FUTURE COMPUTING 2025

The Seventeenth International Conference on Future Computational Technologies and Applications

ISBN: 978-1-68558-259-3

April 6 - 10, 2025

Valencia, Spain

FUTURE COMPUTING 2025 Editors

Hans-Werner Sehring, NORDAKADEMIE Hochschule der Wirtschaft, Germany

FUTURE COMPUTING 2025

Forward

The Seventeenth International Conference on Future Computational Technologies and Applications (FUTURE COMPUTING 2025), held on April 6 – 10, 2025, continued a series of events targeting advanced computational paradigms and their applications. The target was to cover (i) the advanced research on computational techniques that apply the newest human-like decisions, and (ii) applications on various domains. The new development led to special computational facets on mechanism-oriented computing, large-scale computing and technology-oriented computing. They are largely expected to play an important role in cloud systems, on-demand services, autonomic systems, and pervasive applications and services.

Similar to the previous edition, this event attracted excellent contributions and active participation from all over the world. We were very pleased to receive top quality contributions.

We take here the opportunity to warmly thank all the members of the FUTURE COMPUTING 2025 technical program committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to FUTURE COMPUTING 2025.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the FUTURE COMPUTING 2025 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope FUTURE COMPUTING 2025 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of future computational technologies and applications. We also hope that Valencia provided a pleasant environment during the conference and everyone saved some time to enjoy this beautiful city.

FUTURE COMPUTING 2025 Steering Committee

Hiroyuki Sato, The University of Tokyo, Japan
Sergio Ilarri, University of Zaragoza, Spain
Jay Lofstead, Sandia National Laboratories, USA

FUTURE COMPUTING 2025 Publicity Chair

Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
José Miguel Jiménez, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politècnica de València, Spain

FUTURE COMPUTING 2025

Committee

FUTURE COMPUTING 2025 Steering Committee

Hiroyuki Sato, The University of Tokyo, Japan
Sergio Ilarri, University of Zaragoza, Spain
Jay Lofstead, Sandia National Laboratories, USA

FUTURE COMPUTING 2025 Publicity Chair

Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
José Miguel Jiménez, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politècnica de València, Spain

FUTURE COMPUTING 2025 Technical Program Committee

Andrew Adamatzky, University of the West of England, Bristol, UK
Paramasiven Appavoo, University of Mauritius, Mauritius
Ehsan Atoofian, Lakehead University, Canada
Yadu Babuji, University of Chicago, USA
Bernhard Bandow, GWDG, Göttingen, Germany
Kaustav Basu, The Laboratory for Networked Existence (NetXT), USA
Christos J. Bouras, University of Patras, Greece
Massimiliano Caramia, University of Rome "Tor Vergata", Italy
Steve Chan, Decision Engineering Analysis Laboratory - Virginia Tech, USA
Ryan Chard, Argonne National Laboratory, USA
Nan-Yow Chen, National Center for High-Performance Computing (NCHC), Taiwan
Sunil Choenni, Ministry of Justice and Security / Rotterdam University of Applied Sciences, Netherlands
Fabio D'Andreagiovanni, CNRS & UTC - Sorbonne, France
Leandro Dias da Silva, Universidade Federal de Alagoas, Brazil
Chantal Fuhrer, Université de la Réunion, France
Hachiro Fujita, Tokyo Metropolitan University, Japan
Félix J. García Clemente, University of Murcia, Spain
Apostolos Gkamas, University of Ioannina, Greece
Victor Govindaswamy, Concordia University - Chicago, USA
Wei-Chiang Hong, Asia Eastern University of Science and Technology, Taiwan
Sergio Ilarri, University of Zaragoza, Spain
Yuji Iwahori, Chubu University, Japan
Yasushi Kambayashi, Sanyo-Onoda City University, Japan
Mehdi Kargar, Ted Rogers School of Management - Ryerson University, Toronto, Canada
Rethabile Khutlang, Council for Scientific & Industrial Research, South Africa

