

# The Long-Overdue Re-Conception of AI and Robotics

**Roger Clarke**

Xamax Consultancy Pty Ltd, Canberra  
Visiting Professor in Technology & Law, UNSW  
Visiting Professor in Computer Science, ANU

<http://www.rogerclarke.com/EC/AITS.html>, AITS-K.pdf

**UNSW AI Institute – 29 August 2023**

Copyright  
2019-23 **XAMAX**  
Consultancy


1


COMPUTER LAW & SECURITY REVIEW 30 (2014) 230–246


Available online at [www.sciencedirect.com](http://www.sciencedirect.com)  
**ScienceDirect**  
www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

**Understanding the drone epidemic** 

**What drones inherit from their ancestors** 

**The regulation of civilian drones' impacts on public safety** 

**The regulation of civilian drones' impacts on behavioural privacy** 

Copyright  
2019-23 **XAMAX**  
Consultancy

Since 2014 >700 citations

2

## Big Data's Big Unintended Consequences

Marcus R. Wigan, *Oxford Systematics, Swinburne University, and the University of Melbourne*  
Roger Clarke, *Xamax Consultancy, University of New South Wales, and Australian National University*

COMPUTER Published by the IEEE Computer Society 0018-9142/13/3311.00 © 2013 IEEE

*Big data, big risks*

*Info Systems J* (2016) 26, 77–90

COMPUTER LAW & SECURITY REVIEW 16 (2018) 467–476

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)  
**ScienceDirect**  
www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

**Guidelines for the responsible application of data analytics** 

31<sup>st</sup> BLEED CONFERENCE: DIGITAL TRANSFORMATION: MEETING THE CHALLENGES  
JUNE 17–20, 2018, BLEID, SLOVENIA, CONFERENCE PROCEEDINGS  
A. Pucihar, M. Kljajić Borštnar, P. Ravestejn, J. Setz & R. Bons

**Towards Responsible Data Analytics: A Process Approach**  
ROGER CLARKE & KERRY TAYLOR

**Do Ethical Guidelines have a Role to Play in Relation to Data Analytics and AI/ML?**  
For AiCE 2020, UniSA, Adelaide, November 2020

Copyright  
2019-23 **XAMAX**  
Consultancy

Since 2013/16 >500 citations

3

COMPUTER LAW & SECURITY REVIEW 35 (2019) 423–433

**Why the world wants controls over Artificial Intelligence**

**Principles and business processes for responsible AI**

**Regulatory alternatives for AI**

Responsible application of artificial intelligence to surveillance: What prospects?<sup>1</sup>

Information Polity 27 (2022) 175–191

IEEE TRANSACTIONS ON TECHNOLOGY AND SOCIETY, VOL. 4, NO. 1, MARCH 2023

**The Re-Conception of AI:  
Beyond Artificial, and Beyond Intelligence**

Copyright  
2019-23 **XAMAX**  
Consultancy

Since 2019 ~200 citations

4

## The Original Conception of Artificial Intelligence (AI Old)

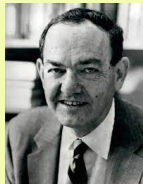


- Based on "the conjecture that every aspect of **learning or any other feature of intelligence** can in principle be so precisely described that a machine can be made to simulate it"
- "The hypothesis is that a physical symbol system [of a particular kind] has the necessary and sufficient means for **general intelligent action**"

## The Original Conception of Artificial Intelligence (AI Old)



- Based on "the **conjecture** that every aspect of learning or any other feature of intelligence **can in principle** be so precisely described that a machine can be made to **simulate** it"
- "The **hypothesis** is that a physical symbol system [of a particular kind] has the necessary and sufficient means for **general intelligent action**"



## From Conjecture and Hypothesis To Belief

"Within the very near future - much less than twenty-five years - **we shall have** the technical capability of substituting machines for any and all human functions in organisations.

