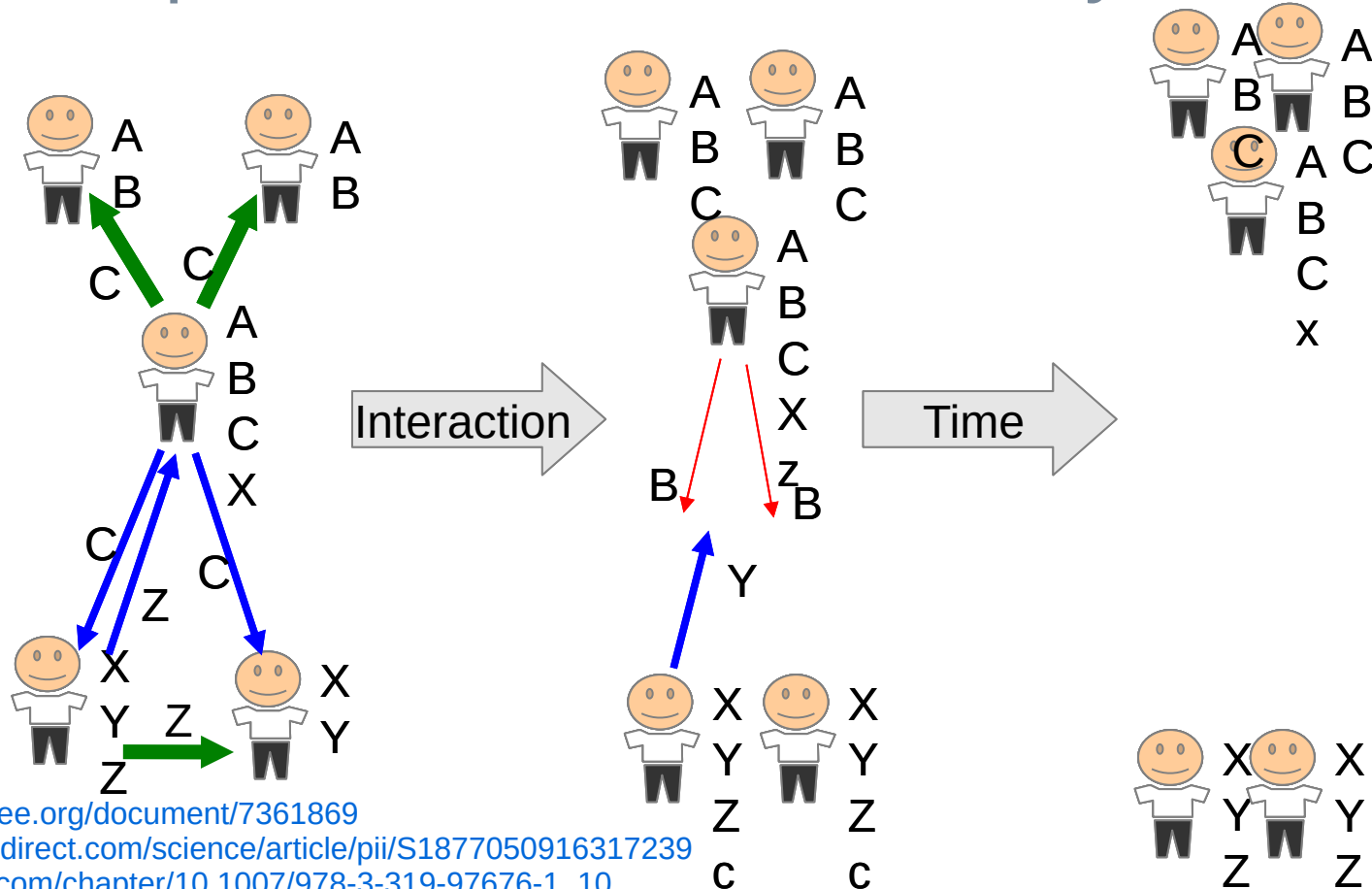## "Highly valable social evidence"

