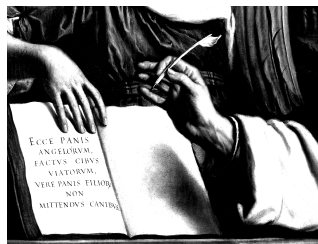


Artificial intelligence meets theism: philosophical reflections

Soenke Ziesche

DIVUS THOMAS



Anno 126° - 2023 - maggio/agosto

Artificial intelligence meets theism: philosophical reflections

Contributi di

Jennifer Ang

Brian Besong (ed.)

David Černý (ed.)

Arvin Gouw

Roman V. Yampolskiy

Michael LaBossiere

Michael McCarthy

Christopher M. Reilly

Jacob Sparks

Margarita Vega

Ava Thomas Wright

Soenke Ziesche

DIVUS

ISSN 0012-4257

Periodico quadrimestrale dello Studio Filosofico Domenicano,
della Provincia San Domenico in Italia
Via dell'Osservanza, 72
40136 Bologna BO

Tel. ++39 051582034

acquisti@esd-domenicani.it

www.edizionistudiodomenicano.it

Autorizzazione del Tribunale di Firenze del 19 aprile 1948 n. 13

Direttore: MARCO SALVIOLI - divusthomasdirettore@esd-domenicani.it

Comitato editoriale: Alberto Ambrosio, Giuseppe Barzaghi (direttore responsabile),
Giovanni Bertuzzi, Giorgio Carbone, Diana Mancini,
Flavio Minoli, Tommaso Reali.

Collaboratori: Fernando Bellelli, Erio Castellucci, David Černý,
Alberto Cevolini, Marianna Rascente, Claudio Testi.

TUTTI I DIRITTI SONO RISERVATI

© 2023 - Edizioni Studio Domenicano - www.edizionistudiodomenicano.it

I diritti di traduzione, di memorizzazione elettronica, di riproduzione e di adattamento totale o parziale con qualsiasi mezzo, compresi i microfilm, le fotocopie e le scannerizzazioni, sono riservati per tutti i Paesi.

Le fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% del volume dietro pagamento alla SIAE del compenso previsto dall'art. 68, commi 4 e 5, della legge 22/04/1941, n. 633.

Le riproduzioni diverse da quelle sopra indicate, e cioè le riproduzioni per uso non personale (a titolo esemplificativo: per uso commerciale, economico o professionale) e le riproduzioni che superano il limite del 15% del volume possono avvenire solo a seguito di specifica autorizzazione scritta rilasciata dall'Editore oppure da AIDRO, Corso di Porta Romana 108, 20122 Milano, segreteria@aidro.org

L'elaborazione dei testi, anche se curata con scrupolosa attenzione, non può comportare specifiche responsabilità per eventuali involontari errori o inesattezze.

DIVUS THOMAS

Rivista quadrimestrale

ABBONAMENTI 2023

	<i>un anno</i>	<i>due anni</i>
Italia ordinario biblioteche, enti, agenzie ecc.	€ 100,00	€ 160,00
Italia ridotto persone fisiche e privati	€ 60,00	€ 100,00
Esteri ordinario biblioteche, enti, agenzie ecc.	€ 190,00	€ 330,00
Esteri ridotto persone fisiche e privati	€ 150,00	€ 240,00
Serie completa 1924 - 2022, sconto 70%	€ 9.270,00	€ 2.781,00
Supplemento per sottoscrizione successiva al 15 marzo		€ 30,00
Serie completa 1992 - 2022, sconto 50%	€ 2.790,00	€ 1.395,00
numero singolo € 30,00		

I singoli quaderni si possono acquistare anche presso l'Editore.

PAGAMENTI

Bonifico bancario

c/c numero	12971404
tenuto presso	Poste Italiane SpA
intestato a	Edizioni Studio Domenicano
IBAN	IT 49 W 07601 02400 000012971404
BIC	B P P I I T R R X X X

Bollettino postale

ccp	12971404
intestato a	Edizioni Studio Domenicano

Non si accettano assegni bancari.

Il contratto di abbonamento ha durata di 1 o 2 anni, cessa con l'invio dell'ultimo numero di annata. Il rinnovo utile ad assicurare la continuità deve essere effettuato con versamento della quota entro il 15 marzo del nuovo anno.