"Duplicating problem-solving and information-handling capabilities of the brain **is not far off** ... surprising if it were not accomplished within the next decade" (1960)

"By the end of the 2020s [computers **will have**] intelligence indistinguishable to biological humans" (2005)

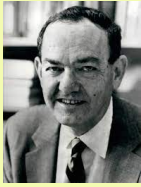


## From Conjecture and Hypothesis To Belief

"**Within** the very near future - **much less than twenty-five years** - we shall have the technical capability of substituting machines for any and all human functions in organisations.

"Duplicating problem-solving and information-handling capabilities of the brain is not far off ... surprising if it were not accomplished **within the next decade**" (1960)

"**By the end of the 2020s** [computers will have] intelligence indistinguishable to biological humans" (2005)



## From Conjecture and Hypothesis To Belief

"Within the very near future - much less than twenty-five years - we shall have the technical capability of **substituting machines for any and all human functions** in organisations.

"**Duplicating problem-solving and information-handling capabilities of the brain** is not far off ... surprising if it were not accomplished within the next decade" (1960)

"By the end of the 2020s [computers will have] **intelligence indistinguishable to biological humans**" (2005)

## Bifurcation of the Field

- The 'grand challenge' aspect:  
'Artificial general intelligence' or 'Strong AI'  
Aspiration to replicate human intelligence
- Human intelligence as Inspiration  
'Weak AI' / 'Narrow AI'

## Separation But Not Divorce

## How to Recognise 'an AI'

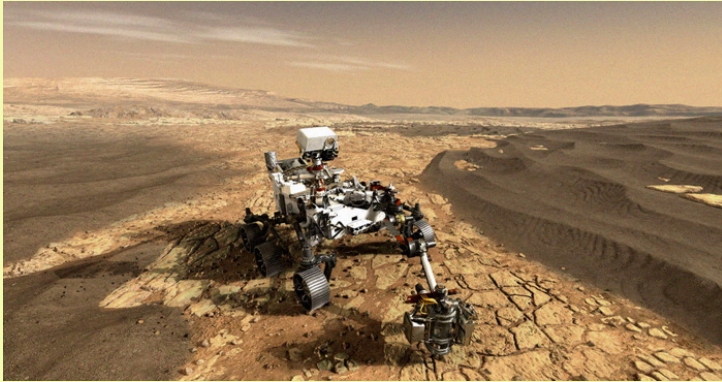
*Intelligence is exhibited by an artefact if it:*

- (1) *evidences **perception and cognition** of relevant aspects of its environment*
  - (2) *has **goals**; and*
  - (3) ***formulates actions** towards the achievement of those goals*
- and?*
- (4) ***implements those actions***

## Embodiments of AI

- **Computers**
- **Robots**  
'A Computer that Does' &  
'A Machine that Computes'
- **Humanoid Robots**  
Androids  
Gynoids / Fembots
- **Vehicles**  
Terrestrial  
– Road, Rail, Off-Road  
Airborne  
Water-borne, Submerged
- **Bus-Stops**  
And other everyday Things
- **Cyborgs**  
A Human whose natural capabilities have been enhanced by technological means  
  
A Hybrid of a human and one or more associated, attached or embedded artefacts

## 'Terrestrial', Off-Road, Remote



Copyright  
2019-23 **XAMAX**  
Consultancy

Aug 2021 – <https://futurism.com/the-byte/nasas-mars-rover-took-selfie-beautiful>