"Redundant social evidence"
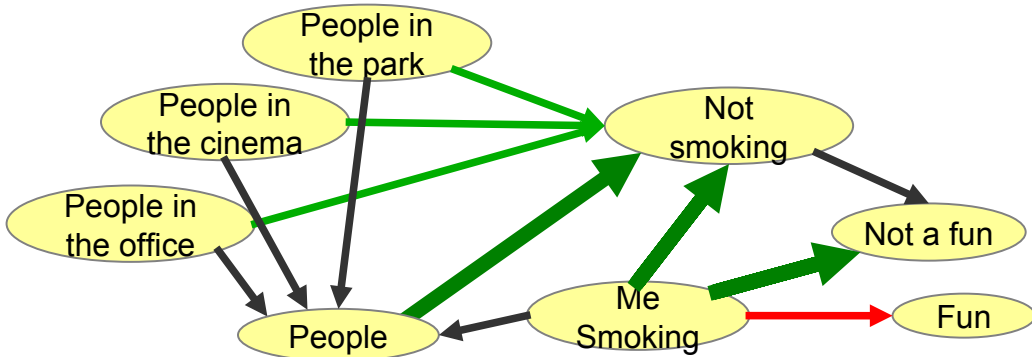
"Implicit evidence injection"
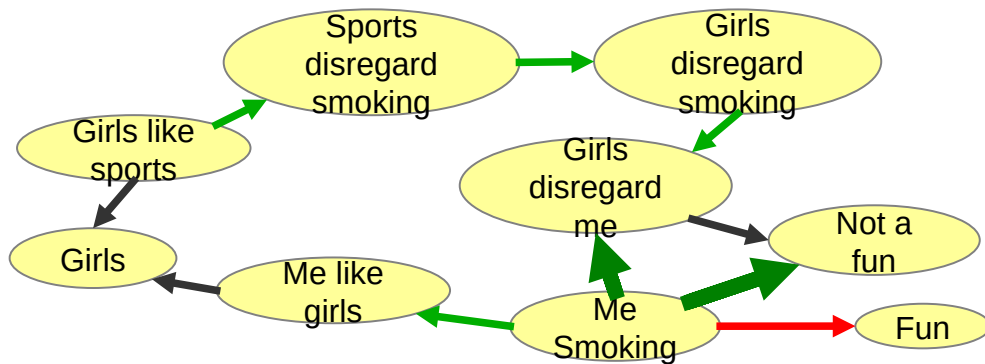
Social evidence-based cognitive-behavioral model
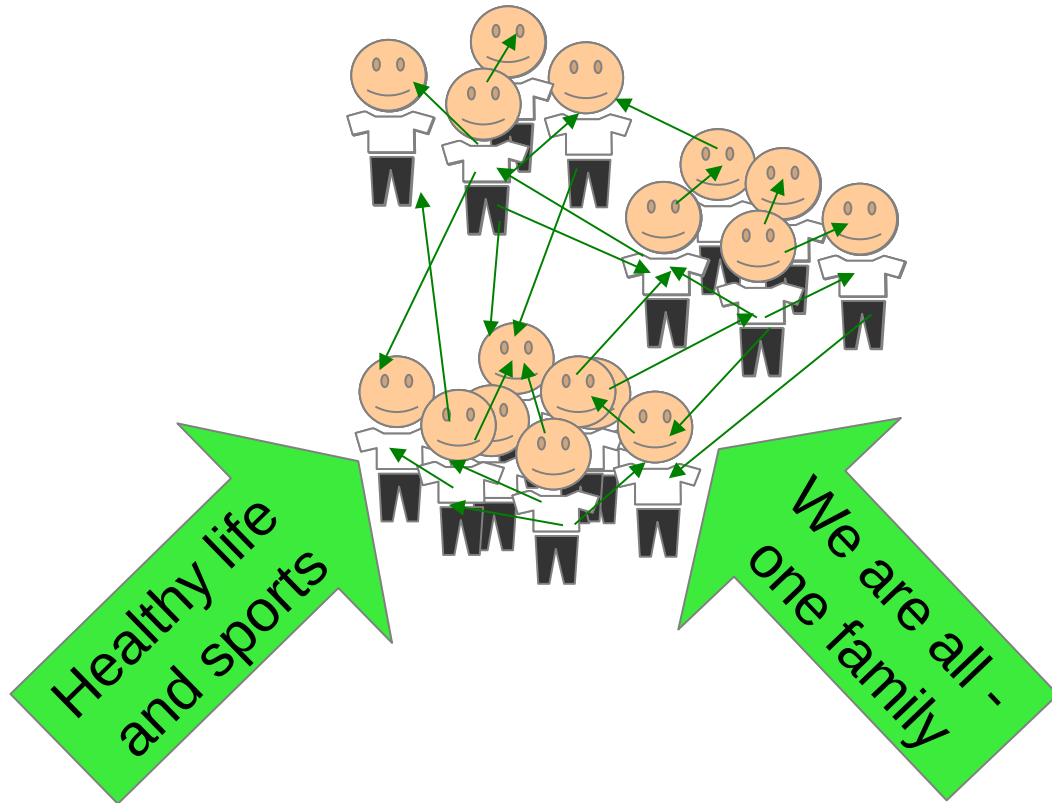## Helps to understand social dynamics