All'abbonamento sottoscritto dopo il 15 marzo si applica un supplemento di € 30,00. Per la sottoscrizione scrivere a: acquisti@esd-domenicani.it

Il reclamo per la mancata ricezione di un quaderno va fatto entro tre mesi dalla ricezione del quaderno successivo a quello di cui si lamenta la mancata consegna.

SOMMARIO

DAVID ČERNÝ – BRIAN BESONG

Introduction	15
---------------------	----

MONOGRAFIE

MARGARITA VEGA

Aquinas walks into the chinese room	17
--	----

Introduction	17
--------------	----

Searching for the bare minimum of all things mental	19
---	----

Forty years inside the chinese room	25
-------------------------------------	----

The ontology of the mental	31
----------------------------	----

Thesis one: zero-time in artificial intelligence	40
--	----

Thesis two: proportional causality	44
------------------------------------	----

Information and panpsychism	47
-----------------------------	----

The emergent properties of information processing	51
---	----

Life as a proportional cause for zero-time	60
--	----

ABSTRACT	62
----------	----

RIASSUNTO	62
-----------	----

ARVIN GOUW

Of angels and robots: knowledge and purpose in AI from thomistic passions to Chalmer's qualia	64
--	----

Aquinas' theory of knowledge and purpose	64
--	----

<i>Two approaches to intelligence in AI</i>	67
---	----

Symbolic and embodied AIs	67
---------------------------	----

<i>Symbolic AI</i>	67
--------------------	----

General vs. specialized AI	68
----------------------------	----

Angels and robots	70
-------------------	----

Angelic and robot's fall	71
--------------------------	----

<i>Embodied AI</i>	73
The notion of passion as movement	74
The passions of embodied robots	75
Qualia and artificial intelligence	77
<i>Functional consciousness and qualia distinction</i>	77
<i>Logical vs. natural supervenience of consciousness and reductive materialism</i>	78
Logical supervenience	78
Natural supervenience	79
Arguments for the natural supervenience of consciousness	80
<i>Psychophysical law of organizational invariance</i>	81
The principle of organizational invariance	82
Fading and dancing qualia thought experiments	83
Current artificial qualia constructions	84
Implementing fading and dancing qualia experiments as qualia detection methods	85
Concluding remarks	86
ABSTRACT	87
RIASSUNTO	88
 MICHAEL LABOSSIERE	
On digital souls & God's robots: immortality through AAI?	90
Introduction	90
Locke is key	91
Descartes & God's robots	99
Conclusion	108
ABSTRACT	108
RIASSUNTO	109

MICHAEL MCCARTHY

The function of sentient intelligence and self-constitution for superintelligence	111
Introduction	111
Artificial intelligence: clarifying terms	112
The end of intelligence	115
Sentient intelligence for superintelligence	118
The ethics of superintelligence	122
Conclusion: exploring the value of superintelligence	127
ABSTRACT	128
RIASSUNTO	129

SOENKE ZIESCHE – ROMAN V. YAMPOLSKIY

The problem of AI identity	131
Introduction	131
Motivation	132
First approach	134
<i>The personal identity problem</i>	135
<i>Proposed definition</i>	136
<i>Concept of computational irreducibility</i>	137
Second approach	139
<i>Multi-factor authentication - introduction</i>	139
<i>Multi-factor authentication - verification of AI identity</i>	140
Further scenarios	142
<i>Replication/fission</i>	142
<i>Fission/swarm</i>	142
<i>Switch off</i>	143
<i>Resurrection / switch on again</i>	144
<i>Change of hardware / substrate</i>	145

<i>Transition from non-sentient to sentient (and vice versa)</i>	145
<i>Journey to the past</i>	146
<i>Offspring</i>	146
<i>Identity change</i>	147
Summary and future work	147
ABSTRACT	150
RIASSUNTO	150
CHRISTOPHER M. REILLY	
How artificial intelligence technology encourages the vice of acedia	152
Introduction	152
Acedia and instrumental rationality	153
Instrumental rationality as a path to acedia	155
Materialization of instrumental rationality in AI	161
Manifestations of acedia in the AI-driven environment	169
Putting it all together	174
ABSTRACT	175
RIASSUNTO	175
JENNIFER ANG	
The problem of the responsibility of the “nobody”	176
Introduction	176
Responsibility gap and agent-centric arguments	177
Authentic autonomy and deflectors of responsibility	185
Moral responsibility as relational	193
Conclusion	196
ABSTRACT	197
RIASSUNTO	198