13

## Mechanical Performance of such Challenging Physical Tasks is GOOD

Copyright  
2019-23 **XAMAX**  
Consultancy

14

## Mechanical Performance of such Challenging Physical Tasks is GOOD

### But Intelligence also requires Second-Order Intellect or Insight

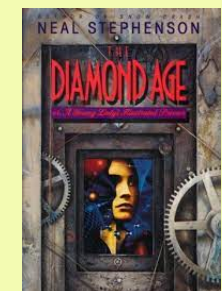
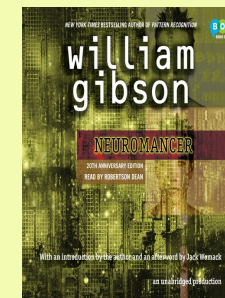
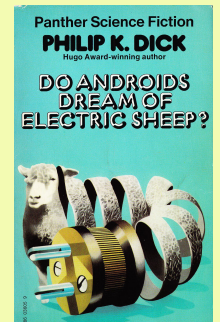
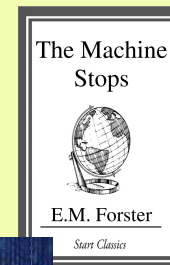
- Values-Driven Formulation of Goals
- Common-Sense Understanding of Context
- Detection of Changes of Relevance
- Ongoing Re-Evaluation of Values
- Ongoing Adaptation of Goals

Copyright  
2019-23 **XAMAX**  
Consultancy

Dreyfus H.L. (1972)  
Weizenbaum J. (1976)

15

## Science Fiction Anticipates Reality



Copyright  
2019-23 **XAMAX**  
Consultancy

16

## AI Sceptics are in Good Company



## A Distillation of the Threats Inherent in AI

1. **Artefact Autonomy**  
Substantial delegation from humans to non-humans
2. **Inappropriate Assumptions about Data**  
Data selectivity, interpolation, incompatibility, quality
3. **... and about the Inferencing Process**  
Uncontrolled environments, unmodelled systems
4. **Opaqueness of the Inferencing Process**  
Unexplainability, procedural fairness, unaccountability
5. **Irresponsibility**  
Everyone in the chain points at everyone else

## Degrees of Autonomy

		Function of the Artefact	Function of the Human
	0	<b>NIL</b>	<b>Analyse, Decide, Act</b>
Decision Support System	1	Analyse Options	<b>Analyse, Decide, Act</b>
	2	Advise re Options	<b>Analyse, Decide, Act</b>
	3	Recommend Act	<b>Analyse, Approve/Reject Act</b>
Decision System	4	<b>Notify Impending Act</b>	Override/Veto Impending Act
	5	<b>Act and Inform</b>	Interrupt/Suspend/Cancel an Act
	6	<b>Act</b>	<b>NIL</b>

Armstrong (2010, p.14),  
Sheridan & Verplank (1978, Table 8.2, pp. 8-17-8.19)  
as interpreted by Robertson et al. (2019, Table 1)

## The Threats Inherent in AI

1. **Artefact Autonomy**  
Substantial delegation from humans to non-humans
2. **Inappropriate Assumptions about Data**  
Data selectivity, interpolation, incompatibility, quality
3. **... and about the Inferencing Process**  
Uncontrolled environments, unmodelled systems
4. **Opaqueness of the Inferencing Process**  
Unexplainability, procedural fairness, unaccountability
5. **Irresponsibility**  
Everyone in the chain points at everyone else



## Data & Information Quality Factors

### Assessable at time of collection

- D1 – Syntactic Validity
- D2 – Appropriate (Id)entity Association
- D3 – Appropriate Attribute Association
- D4 – Appropriate Attribute Signification
- D5 – Accuracy
- D6 – Precision
- D7 – Temporal Applicability

### Assessable only at time of use

- I1 – Theoretical Relevance
- I2 – Practical Relevance
- I3 – Currency
- I4 – Completeness
- I5 – Controls
- I6 – Auditability

## The Threats Inherent in AI

1. **Artefact Autonomy**  
Substantial delegation from humans to non-humans
2. **Inappropriate Assumptions about Data**  
Data selectivity, interpolation, incompatibility, quality
3. **... about the Inferencing Process**  
Uncontrolled environments, unmodelled systems
4. **Opaqueness of the Inferencing Process**  
Unexplainability, procedural fairness, unaccountability
5. **Irresponsibility**  
Everyone in the chain / network points at everyone else