Interaction → Time →

20

Social evidence-based cognitive-behavioral model
Makes it possible to build or destroy societies
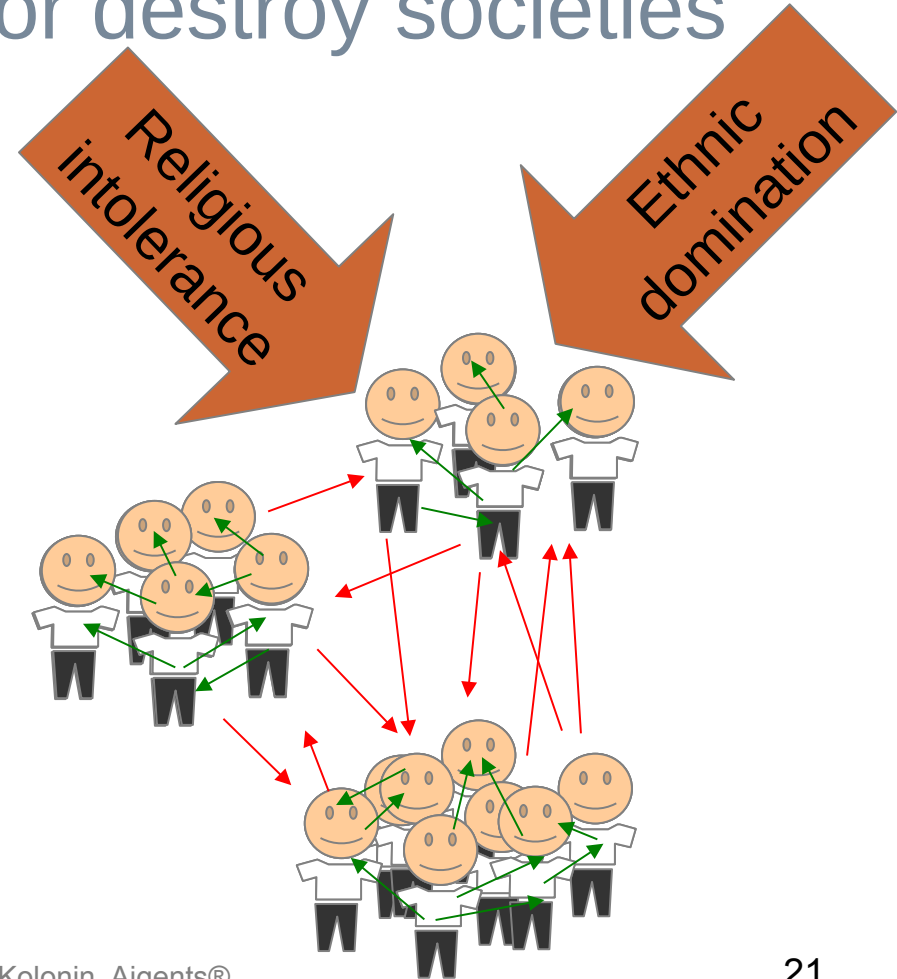
Healthy life and sports

