

Supporting Sustainable Decision-Making with Value-Sensitive Artificial Intelligence

Doctoral Thesis

Author(s):

Asikis, Thomas

Publication date:

2022

Permanent link:

<https://doi.org/10.3929/ethz-b-000532633>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

SUPPORTING SUSTAINABLE DECISION-MAKING
WITH VALUE-SENSITIVE ARTIFICIAL INTELLIGENCE

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

THOMAS ASIKIS

Master of Science in Information Systems,
Athens University of Economics and Business,
born on 11 August 1991,
citizen of Greece

accepted on the recommendation of
Prof. Dr. Dr. h. c. Dirk Helbing, examiner
Prof. Dr. Petros Koumoutsakos, co-examiner
Prof. Dr. Avishek Anand, co-examiner

Abstract

Sustainability is a term that is becoming increasingly prevalent, as several recent catastrophic events are often attributed to the impact that modern lifestyle has on the environment. Moreover, the term extends also to sustainability of operational democratic societies, where availability and accessibility to several critical infrastructures is ensured for individuals. Yet, achieving sustainability can be challenging. Recent reports from the United Nations conclude that improving decision-making processes on several levels, from institutional level policy making to individual level everyday decisions, supports sustainable development. Typically, deciding on more “sustainable” solutions often leads to complex multi-objective optimization problems, which are not trivial to solve. Modern artificial intelligence (AI) provides several methods to handle complex problems, particularly within the fields of machine learning and optimization. Nevertheless, AI has also given rise to new challenges, especially related to privacy, autonomy, and personal values and morals. Thus, the need for value-sensitive AI systems arises, where values and preferences are included directly in the system design.

The current thesis pursues the paradigm of efficient value-sensitive AI, mainly focusing on sustainable decision-making problems, by providing experimental, empirical, and methodological arguments. Three main design approaches are considered: centralized, decentralized, and a hybrid combination of both. First, an AI system applies centralized controls to simulations of critical infrastructure components, such as power grids. The results show the ability of said AI controls to stabilize and sustain the operation of critical infrastructures. Yet, centralized approaches may restrict and suppress personal freedoms and autonomy, especially when applied on individuals. Thus, a decentralized approach is evaluated next in the domain of sustainable product recommendations. The proposed AI system receives explicit input from individuals regarding their morals and values, and then calculates personalized ratings for sustainable products. Interactions between the decentralized system and individuals are evaluated in a field-study on two grocery stores. Statistically significant results confirm the system’s ability to support individuals towards more sustainable purchases.

Finally, hybrid combinations of centralized and decentralized approaches are evaluated. A novel privacy-preserving framework is proposed to calculate accurate aggregations of individual data without exposing the actual individual data to centralized systems. Additionally, a novel hybrid AI system is introduced and combined with the privacy-preserving framework to generate sustainable basket recommendations based on personal values and environmental objectives. Quantitative results on a synthetic dataset show a considerable reduction of environmental impact, even when users adopt only a fraction of the recommendations.

Zusammenfassung

Nachhaltigkeit ist ein Begriff, der zunehmend an Bedeutung gewinnt. Dies aufgrund immer häufiger vorkommenden, katastrophalen Umweltereignissen, welche auf die Auswirkungen des modernen Lebensstils zurückzuführen sind. Darüber hinaus erstreckt sich der Begriff auch auf die Nachhaltigkeit funktionierender demokratischer Gesellschaften, in denen die Verfügbarkeit und Zugänglichkeit verschiedener kritischer Infrastrukturen für jeden Einzelnen gewährleistet ist. Jedoch ist das Erreichen von Nachhaltigkeit eine Herausforderung. Aktuellste Berichte der Vereinten Nationen kommen zum Schluss, dass eine Verbesserung der Entscheidungsprozesse auf verschiedenen Ebenen, von der institutionellen Ebene der Politik bis hin zu den alltäglichen Entscheidungen eines jeden Einzelnen, eine nachhaltige Entwicklung unterstützen könnte. Typischerweise führen "nachhaltigere" Lösungen jedoch oft zu komplexen, multi-dimensionalen Optimierungsproblemen, die nicht mit einfachen Mitteln zu lösen sind. Moderne künstliche Intelligenz (KI) bietet verschiedene Methoden zur Lösung komplexer Probleme, insbesondere in den Bereichen maschinelles Lernen und Optimierung. Dennoch hat KI auch einige Herausforderungen mit sich gebracht, insbesondere in Bezug auf den Schutz der Privatsphäre, der Gewährleistung von Autonomie und der Achtung persönlicher Werte und Moralvorstellungen. Deshalb steigt der Bedarf nach wertsensitiven KI-Systemen, bei denen Werte und Präferenzen direkt in das Systemdesign miteinbezogen werden.

Die vorliegende Dissertation verfolgt dieses Paradigma einer effizienten und wertsensitiven KI. Mit Fokus auf Entscheidungsfindungsproblemen im Bereich der Nachhaltigkeit, werden experimentelle, empirische und methodologische Argumente, Resultate und Lösungen aufgezeigt. Es werden drei Designansätze betrachtet: Zentralisiert, dezentralisiert und eine hybride Kombination aus beiden Ansätzen. Zunächst wendet ein zentralisiertes KI-System Kontrollen in Simulationen kritischer Infrastrukturkomponenten an, wie z.B. Stromnetzen. Die Ergebnisse zeigen, dass diese KI-Kontrollen in der Lage sind, den Betrieb kritischer Infrastrukturen zu stabilisieren und aufrechtzuerhalten. Zentralisierte Ansätze können jedoch persönliche Freiheiten und Autonomie einschränken und gar unterdrücken, insbesondere wenn sie auf Einzelpersonen angewendet werden. Daher wird als nächstes ein dezentralisierter Ansatz im Bereich der nachhaltigen Produktempfehlungen evaluiert. Das vorgeschlagene dezentralisierte KI-System erhält von Einzelpersonen Informationen zu ihren Moralvorstellungen und Werten und berechnet darauf basierend personalisierte Empfehlungen nachhaltiger Produkte. Die Interaktionen zwischen dem dezentralisierten System und Einzelpersonen werden in einer Feldstudie in zwei Lebensmittelgeschäften untersucht. Statistisch signifikante Ergebnisse bestätigen die Fähigkeit des dezentralisierten KI-Systems Individuen zu nachhaltigeren Einkäufen zu bewegen.

Schliesslich werden hybride Kombinationen von zentralisierten und dezentralisierten Ansätzen evaluiert. Ein neuartiges Framework zur Wahrung der Privatsphäre wird vorgeschlagen, um genaue Aggregationen individueller Daten zu berechnen ohne die tatsächlichen individualbasierten Daten an zentralisierte Systeme weiterzugeben. Darüber hinaus wird ein neuartiges hybrides KI-System eingeführt und mit dem datenschutzfreundlichen Framework kombiniert, um nachhaltige Warenkorbeempfehlungen auf der Grundlage persönlicher Werte und ökologischer Ziele zu generieren. Quantitative Ergebnisse aus einem synthetischen Datensatz zeigen eine beträchtliche Verringerung der Umweltbelastung, selbst wenn die User nur einen Bruchteil der Empfehlungen annehmen.