## Assumptions Often Implicit in AI/ML

- An underlying model of reality
- Near-enough correspondence with reality
- Adequate training-set quality
- Adequate data-item quality
- Adequate data-item correspondence to the phenomenon it purports to represent
- No material training-set bias
- No learning algorithm bias
- Compatibility of data and 'model'
- Logically valid inferences
- Empirically checked inferences

## Risk Factors in AI/ML

- **Insufficient, active and careful modelling**  
of real-world problem-solutions, problems, or problem-domains  
cf. lists of input and output variables, (plus intermediating/hidden variables, if 'deep')  
cf. implicit variables ('unsupervised' ML)
- **No explicit, designed-in real-world relationship**  
And/or inadequate audit of the relationship
- **Loss of the Theory-Empiricism partnership**  
i.e. Empiricism may dominate Theory

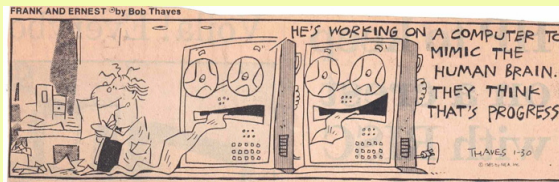
## Socio-Political Impacts and Implications

- **De Facto Delegation**  
"The computer says no"
- **Unexplainability**  
Accountability Undermined
- **Unfair Decisions, Actions**  
Discriminatory Behaviour
- **Economic, Social Scoring**  
Non-Conformist Victimisation
- **Undefendable Accusations**  
Power, Information Asymmetry
- **'Predestination'**  
Predictive Policing
- **People-Replacement**  
Effect on Income Distribution
- **Denial of Services, of Movement, of Identity**  
Public Resentment, Violence

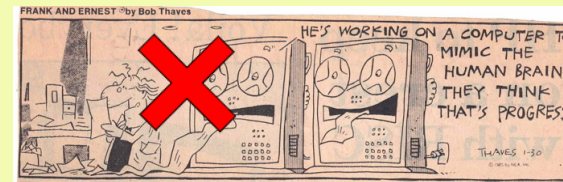
## 'Artificial'? Or 'Artefactual'? 'Intelligence' What Do We Want From It?

- There are 8 billion people and we're multiplying (too) fast
- Why would we want yet more Natural Intelligence?

## 'Artificial'? Or 'Artefactual'? 'Intelligence' What Do We Want From It?



## 'Artificial'? Or 'Artefactual'? 'Intelligence' What Do We Want From It?



- Do things well that humans do poorly, or cannot do at all
- Perform functions within systems that include both humans and artefacts
- Interface effectively, efficiently and adaptably with both humans and other artefacts

## ChatGPT / LLM's Achilles Heel

- Unseptical and unbridled enthusiasm was quickly followed by recriminations:
  - Gamma testers conducted serious testing
  - Students submitted mistaken assignments
  - Journals required declarations of 'no LLM'
  - Lawyers submitted briefs with invented cases
  - ARC Assessors submitted facile reports

## ChatGPT / LLM's Achilles Heel

- Unseptical and unbridled enthusiasm was quickly followed by recriminations:
  - Gamma testers conducted serious testing
  - Students submitted mistaken assignments
  - Journals required declarations of 'no LLM'
  - Lawyers submitted briefs with invented cases
  - ARC Assessors submitted facile reports

### Government warns on generative AI use

Don't use ChatGPT to make decisions, write code, or prepare tenders.