We are all - one family

Religious intolerance

Ethnic domination

21

Networks of Beliefs: An Integrative Theory of Individual- and Social-Level Belief Dynamics

Networks of Belief Theory Premises Illustrated on a Minimal External Network of Two People, Seen From the Perspective of the Focal Person on the Left
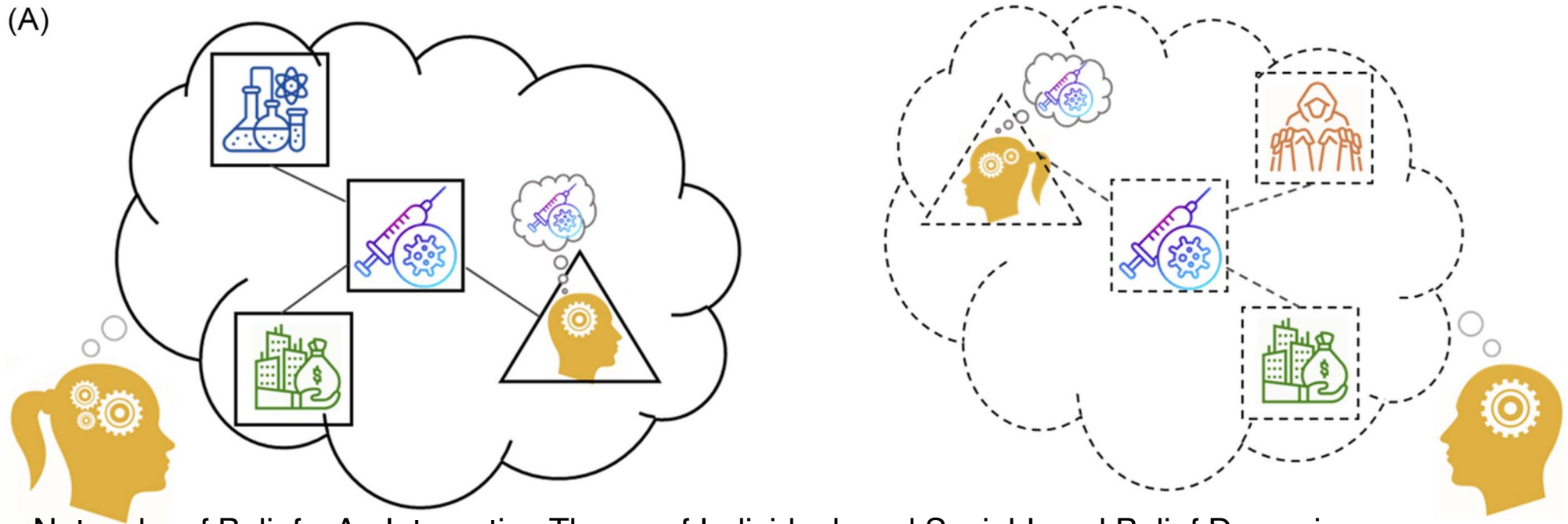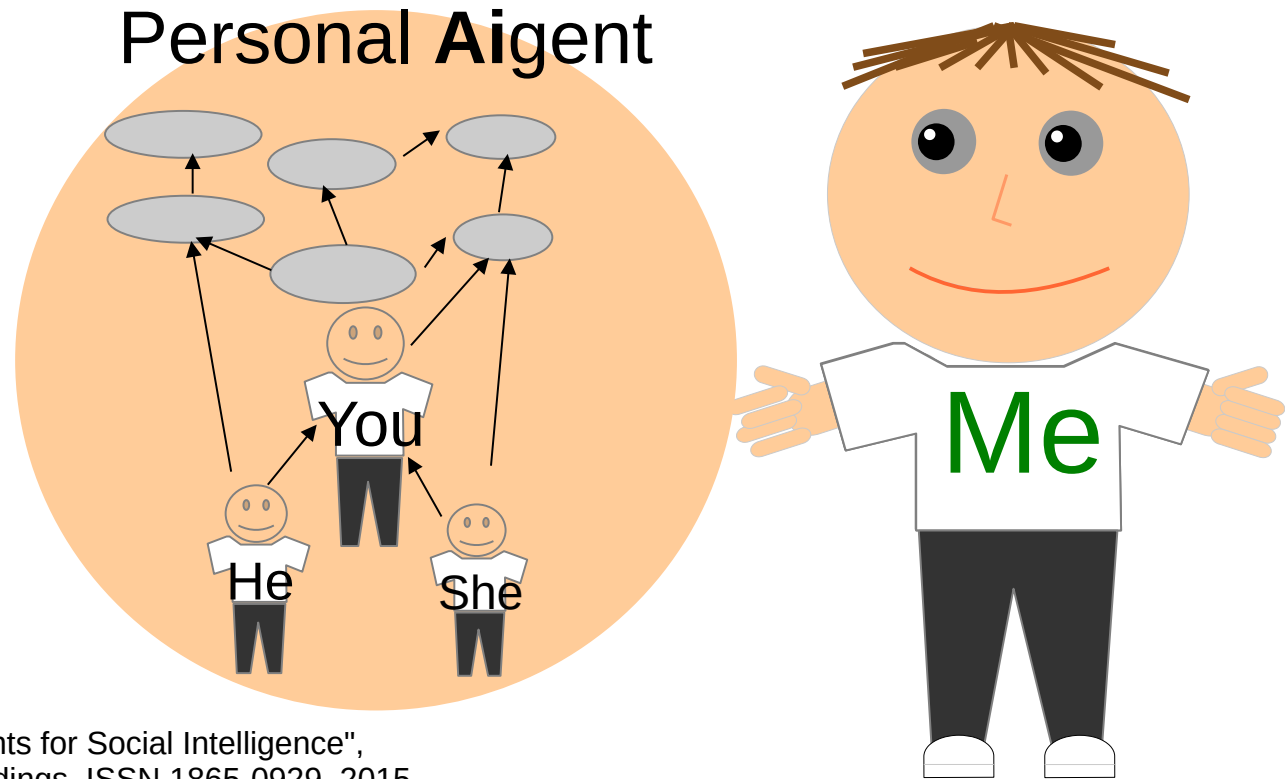