Michihiro Koibuchi, National Institute of Informatics, Japan
Zbigniew Kokosinski, Cracow University of Technology, Poland
Carlos León-de-Mora, Universidad de Sevilla, Spain
Christoph Lipps, German Research Center for Artificial Intelligence, Germany
Jay Lofstead, Sandia National Laboratories, USA
Giuseppe Mangioni, DIEEI - University of Catania, Italy
Wail Mardini, Jordan University of Science and Technology, Jordan
Anuja Meetoo-Appavoo, University of Mauritius, Mauritius
Yassine Mekdad, Florida International University, USA
Isabel Muench, German Federal Office for Information Security, Germany
Anand Nayyar, Duy Tan University, Vietnam
Sithembile Nkosi, Mangosuthu University of Technology / University of KwaZulu Natal, South Africa
Kendall E. Nygard, North Dakota State University - Fargo, USA
Carla Osthoff, National Laboratory for Scientific Computing, Brazil
Fred Petry, Naval Research Laboratory, USA
Wajid Rafique, Nanjing University, China
Eric Renault, Télécom SudParis | Institut Polytechnique de Paris, France
Hiroyuki Sato, The University of Tokyo, Japan
Andrew Schumann, University of Information Technology and Management in Rzeszow, Poland
Friedhelm Schwenker, Ulm University, Germany
Xiaojun Shang, University of Texas at Arlington, USA
Massimo Torquati, University of Pisa, Italy
Carlos M. Travieso-González, University of Las Palmas de Gran Canaria, Spain
Ndivhoniswani Aaron Tshidzumba, University of South Africa & North-West University, South Africa
Eirini Eleni Tsiropoulou, Arizona State University, USA
Teng Wang, Oracle, USA
Alex Wijesinha, Towson University, USA
Hongji Yang, Leicester University, UK
Aleš Zamuda, University of Maribor, Slovenia
Claudio Zandron, University of Milano-Bicocca, Milan, Italy
Albert Zomaya, University of Sydney, Australia

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Treatment of the Multi-Attribute Decision-Making Rank Reversal Problem for Real-World Systems

Steve Chan
VTIRL, VT/DE-STEA
Orlando, USA
Email: stevec@de-stea.tech

***Abstract*—This paper describes an enhanced approach towards considering the Rank Reversal (RR) problem for certain Multi-Attribute Decision-Making (MADM) methods critical to Multi-Criteria Decision-Making (MCDM) systems. Prototypical testing environments for RR usually do not include key facets of Real-World Systems (RWS), such as the treatment of time, prospective Influence Dominating Sets (IDS) at play, sub-biases throughout the system, involved Decision Engineering Pathways (DEP) for consortial environments, and a more Transparent, Explainable, and Accountable (TEA)-oriented architectural construct, which are all desired in these contemporary times. These facets have been considered as Extrapolated Decision Quality (DQ) Thematics (EDQTs) of the Howard & Abbas six classically understood facets of DQ, and they are critical for MCDM RWS. Since various MADM methods vary in performance against the EDQTs, the approach utilized is to employ a robust Multi-Objective Decision-Making (MODM) module to discern the more optimal MADM methods to utilize in an ongoing fashion.**

Keywords-decision engineering pathway; decision-making; multi-criteria decision-making; multi-attribute decision-making; rank reversal; multi-objective decision-making; decision quality; artificial intelligence; machine learning; epistemic transparency.

I. INTRODUCTION

The issue of bias in Artificial Intelligence (AI) systems has been a prevalent topic. Major companies, such as in the 2019 to 2020 time frame, had withdrawn a number of AI/Machine Learning (ML) offerings from the marketplace due to the fact that mitigation against prospective biases (e.g., gender, racial/ethnic, etc.) had not been robustly considered in the design of those systems. Since that time, algorithmic bias has become an acknowledged issue, and the notion of equitable outcomes (as contrasted to "unfair" or "privileged" outcomes) has become an important aspect in the design of AI/ML-centric Tools, Platforms, Methodologies, Frameworks, and Systems (TPMFS). Confalonieri notes that while the Explainability in AI (XAI) movement has resurged in recent times, its origins trace back a number of decades via various research Lines of Effort (LOEs), such as "expert systems," "recommender systems," "neural-symbolic learning and reasoning," etc. [1]. Heder notes that Winograd had investigated the "issues of explanations and transparency" (critical to XAI) via LOEs, such as "phenomenology" and "cognitive science," and Hosain underscores Winograd's contributions [2][3]. Heder also investigated the notions of "epistemic opacity" (i.e., wherein functional details may not be clear, such as in a "black box" architectural construct) and the criticality of moving towards "epistemic transparency" [4]. The IEEE Standards Association has also opined the need for Autonomous and Intelligent Systems (AIS) to be comprehensible so as to be accountable, and standards, such as IEEE P7001 [Standard for Transparency of Autonomous Systems] (one of the P70XX series of standards), have emerged, received approval (e.g., 2021), and published (e.g., 2022) so as to put forth a delineation of Transparency/Explainability (T/E); for example, P7001 has a T/E scale of 0 (no T/E) to 5 (fullest attainable extent of T/E) [5]. While P7001 seems to have gained some traction in areas, such as robotics, advances in the area of XAI are still nascent/ongoing [6]. Winfield points out that P7001 is a process standard, wherein the involved T/E measures are not specified, and Winfield further asserts that the principal role of P7001 is to serve as a System Transparency Specification (STS) and as a System Transparency Assessment (STA) [7].