By David Braue on Jul 11 2023 10:56 AM

## ChatGPT / LLM's Achilles Heel

- Unseptical and unbridled enthusiasm was quickly followed by recriminations:
  - Gamma testers conducted serious testing
  - Students submitted mistaken assignments
  - Journals required declarations of 'no LLM'
  - Lawyers submitted briefs with invented cases
  - ARC Assessors submitted facile reports
  - Aust Govt places tight limits on its use
- It was designed as a Decision Tool
- It should be designed as a Decision Support Tool

Human  
Intelligence



Augmented  
Intelligence



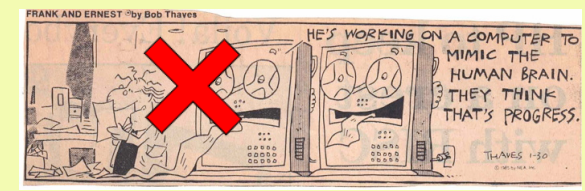
## Augmented Intelligence

- Ashby (1956) on 'intelligence amplification'
- Engelbart (1962) on 'augmenting human intellect'
- Mann (2001) on wearable/body-borne computing, augmented / diminished / mediated reality, sur- / sous- / meta- / equi-veillance, ...
- Araya (2019) on 'augmented intelligence' as "an alternative conceptualization of AI that focuses on its assistive role in advancing human capabilities"
- IEEE Council on Extended Intelligence (2017-19) "it is not AI in isolation, but the social, economic, political, and cultural systems within which these tools are integrated that must be addressed to avoid reductionist outcomes"



Human Intelligence & Complementary Artefactual Intellectics → Augmented Intelligence

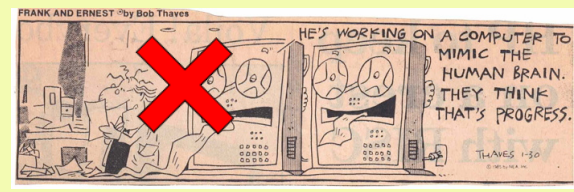
## Complementary Artefactual Intellectics What Do We Want From It?



Do things well that humans do poorly, or cannot do at all:

- Dull
- Dirty
- Dangerous

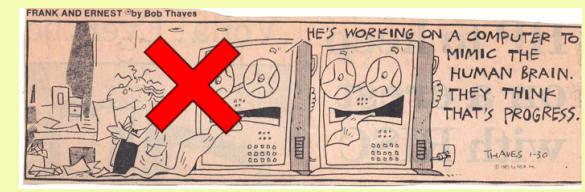
## Complementary Artefactual Intellectics What Do We Want From It?



Do things well that humans do poorly, or cannot do at all:

- Dull
- Dirty
- Dangerous
- P
- S

## Complementary Artefactual Intellectics What Do We Want From It?



Do things well that humans do poorly, or cannot do at all:

- Dull
- Dirty
- Dangerous
- Precision
- Speed

Effectors & Actuators ⇒ Capability



Human Intelligence & Complementary Artefactual Intellectics ⇒ Augmented Intelligence  
 & & &  
 Effectors & Actuators ⇒ Action Capability

Human Intelligence & Complementary Artefactual Intellectics ⇒ Augmented Intelligence  
 & & &  
 Effectors & Actuators ⇒ Action Capability  
 ↓ ↓ ↓  
 Human Action Capability & Complementary Artefactual Capability ⇒ Augmented Action Capability

Means & Ends  
 Human Intelligence & Complementary Artefactual Intellectics ⇒ Augmented Intelligence  
 & & &  
 Effectors & Actuators ⇒ Action Capability  
 ↓ ↓ ↓  
 Human Action Capability & Complementary Artefactual Capability ⇒ Augmented Action Capability

Human  
Intelligence

&

Complementary  
Artefactual  
Intellectics



Augmented  
Intelligence

&

&

&

Effectors

&

Actuators



Action  
Capability



Human  
Action  
Capability

&

Complementary  
Artefactual  
Capability



Augmented  
Action  
Capability

Which of these is/are Robotics cf. Cobotics ?  
And 'AI-based autonomous systems'?

## The Long-Overdue Re-Conception of AI and Robotics

**Roger Clarke**

Xamax Consultancy Pty Ltd, Canberra  
Visiting Professor in Technology & Law, UNSW  
Visiting Professor in Computer Science, ANU

<http://www.rogerclarke.com/EC/AITS.html>, AITS-K.pdf

**UNSW AI Institute – 29 August 2023**