Networks of Beliefs: An Integrative Theory of Individual- and Social-Level Belief Dynamics
Jonas Dalege, Mirta Galesic, and Henrik Olsson
Online First Publication, September 19, 2024. https://dx.doi.org/10.1037/rev0000494

2024

Aigents® - Personal Assistant

Personal **Ai**gent

You

He   She

Me

A. Kolonin, "Aigents: Adaptive Personal Agents for Social Intelligence",
KESW, 6th International Conference Proceedings, ISSN 1865-0929, 2015
https://link.springer.com/chapter/10.1007/978-3-319-24543-0_22

A.Kolonin, "Adaptive experiential learning for business intelligence
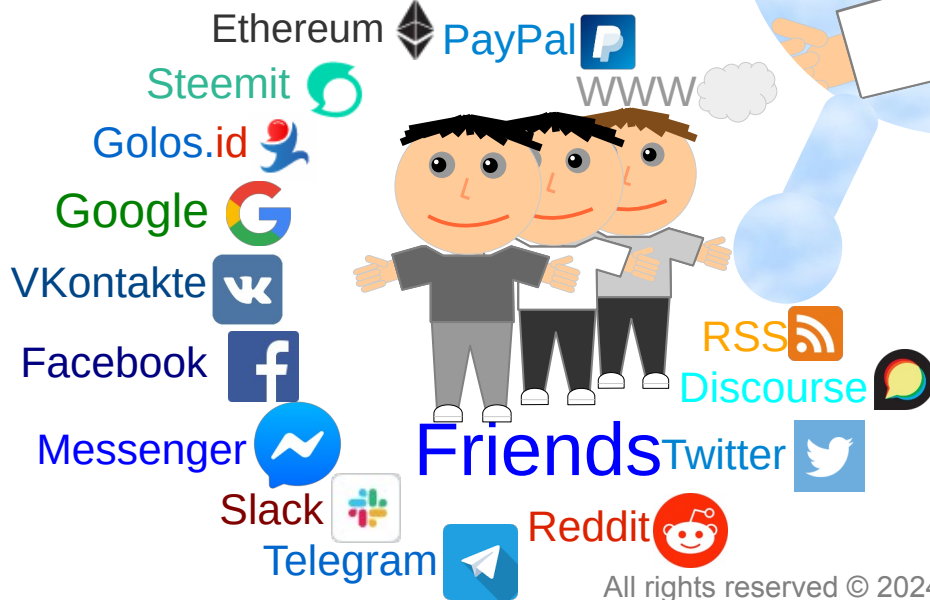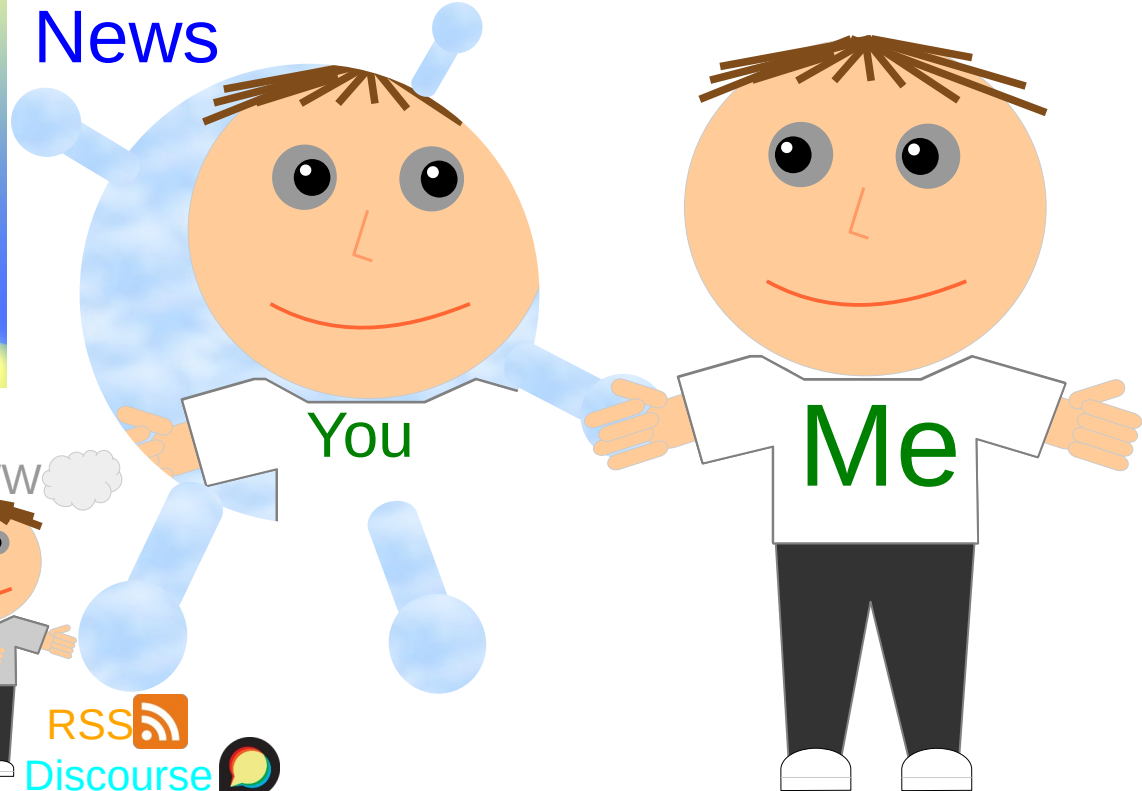agents", CSGB Symposium Proceedings, 2016
https://ieeexplore.ieee.org/document/7587679

2015-2016

Aigents® - Social computing platform

News

You

Me

Ethereum
PayPal
Steemit
Golos.id
Google
VKontakte
Facebook
Messenger
Slack
Telegram

WWW

RSS
Discourse

Friends
Twitter
Reddit

https://aigents.com/
https://github.com/aigents

Thank you for attention!
Full version:

Anton Kolonin

akolonin@aigents.com

Telegram: akolonin

AIGENTS

https://aigents.com

N* Novosibirsk State University

*THE REAL SCIENCE

https://www.nsu.ru

AGI IN RUSSIAN

https://agirussia.org