Beyond STA and the issue of transparency, the Association of Computing Machinery (ACM) accentuates *explanation* in its "Principles for Algorithmic Transparency and Accountability" [8]. Also, "the European Union's [General] Data Protection Regulation (GDPR) stipulates a right" "for consumers affected by an automatic decision" "to obtain 'meaningful information about the logic involved'" [1]; Confalonieri notes that this equates to a "right to *explanation*" [1]. Along this vein, Winfield notes that "P7001 recognises that AI technology cannot be separated from the larger Socio-Technical System [STS] of which it is a component" [7]. STS encompasses the interplay among *humans*, technology, and the environs, and while the overarching XAI and P7001-type movements further mature and burgeon, it is interesting to note that for some ecosystems, there has been a predilection for increasing the utilization of *humans*-in-the-loop for Decision Engineering (DE)/Decision-Making (DM) (to mitigate against "non-perfect" algorithmic and AI/ML-centric paradigms), particularly for "high-stakes tasks" [9]. The arena encompassing this Human-Computer Interaction (HCI)-centric DE/DM begets a new set of challenges, such as in the case of Multi-Criteria Decision-Making (MCDM) for AI Technology-Related Investment Decisions (TRID). This might beget the use of *human* evaluators, who in a number of cases, such as within the reviewer ecosystem, self-assess their own level of expertise in a Subjective Measure/Methodology (SM) fashion. Yet, the level of

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. **ISBN: 978-1-68558-259-3**

expertise should be context dependent; for example, various reviewers may rate their "AI hardware expertise" at the same level — such as when reviewing an AI whitepaper involving massive datasets, intricate Deep Learning (DL) (as contrasted to the less intricate methods of ML), accelerated computational performance, and energy efficiency — but in actuality, those with Tensor Processing Unit (TPU) and Graphics Processing Unit (GPU) proficiency may be better suited than those with simply Central Processing Unit (CPU) experience. After all, it is now generally understood that GPUs may offer better performance speeds for DL models with large datasets over CPUs (as the size of the involved dataset increases, CPU performance may decrease due to its constrained parallel processing capabilities) and for large-scale computation, TPUs may offer accelerated performance (as well as better energy efficiency "without jeopardizing the model's accuracy") over GPUs and CPUs [10][11]. Likewise, "technical expertise regarding AI" may also vary depending upon time frame and macro trends, such as those which can be gleaned from the U.S. Patent Trademark Office (USPTO) AI Patent Dataset (AIPD) and PatentsView Data, World Intellectual Property Organization (WIPO), etc. In many cases, this information is not being robustly considered for TRID-related reviewer assessments; indeed, the realm of assessments is heavily beset with SM, which are infrequently counterbalanced with Objective Measures/Methodology (OM) approaches.

A well-counterpoised Dynamic Assessment and Weighting System (DAWS) can be utilized to derive more appropriate weights, such as when considering the SM-centric self-assessment of the reviewers and OM-centric macro trend utilization. For example, during the time period 2000-2020, according to the USPTO AIPD and PatentsView data, the AI component technologies with the highest number of patents (with a government interest) were, in descending order, "Knowledge Processing (KP), [Computer] Vision (CV), Planning & Control (P&C), AI Hardware (AIH), ML, Natural Language Processing (NLP), and Evolutionary Computation (EC)" [12][13]; of course, the order changes depending upon the time frame chosen (e.g., 2012-2016 might differ from 2016-2020). In addition, there was a "2023 update to the AIPD" that incorporates various refinements (e.g. BERT for Patents) and overcomes prior limitations that might affect the sorting order [14]. The relative ranking of KP, CV, P&C, AIH, ML, NLP, and EC, among others, is likely to be significant for the review of a TRID, particularly if there is an accompanying supposition/reliance upon future governmental funding [12]. The DAWS, which is also referred to by various other terms of art, such as Adaptive Weighting Schema (AWS), Adaptive Weighting Methodology (AWM), Adaptive Assessment & Weighting Methodology (A2WM), Adaptive Criteria Weighting System (ACWS), etc., endeavors to overcome the SM biases with OM input. Moreover, the DAWS construct is also envisioned to have an enhanced T/E posture. To address the research goal and problem statement of achieving not only a more robust T/E, but also a DAWS that demonstrates more responsibility (more aspirational at this point), the paper delineates an innovative approach towards devising a construct with more epistemic Transparency, Explainability, and Accountability (TEA). The aspects discussed within this paper are presented in Table I (with utilized acronyms).