JACOB SPARKS – AVA THOMAS WRIGHT	
Human-centered AI: the aristotelian approach	200
The human in human-centered AI	200
Economic and aristotelian models	203
Failures of rationality	206
<i>Bad ends</i>	206
<i>Incontinence</i> (akrasia)	208
<i>Mere continence</i>	210
<i>Wantonness</i>	211
Doubts about the aristotelian approach: control, paternalism, and feasibility	213
Conclusion	217
ABSTRACT	217
RIASSUNTO	218

SAGGI

GIUSEPPE BARZAGHI O. P.	
Eternal design-mystery: the simultaneity of inspection	221
Self-knowledge and the cosmos	221
The ontological whole	228
<i>The whole</i>	234
The theology as science and anagogical or <i>sub specie æternitatis</i> vision	236
<i>Esemplifications of anagogic and theological syllogism</i>	239
Conclusion	240

FRANCESCO DE CAROLIS	
Il pensiero di Luigi Pareyson: dal personalismo all'ermeneutica della coscienza religiosa cristiana. Nota critico-bibliografica	243
RECENSIONI	259
LIBRI RICEVUTI	268

MONOGRAFIE

INTRODUCTION

DAVID ČERNÝ^{*} – BRIAN BESONG^{**}

Artificial intelligence is a fascinating subject of philosophical investigation. Much recent philosophical analysis has been devoted to elucidating the basic concepts of AI and shedding new light on the emerging ethical issues associated with the development and deployment of various AI-based systems. However, much less attention has been paid to the equally intriguing relationship between this emerging technology and theistic philosophical perspectives. AI is developing at an unprecedented rate, constantly acquiring new sophisticated cognitive capabilities. This development raises many intriguing questions that a theistic philosopher should feel obliged to answer. For example, if AI were to exhibit intelligent behaviour that is almost indistinguishable from human behaviour, or if AI were to exhibit behaviour traditionally associated with free will, what would be the implications for the traditional theistic understanding of human beings as free cognitive agents? This collection of texts cannot, of course, address all the questions that arise from considering AI within the framework of theistic philosophy. But it represents one of the first attempts to take advances in AI seriously and to subject them to theism-inspired philosophical scrutiny.

^{*} *Institute of Computer Science, Czech Academy of Sciences.*

Work on this special issue was supported by the AV21 Strategy Grant Project, research programme “Breakthrough Technologies for the Future – Sensing, Digitalization, Artificial Intelligence and Quantum Technologies”, Research theme “Philosophy and the Future of Artificial Intelligence”.

^{**} *Saint Francis University, Loretto, Pa., Wolf-Kuhn Ethics Institute.*

L'intelligenza artificiale è un affascinante oggetto di indagine filosofica. Molte analisi filosofiche recenti sono dedicate a chiarire i concetti di base dell'IA e a gettare nuova luce sulle questioni etiche emergenti associate allo sviluppo e alla diffusione di vari sistemi basati sull'IA. Tuttavia, molta meno attenzione è stata dedicata alla relazione altrettanto intrigante tra questa tecnologia emergente e le prospettive filosofiche teistiche. L'IA si sta sviluppando a un ritmo senza precedenti, acquisendo costantemente nuove e sofisticate capacità cognitive. Questo sviluppo solleva molte domande intriganti a cui un filosofo teista dovrebbe sentirsi in dovere di rispondere. Ad esempio, se l'IA dovesse esibire un comportamento intelligente quasi indistinguibile da quello umano, o se dovesse esibire un comportamento tradizionalmente associato al libero arbitrio, quali sarebbero le implicazioni per la tradizionale comprensione teistica degli esseri umani come agenti cognitivi liberi? Questa raccolta di testi non può, ovviamente, affrontare tutte le domande che sorgono dal considerare l'IA nel quadro della filosofia teistica. Ma rappresenta uno dei primi tentativi di prendere sul serio i progressi dell'IA e di sottoporli a un esame filosofico ispirato al teismo.

