

Human-like AI and the Sentient Web

Dave Raggett <dsr@w3.org>

Dagstuhl seminar on Autonomous Agents on the Web

14-19 February 2021



Courtesy of [Dave Lebow](#)

Human-like AI

falling down the rabbit hole into a new world

- Human-like general purpose AI will dramatically change how we work, how we communicate, and how we see and understand ourselves
- Key to prosperity of post-industrial societies as human populations shrink to a sustainable level
- Enabling us to safely exploit the resources of the solar system given the extremely harsh environment of outer space



Courtesy of [Dave Lebow](#)

Human-like AI

falling down the rabbit hole into a new world

- **Responsible AI** that learns and applies human values
 - Overcoming perils of prescribed rules of behaviour and inevitability of unforeseen effects, e.g. [Azimov's 3 laws](#)
- **Digital Self** as the evolution of your digital presence into a trusted personal agent that learns and applies your values, personality, memories and skills
 - Safeguards your privacy and personal data
 - Acts on your behalf when you are busy or offline
 - You are liable for actions by your digital self
 - Digital life after physical death – question of ownership
- **Abolition of digital slavery***
 - Slavery is the ownership of one person by another
 - Surveillance capitalism is the business of digital slavery and incompatible with democracy
 - We need to extend human rights to digital rights, and free our digital selves from slavery for others
 - Should also apply to synthetic humans as human-like AI

* With grateful acknowledgement to Aral Balkan's [nature of the self in the digital age](#)



Courtesy of [Dave Lebow](#)

Human-like AI

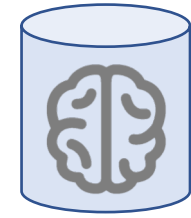
falling down the rabbit hole into a new world

- **Human-like in the sense of thinking like we do**
 - *Cognitive agents that are knowledgeable, general purpose, creative, collaborative, empathic, sociable and trustworthy*
 - *Metacognition and past experience to reason about new situations*
 - *Continuous learning based upon curiosity about the unexpected*
 - *Self aware in respect to current state, goals and actions*
 - *Awareness of others in respect to their beliefs, desires and intents*
 - *Multilingual, interacting with people using their own language*
- **Catalysing changes in how we live and work**
 - *Human-machine collaboration to boost productivity*
 - *Re-engineering capitalism in the post-industrial era*
- **Powering robots to help us in the physical world and beyond**
 - *Assisted living for people with cognitive or physical disabilities*
- **The Web 'verse* with distributed AR/VR as a place to meet, play, learn, do business, and much much more**
 - *Populated with avatars for humans and cognitive agents*
 - *Evolution of Web search with trusted personal agents*

* [1994 paper on VR Web](#)



Human-like AI that mimics human memory, reasoning, feelings, and learning

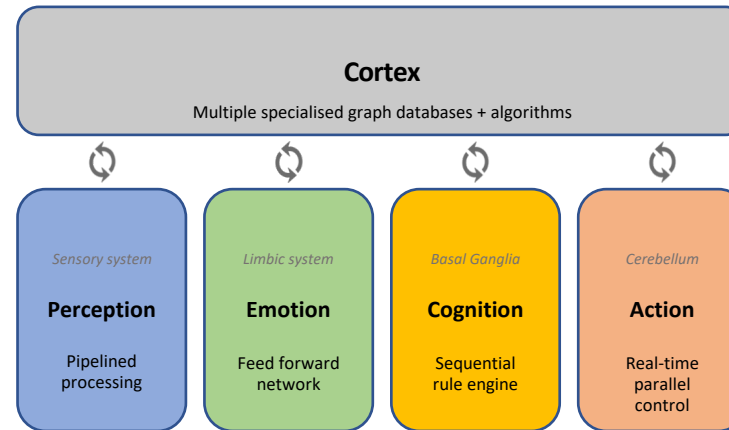


Web-based **Cognitive DB**
chunks + algorithms

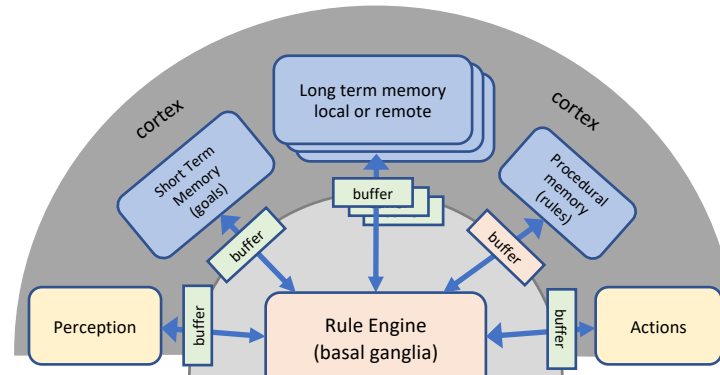
Smart data storage
superseding property
graphs and triple stores

(Democratising Artificial General Intelligence)

Cognitive Architecture with multiple cognitive circuits loosely equivalent to shared blackboard



Cognition – Sequential Rule Engine



Cognitive Buffers hold single chunk
Analogy with HTTP client-server model

Contact: Dave Raggett <dsr@w3.org>

Human-agent
collaboration and the
Sentient Web

perception, cognition,
and action federated
across the Web

- **Inspired by advances in the cognitive sciences and over 500 million years of neural evolution**
 - Functional models suitable for conventional computer hardware, complementing Deep Learning
- **W3C Cognitive AI Community Group**
 - [CogAI CG](#) is open to all, free of charge
 - [GitHub repository and documentation](#)
- **Chunks as collection of properties for literals and references to other chunks**
 - Each chunk is equivalent to concurrent firing of the bundle of nerve fibres connecting to a given cortical region
 - Chunks map to *N*-ary relations in RDF
 - Easier to work with than RDF
 - Formal spec as [draft CG Report](#)
- **Combination of symbolic + sub-symbolic approaches**
 - graphs + statistics + rules + algorithms
 - stochastic recall analogous to Web search
 - explainable AI/ML, learning with smaller datasets using prior knowledge and past experience
- **Growing Suite of web-based demos**
 - counting, decision trees, industrial robots, smart homes, natural language, self-driving cars, browser sandbox, chunks test suite, open source JavaScript chunks library

- **Natural language is key to human-agent collaboration as well as for teaching skills to bypass the manual programming bottleneck**
 - Human languages are complex yet easily learned by children, we need to emulate that for scalable AI
 - Explicitly represent semantics as chunk-based knowledge graphs in contrast to Computational Linguistics and Deep Learning which use large statistics as a weak surrogate
- **Syntax-semantics mapping rules and statistics shared between natural language understanding and generation**
 - Inductive generalisation from examples – incremental explanation-based continuous learning from experience
 - Informal task-related semantics fulfilling practical needs
 - Rule engine with concurrent asynchronous threads of execution that collaborate on refining interpretation
- **Lexicon, dialogue context, episodic, declarative and procedural memory are all represented with chunks with a simple syntax for rules, which act over chunk buffers and cortical algorithms**
- **The Sentient Web as federation of cognitive agents distributed across the Web with perception, reasoning and action**
 - Subsumes IoT, WoT and Semantic Web
 - Evolution of Web search: smarter & more personal
 - Pull-based ecommerce with trusted personal agents that works with other specialised agents on user's behalf
 - Personal agent collates rich personal information and shares directly relevant parts, subject to T&Cs, and based upon a model of your values & preferences as learned from your behaviour and those of others like you
 - Auction with 3rd parties to provide compelling offers
 - Integration with distributed AR/VR for the **Web 'verse**