TABLE I. TABLE OF ACRONYMS

Acronym	***Full Form***
A&F	Aires & Ferreira
A2WM	Adaptive Assessment & Weighting Methodology
ACM	Association of Computing Machinery
ACWS	Adaptive Criteria Weighting System
AI	Artificial Intelligence
AIH	AI Hardware
AIPD	AI Patent Dataset
AIS	Autonomous and Intelligent System
AWM	Adaptive Weighting Methodology
AWS	Adaptive Weighting Schema
C&L	Cascales & Lamata
C&W	Choo & Wedley
CPU	Central Processing Unit
CV	Computer Vision
DAWS	Dynamic Assessment and Weighting System
DE	Decision Engineering
DEP	Decision Engineering Pathway
DL	Deep Learning
DM	Decision-Making
EC	Evolutionary Computation
EDQ	Extrapolated Decision Quality
EDQT	EDQ Thematic
F&H	Finan & Hurley
GDPR	General Data Protection Regulation
GPU	Graphics Processing Unit
HCI	Human-Computer Interaction
IDS	Influence Dominating Set
K&U	Kwiesielewicz & Uden
KP	Knowledge Processing
L&N	Liberatore & Nydick
LOE	Line of Effort
MADM	Multi-Attribute Decision-Making
MCDC	Multi-Criteria Decision-Making
ML	Machine Learning
MODM	Multi-Objective Decision-Making
MVP	Minimum Viable Product
NLP	Natural Language Processing
OM	Objective Measure/Methodology
P&C	Planning & Control
RR	Rank Reversal
RWS	Real-World System
S&V	Saaty & Vargas
SM	Subjective Measure/Methodology
SOTA	State-of-the-Art
STA	System Transparency Assessment
STS	System Transparency Specification
STS	Socio-Technical System
T/E	Transparency/Explainability
TEA	Transparency, Explainability, and Accountability
TPMFS	Tools, Platforms, Methodologies, Frameworks, and Systems
TPU	Tensor Processing Unit
TRID	Technology-Related Investment Decision
USPTO	U.S. Patent Trademark Office
W&W	Wijnmalen & Wedley
WIPO	World Intellectual Property Organization
XAI	Explainability in AI

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. **ISBN: 978-1-68558-259-3**

Section I presented the narrative arc, which explains the title of the paper. Section II provides pertinent background information. Section III provides aspects of the theoretical foundations, which underpin the paper, as well as delineates some of the precursor research LOEs leading up to this point. Section IV presents an experimental construct. Section V summarizes with some reflections and puts forth future work.

II. Background

Schmidt notes that current funding schemas (e.g., seed capital) may no longer suffice since the "next generation of technologies" (e.g., AI) will "increasingly require sustained and substantial amounts of resources to reach commercial scale" [15]; this alludes to the paradigm, wherein AI TRID might carry higher thresholds of risk/reward. To address this, Boucher and others have underscored the use of MCDM "in the evaluation of technology investment decisions" [16]. In addition, Triantaphyllou notes that "pertinent data are very expensive to collect," so a robust utilization/evaluation of this data, such as via MCDM, seems prudent [16].

A. MCDM

Fattoruso (as well as Rao, Sitorus, and of course, Hwan & Yoon) construe MCDM as being comprised of Multi-Attribute Decision-Making (MADM) and Multi-Objective Decision-Making (MODM) [17]. MADM involves "discrete decision spaces" (i.e., the number of alternatives is "finite and predetermined") [18]; in contrast, for MODM, "the decision space is continuous" (i.e., "the number of alternatives is infinite" and undetermined) [19]. Restated, MODM tends to contend with multiple objectives (often conflicting) and seeks to ascertain an optimal solution set among "undetermined continuous alternatives" while MADM tends to contend with a single objective and sorts/ranks so as to determine the optimal solution among "a finite set of discrete alternatives" [20]. MADM and MODM each have SMs and OMs that can be leveraged. Ideally, the OMs can somewhat mitigate against the SMs, and three distinct scenarios are presented, wherein this counterpoising would be invaluable.

1) Scenario #1

In a number of cases, reviewer evaluations (at the "same level of expertise") may be diametrically opposed. The choice of OM is non-trivial, as conventional generalized measures, such as h-index or i-index may be specious in deciding how to re-weight the reviewer's self-assessment [21]. In the case of an AI technology firm (e.g., whose intended market is, say, Japan and/or Germany) seeking funding for the advancement of the AI technique of, say, fuzzy logic, the reviewer with the stronger background in fuzzy logic might be of higher criticality and weighted more, as the need to determine the competitive barrier to entry in the involved countries is significant, particularly as the WIPO indicates that the referenced countries have notable strengths in the area of fuzzy logic [12][13][14][22]. Hence, the reviewer's expertise level varies by the involved locale, as what constitutes State-of-the-Art (SOTA) may vary geographically.