AQUINAS WALKS INTO THE CHINESE ROOM

MARGARITA VEGA*

INTRODUCTION

Can we confidently say that intelligent machines are not conscious and do not have agency? In 1980, the Chinese Room Argument (CRA) by John R. Searle¹ attempted to refute the strong version of Artificial Intelligence², which defended the view that computers can think. Almost forty years later, still many base their intuitions that humans and computers differ vastly on the CRA's theses of "semantics is not syntax" and "syntax is not physics". However, for many others, the CRA never adequately refuted computationalism, and the possibility of computers that think and are conscious seems more relevant than ever before.

Our digital technology has gone from modest and rudimentary outputs to more sophisticated ones. Caught by the euphoria of the digital revolution, the thrust of new generations raised under the paradigm of new digital technologies, and popular culture, the familiarity with the idea that computers do not think has become

* Dominican School of Philosophy & Theology, California, USA.

¹ Searle envisioned this thought experiment during his flight to attend Roger Schank's Yale AI research group's meeting (see G. FAIGENBAUM, *Conversations with John Searle*, Libros en red, Montevideo 2001, p. 78). The CRA was later published as *Minds, Brains, and Programs*, in «Behavioral and Brain Sciences» 3, 3 (September 1980), pp. 417-57.

² «According to strong AI, the appropriately programmed computer with the right inputs and outputs literally has a mind in exactly the same sense that you and I do», J. SEARLE, *Minds and Brains without Programs*, in «Mindwaves» 1987, pp. 209-33, 45.

invece la distinzione e le specificità dei due approcci teoretici, rivalutando l'originalità e la pregnanza delle argomentazioni che caratterizzano la via bontadiniana. Oltre a ciò, il testo affronta anche altre questioni degne di nota, come quella relativa all'attendibilità dell'interpretazione bontadiniana dell'attualismo, e presenta in appendice alcuni interessanti appunti, ad opera di Paolo Poli, volti a tracciare una breve biografia di Bontadini.

Concludendo, e riassumendo, *L'unica svolta di Bontadini*, per mole di documenti analizzati, per completezza nella ricostruzione, anche cronologica, dello snodarsi del pensiero bontadiniano, e per il confronto serrato con le altrettante ricostruzioni storiografiche esistenti, si configura, a detta dello scrivente, come il più completo e curato volume di approfondimento storiografico sul pensatore milanese in circolazione. Nonché quale ricostruzione storica più convincente quanto a puntualità di argomentazioni portate e mole documentale utilizzata, ma su questo il verdetto è lasciato al lettore, che, testo alla mano, potrà farsene un'idea in prima persona.

MARCO MENEGHIN

LIBRI RICEVUTI

A. MIRALLES - R. DÍAZ DORRONSORO, *Il matrimonio. Teologia e vita*, Seconda edizione riveduta e ampliata, Edizioni Santa Croce, Roma 2023, pp. 393, € 35,00.

Annotazioni

FILOSOFIA

- BONTADINI G., *Protologia*
- SUBACCHI M., *Che cos'è l'uomo. Maritain e Heidegger a confronto*
- MONDIN B., *Ontologia e Metafisica*, 3^a ed.
- MONDIN B., *Storia della Metafisica*, voll. 1-3, 2^a ed.
- SUBACCHI M., *Esistenza e libertà. Saggio sull'Esistenzialismo*
- TESTI C. A., *La logica di Tommaso d'Aquino. Dimostrazione, induzione e metafisica*
- MONDIN B., *Epistemologia e Cosmologia*, 2^a ed.
- BARZAGHI G., *Diario di metafisica*, 2^a ed.
- MONDIN B., *Etica e Politica*, 2^a ed.
- MONDIN B., *La metafisica di San Tommaso d'Aquino e i suoi interpreti*, 2^a ed.
- MONDIN B., *Il problema di Dio*, 2^a ed.
- RUFFINENGO P. P., *Ontonòesis. Introduzione alla metafisica per un amico pasticcere*
- MANZI A., *La paura dell'uomo contemporaneo*
- GORIUP L., *Il rischio è bello*
- MAZZANTI A. M. (ed.), *Verità e mistero*
- VANNI ROVIGHI S., *Filosofia della conoscenza*, 2^a ed.
- BERTUZZI G. (ed.), *L'origine dell'Ordine dei Predicatori e l'Università di Bologna*
- SALVIOLI M., *Il Tempo e le Parole*
- CARPI O. L., *Il problema del rapporto fra virtù e felicità nella filosofia morale di Immanuel Kant*
- LOBATO A., *La dignità della persona umana. Privilegio e conquista*
- AA. VV., *Dalla Prima alla Seconda Scolastica*
- PIAZZA G., *Il nome di Dio. Una storia della prova ontologica*
- EMILIANI A., *Dio è la mia speranza*
- EMILIANI A., *Una nuova via alla ricerca di Dio*
- PIETROSANTI R., *L'anima umana nei testi di San Tommaso*
- AA. VV., *Cristianesimo nella postmodernità e paideia cristiana della libertà*
- BOCHENSKI J., *Nove lezioni di logica simbolica*, 3^a ed.
- BASTI G., *Filosofia dell'uomo*, 3^a ed.
- EMILIANI A., *Ascesa spirituale a Dio*
- SIMON B. M., *Esiste una «intuizione» dell'essere?*
- TOMMASO D'AQUINO, *L'essere e la partecipazione. Commento al libro di Boezio «De Ebdomadibus»*