2) Scenario #2

As noted by various repositories on GitHub, startups and lean engineering teams seeking to develop the Robinson-Blank-Ries notion of Minimum Viable Products (MVP) might use various packages from Github for more Rapid Application Development (RAD) [23]. However, in some cases, technical issues for the package may abound (e.g., "signature consistency and dependency intricacies have been shown to result in errors and/or incorrect results") and may constitute "glass ceilings" (until resolved) [24]. In this case, the reviewer with the higher proficiency in numerical methods and experience with various libraries, toolkits, and frameworks (e.g., PyTorch, Tensorflow, etc.) might be of higher criticality (e.g., for having previously contended with incompatibility issues, conflicts with required libraries, as well as an assortment of "glass ceiling" matters) and, likely, should be re-weighted accordingly [25].

3) Scenario #3

In a number of cases, professional investors endeavor to mitigate against bias so as to enhance investment discipline and achieve a better Return on Investment (ROI). The use of Behavioral/Emotional Analytics (BEA) within this ecosystem has been increasing, and there have been some explorations with using BEA Multimedia (MM) feeds for re-weighting the self-assessment of reviewers involved with TRID. Differing from the predominantly volunteer reviewers within the academic community, reviewers for TRID tend to be paid professionals, and accordingly, they are more amenable to the stipulations of the investment firms, who engage their services. MM-based BEA has improved since the 1990s with enhanced resolution and more robust time series analytical tools to discern, among other measures, Duchenne indicators — "lip corner puller action unit (AU12)," "cheek raiser action unit (AU6)," lip corners pulled "towards the ears" (AU12), etc. — so as to, potentially, posit how fervently/sincerely the reviewer subscribes to his/her own self-evaluation of expertise on a topic [26]. The use of Duchenne indicators seems to supported by the increasing use of the "Automatic Facial Expression Analysis (AFEA), which automates the Facial Action Coding System (FACS)," and is noted by Clark and others as being "the most comprehensive, psychometrically rigorous, and widely used system to describe facial activity in terms of visually observable facial muscle actions (i.e., [Action Units or] AUs)" [27]. As TEA accountability can lead to more "trustworthy" TPMFS, it should be of no surprise that the use of Duchenne (e.g., "genuine") and non-Duchenne ("non-genuine)" indicators (e.g., smiles) have been of great interest as a prospective OM-centric MM feed [28].

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

B. Effective MODM & MADM SMs/OMs

Lyons-Padilla notes that “asset allocators manage more than $69.1 trillion dollars globally on behalf of governments, universities, charities, foundations, and companies” and retain “professional managers to generate returns” (i.e., ROIs) [29]. Despite the anticipated investment discipline, particularly given the magnitude of funds at stake, Lyons-Padilla and others have reported that professional investor human review teams remain subject to bias in their financial decisions [29]. Along this vein, a TRID human review team may be beset by a variety of predilections. For example, the teams may have been assembled using a variety of 360 evaluation, personality type, and conflict mode/management assessments that are predominantly SM-based (and, thereby, subject to inherent biases). In many cases, these assessment tools were matured/utilized, such as in the 1950s, 1950s/60s, 1970s, respectively, although the developmental origins tend to trace back to the 1930s and 1940s (particularly during the World War II time frame) [30][31][32]. As this was prior to the more prevalent use of AI/ML (since the 1990s), the counterpoising of SM with OM-based approaches remains a relatively unsaturated/nascent area. However, the arena of MCDM endeavors has leveraged both SM and OM so as to formulate a more practical/logical weighting, such as noted by Taherdoost (as well as Hwang & Yoon and others) [33]. Prior experimentation has shown that particular combinations of MADM/MODM SMs/OMs can achieve a modicum of efficacy; exemplars are shown in Table II.

TABLE II. EXEMPLAR MADM/MODM SMS/OMS

#	TPMFS	MADM/ MODM	SM/ OM
1	Analytic Hierarchy Process (AHP)	MADM [34]	SM [35]
2	Weighted Aggregated Sum Product Assessment (WASPAS)	MADM [36]	SM [37]
3	CRiteria Importance through Intercriteria Correlation (CRITIC)	MADM [38]	OM [39]
4	Data Envelopment Analysis (DEA)	MADM [38]	OM [38]
5	Technique of Order Preference by Similarity to an Ideal Solution (TOPSIS)	MADM [40]	OM [41]
6	Fuzzy VIseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR)	MADM [42]	SM/OM [43]
7	Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) (e.g., I and II)	MADM [44]	SM/OM [45][46] [47][48] [49]
8	ELimination Et Choix Traduisant la Realité (ELECTRE)	MADM [50]	SM/OM [47]
9	Multi-Objective Optimization by a Ratio Analysis plus the Full Multiplicative Form (MULTIMOORA)	MODM [51]	SM [51]
10	Goal Programming (GP) Method	MODM [52]	OM [52]

Yet, even for the case of a well-counterpoised construct, the matter of TEA is a separate matter, and architectural constructs, from previous experimentation, are often not evaluated for TEA. This segues to the need for an experimental TEA construct, which is described in Section IV. Some of the theoretical foundations are delineated in Section III below.

III. THEORETICAL FOUNDATIONS FOR THE EXPERIMENT

Abbas and Howard had noted that there are, fundamentally, “six elements of Decision Quality” (DQ) (although Abbas later expands this to eleven elements) [53]. These include: (1) an understanding of the involved “uncertainty,” (2) a grasp of the problem boundaries (e.g., including the temporal constraints of (1)) and the “perspectives involved,” (3) identification of the reasoning involved (e.g., “values,” “trade-offs,” prioritization schemas, etc.), (4) the “commitment to action” by the Decision Maker (DM) “and the stakeholders…affected by the decision,” (5) the determination of “feasible” alternatives, and (6) the “choice criterion” to “choose the alternative with the highest expected utility” (e.g., use of the Neumann-Morgenstern utility function) [53][54]. Along this vein, various DQ dimensions have been explored, as is shown in Table II, by way of extrapolated LOEs/EDQTs. For example, (1) has been extended to the notion of ambiguity/uncertainty, (2) has been extended to more fully contextualize the “perspectives involved,” via Spatial-Temporal Knowledge Graph (STKG) Completion (STKGC)/STKG Reasoning (STKGR), (3) has been extended to contextualize the involved rationale, via DAWS (e.g., AWS/AWM/A2WM/ACWS), (4) has been extended to better comprehend the potential DE Pathways (DEP) and the accompanying operationalization schemas (e.g., Command and Control or C2) by the DM and/or the notion of Multi-Partner Enclaves (MPEs) or “coalitions of the willing,” (5) has been extended to better organize/sort/rank the prospective alternatives, via a Counterpoised MCDM (C-MCDM) (e.g., a balancing of MADM/MODM SM/OM), and (6) has been extended to consider the most apropos MADM (given RR considerations along with the considerations of (1) through (5)). The EDQTs are clarified in the following subsections A through F.

A. LHM, an extrapolation of DQ#1

The notion of “uncertainty” should not be treated in isolation, particularly when there is a temporal element. Time can be classified as Compressed Decision Cycles (CDC) (i.e., a “paradigm of ‘tight time constraints’”) and Uncompressed Decision Cycles (UDC) (i.e., a paradigm, wherein time is not necessarily of the essence) [56]. In a situation of CDC, the DM may tolerate “higher uncertainty (i.e., sparse data) given the condition of lower ambiguity” (i.e., a similar situation has happened before, so there is some prior experience on how to react) [56]. This paradigm of Lower Ambiguity and Higher Uncertainty is referred to as LAHU. In contrast, in a situation of UDC, the DM may

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

not "simply accept the higher uncertainty" and might "proactively seek to use 'more data to lower uncertainty'," particularly given the condition of higher ambiguity (i.e., a comparable situation has not been encountered before, so there is no apriori experience of how to react) [56]. This paradigm of Higher Ambiguity and Lower Uncertainty is referred to as HALU. When conjoined, a LAHU HALU Module is referred to as an LHM [55][56].

B. *Higher-Order Networks (HON), an extrapolation of DQ#2*

Tian asserts that a Knowledge Graph (KG) "describes the objective world's concepts, entities, and their relationships in the form of graphs" [57]. The procedure of positing links and nodes is known as KG Completion (KGC). Building upon this, Chen and Ji assert that KG Reasoning (KGR) can "discover new knowledge from existing knowledge" [58][59]. However, in its base form, KGs are static, as they lack temporal information [60]. In turn, Temporal KGs (TKs) are critiqued for their lack of spatial information [61]. Also, Spatial-Temporal KGs (STKGs) are critiqued against the backdrop of Positive Influence Dominating Sets (PIDS) as well as Negative Influence Dominating Sets (NIDS), and the PIDS/NIDS effects are considered against the Abelian Sandpile Model (ASM) or Bak–Tang–Wiesenfeld (BTW) phenomenon of non-equilibrium systems so as to ascertain the prospective "Higher-Order Networks" (HONs) at play (i.e., other stakeholders), which is of critical import to discern. After all, without being cognizant of the potential HONS at play (as well as identifying the likely HONs at play), delineation of the boundaries and the framing of the problem will not be correct [62].

C. *DAWS, an extrapolation of DQ#3*

In a substantial number of cases, TPMFS are beset by selection bias (e.g., the choice/formulation of heuristics). In a number of these cases, even the DAWS involved are beset with confirmation bias (e.g., the choice/amalgamation of parameters). This effect is further aggravated when the utilized AI/ML is also beleaguered with inherent inclinations. A mitigation approach that has been utilized with some efficacy has been to utilize the Type-2 Fuzzy Sets (T2FS) and Spherical Fuzzy Set (SFS) versions of the TPMFS approaches of Table 1. Other enhancements include utilizing an Extended Matrix Shanks Transformation Accelerant (EMSTA).