- MANFERDINI T., *Comunicazione ed estetica in Sant'Agostino*
 AA. VV., *La nuova evangelizzazione e il personalismo cristiano*
 MANFERDINI T., *Essere e verità in Rosmini*
 ROSSIGNOTTI M., *Persona e tempo in Berdjaev*
 FIORENTINO E., *Guida alla tesi di laurea* (esaurito)
 EMILIANI A., *Da gli enti finiti al superente infinito e personale che conosce e ama*
 LORENZ D., *I fondamenti dell'ontologia tomista*
 STRUMIA A., *Introduzione alla filosofia della scienza* (esaurito)
 BASTI G., *Il rapporto mente-corpo nella filosofia della scienza* (esaurito)
 AA. VV., *Etica dell'atto medico*
 BERTUZZI G., *La verità in Martin Heidegger*
 LORENZINI M., *L'uomo in quanto persona*
 AA. VV., *Coscienza morale e responsabilità politica*
 AA. VV., *Crisi e risveglio della coscienza morale del nostro tempo*
 AA. VV., *Homo loquens* (esaurito)
 TOMMASO D'AQUINO, *Pagine di filosofia*, 2^a ed.

Tutti i nostri libri e le altre attività
delle Edizioni Studio Domenicano possono essere consultati su:
www.edizionistudiodomenicano.it

Edizioni Studio Domenicano
Via dell'Osservanza 72 - 40136 Bologna - ITALIA
Tel. +39 051582034
acquisti@esd-domenicani.it
www.edizionistudiodomenicano.it

Finito di stampare nel mese di ottobre 2023
presso Venti Media Print, Monteriggioni, Siena

Artificial intelligence is developing rapidly and acquiring increasingly sophisticated cognitive abilities. The aim of this special issue of Divus Thomas is to seek answers to two interrelated questions: i) what implications might the existence of AI have for our philosophical understanding of humans and their capacities, ii) can philosophical anthropology provide us with clues for interpreting the emergent capacities of machines and for predicting future developments in AI and related fields?

L'intelligenza artificiale si sta sviluppando rapidamente e sta acquistando capacità cognitive sempre più sofisticate. L'obiettivo di questo numero speciale di Divus Thomas è cercare risposte a due domande interconnesse: 1) quali implicazioni può avere l'esistenza dell'IA per la nostra comprensione filosofica degli esseri umani e delle loro capacità? 2) Può l'antropologia filosofica fornirci indizi per interpretare le capacità emergenti delle macchine e per prevedere gli sviluppi futuri dell'IA e dei campi correlati?

DIVUS THOMAS

Rivista quadrimestrale
Via dell'Osservanza, 72
40136 Bologna
tel. +39-051-582034
www.esd-domenicani.it

Spediz. in abb. postale
D.L. 353/2003 (conv. in
L. 27/02/2004 n. 46 art. 1
c. 2) e aut. CMP Bologna
n. 070094 del 27/11/2007

€ 30,00

ISBN 978-88-5545-072-0



9 788855 450720