D. *C2, an extrapolation of DQ#4*

DEPs may vary for the DM and the stakeholders of the MPE (and for the MPE itself, as it evolves or devolves); DEPs for urgent situations (i.e., "exigency circumstances") and non-urgent situations (i.e., "non-exigency circumstances) may differ greatly. This also relates to the notion that the Minimum Controllability Problem (MCP) is quite different from the Efficient Controllability Problem (ECP) (since ECP is more desirable for exerting control, when desired, over a more elongated period of time). In particular, control may need to be exercised during "exigency circumstances." DEPs may also vary depending upon the degree of resiliency incorporated into the involved system/paradigm. The ability to exercise action/operationalize, via the involved/available C2, is highly dependent upon the involved DEPs, MCP/ECP, and circumstances (e.g., exigency/non-exigency).

E. *TEA, an extrapolation of DQ#5*

Prior research had found that a cascading class of "ever smaller" convolutional filters is well-suited for DL (and the implementation of C-MCDM) since they well mimic a Convolutional Wavelet Transform (CWT) approach, which unlike other types of transforms, do not necessarily suffer as much from truncation, leakage, and other issues [63]. Hence, there is an advantage to leveraging "cascading 'CWT-like' convolutional filters" [63]. Also, bounds tightening can be employed (e.g., such as by a bespoke convex relaxations framework for the "tightest possible relaxation") so as to further delineate the successive steps being taken. This can be achieved via a Bespoke Implementation (BI), which was delineated in prior work and also lends towards operationalizing the MODM OM. Collectively, the approach lends to TEA.

F. *Rank Reversal (RR) Challenge, an extrapolation of DQ#6*

Despite the generalized promise of the MADM TPMFS of Table 1, the specific implementation is crucial. For example, in some instances, the MADMs of Table 1 can experience a "Rank Reversal" (RR) phenomenon and yield incorrect results. Belton and Gear (B&G) had first noted the RR dilemma, and recognition of the problem was affirmed by Triantaphyllou and others across the gamut of MADM approaches. Even newly introduced MADM methods are beset by the RR challenge. However, it has been reported that, among others, the Ranking of Alternatives through Functional mapping of criterion sub-intervals in a Single Interval (RAFSI) method can somewhat mitigate against the RR challenge, and it is also mathematically straightforward so as not to worsen the TEA goal [64].

Garcia-Cascales describes RR as a paradigm that manifests when a DM "is confronted with new alternatives that were not thought about" or available "when the selection process was initiated" [65]. Aires adds to this by noting that "RR refers to a change in the ordering among alternatives previously defined after the addition or removal of an alternative from the group previously ordered" and pointed out that the primary methods of MCP (e.g., "AHP, TOPSIS, ELECTRE, PROMETHEE and combinations thereof") "have been criticized due to the occurrence of" RR [66]. By way of background, RR discussions had commenced via Saaty, B&G, and Saaty and Vargas (S&V) in 1980, 1983, and 1984, respectively. The dialectic prompted others, such as Triantaphyllou, Finan & Hurley

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

(F&H), Liberatore & Nydick (L&N), Wijnmalen & Wedley (W&W), and others to engage in RR research. Simplistically, B&G argued that RR can manifest "when a new alternative is added or deleted," S&V argued that RR "can occur due to the presence of near or similar copies within the set of alternatives," Cascales & Lamata (C&L) asserted that "it is well known that when the projects are very close[,] the order between them can depend on the method used on their evaluation," Fedrizzi argued that RR "depends on the distribution of criteria weights" (i.e., entropy of the weight distribution") and that "the estimated probability of" RR "increases with the weights entropy," and Choo & Wedley (C&W), Lin, as well as others worked on "deriving the priority values from the pair-wise comparison matrix," but Kwiesielewicz & Uden (K&U) showed that the "pair-wise comparison matrix can be contradictory (inconsistent), yet it can pass the consistency check" [67][68]; this list goes on. Proposed RR mitigation methods, among others, have been put forth by Zizovic in the form of "the lattice MADM method," Kizielewicz's "Characteristic Objects method (COMET)," Dezert's Stable Preference Ordering Towards Ideal Solution Method (SPOTIS), and others [69][70][71][72]. Wieckowski points out that theoretical mitigation and practical mitigation for RWS are quite different and uses varying sensitivity analysis results to underscore the point [72]. Yet, "despite the great interest" in RR, Aires asserts that "given its importance for addressing the reliability of MCDM methods, there is still a paucity in the literature regarding this subject" [66]. This assertion was made despite the fact that Maleki & Zahir had "evaluated 61 papers…from 18 international journals," Aires & Ferreira (A&F) had evaluated "130 articles…from 37 journals," and others (e.g., Yu) [66][73][74].

A key factor for ascertaining the latent stability of MADM methods is to inject replacement alternatives into (or by removing alternatives from) the original set. Ideally, the MADM method would not exhibit any substantive change in the organizing/sorting/ranking of the alternatives. Zizovic's RAFSI constitutes a foray into better contextualizing resistance to RR; this paper endeavors to continue that foray. The research of this paper also considers the elements of: (1) time (e.g., CDC/UDC), such as in the case of LHM, (2) HON (e.g., PIDS/NIDS), such as in the case of STKGC/STKGR, (3) biases/sub-biases (e.g., chosen parameters, indices, heuristics, etc.), such as in the case of the DAWS utilized, (4) involved DEPs and the ability to exert C2, whether DM/MPE and/or MCP/ECP during varied circumstances (e.g., exigency/non-exigency), and (5) involved architectural construct (e.g., for the treatment of TEA), which needs to consider both the Method (M) and Architecture (A) involved. The prior research relating to the EDQTs atop the fundamentals of DQ#1 through 5, which segue to the novelty and contribution of this paper, is shown in Table III below.

TABLE III. EDQTS FOR THE VARIOUS DQ DIMENSIONS

DQ #	***DOI***	***EDQTs***
1	• 10.1109/GEM61861.2024.10585580 • 10.1109/IAICT62357.2024.10617473	LHM (UDC/CDC)
2	• 10.1109/AIIoT61789.2024.10579029 • 10.1109/IBDAP62940.2024.10689701	HON (PIDS/NIDS)
3	• 10.1109/CyMaEn57228.2023.10051057 • 10.1109/ICPEA56918.2023.10093212 • 10.1109/ICSGTEIS60500.2023.10424230 • 10.1109/AIIoT61789.2024.10579033	DAWS
4	• 10.1109/IEMCON.2019.8936241 • 10.1109/IAICT62357.2024.10617473	C2 (MCP/ECP)
5	• 10.1109/ICPEA56918.2023.10093212 • 10.1109/AIIoT61789.2024.10579033 • 10.1109/ICDCSW53096.2021.00014 • 10.1109/IEMCON53756.2021.9623140 • 10.1109/OETIC57156.2022.10176215	TEA (M/A)
6	This paper.	RR

For this paper, a particular Achilles heel of MCDM systems was explored and addressed. Hwang and Yoon had previously noted that the most utilized facet of MCDM was that of MADM, and Fattoruso had found that AHP was the most prevalent method utilized for MADM [17]. In addition, Fattoruso noted that methods, such as PROMETHEE and ELECTRE, were minimally used in various sectors; TOPSIS was used slightly more often, but its use still paled in comparison to AHP [17]. Despite the widespread use of AHP, ironically, Aazadfallah asserts that AHP is the most sensitive to RR, while TOPSIS, PROMETHEE II, and ELECTRE are more resistant/stable (yet still susceptible to RR as well) [75]. Other MADM methods are also sensitive to RR [76]. Even after B&G noted the AHP susceptibility to RR and the creator of AHP, Saaty, unveiled an updated version, B&G pointed out that Saaty's updated version was still susceptible under particular conditions; B&G released a version that was supposedly resistant to RR, but S&V asserted that the B&G version was susceptible as well [77]. Bottom line, AHP is still deemed to be susceptible to RR. Moving beyond the catch-all generalizations of RR, Resistance/Stability (R/S) is also subject to the RR Type (RRT), as shown in Table IV.

TABLE IV. TYPES OF RR (RRT)

RRT #	***Initial Ranking***	***Expected Ranking after change***	***Exemplar Manifested RR***
1	DEP_3,DEP_1,DEP_2	($DEP_1 \sim DEP_4$); DEP_3,DEP_4,DEP_2	DEP_2,DEP_4,DEP_3
2	DEP_3,DEP_1,DEP_2	($DEP_1 > DEP_4$); DEP_3,DEP_4,DEP_2	DEP_2,DEP_4,DEP_3
3	DEP_3,DEP_1,DEP_2	($DEP_1 \sim DEP_4$); $DEP_3 > DEP_4$ $DEP_4 > DEP_2$; DEP_3,DEP_4,DEP_2;	$DEP_3 > DEP_4$ $DEP_2 > DEP_4$; ($DEP_3 \sim DEP_2$); $DEP_3 \sim DEP_2 > DEP_4$
4	DEP_3,DEP_1,DEP_2	$DEP_3 > DEP_4$ $DEP_4 > DEP_2$; $DEP_3 > DEP_4 > DEP_2$	$DEP_3 > DEP_2$ $DEP_2 > DEP_4$; $DEP_3 > DEP_2 > DEP_4$

In the case of RRT#1, let us take the classical case of a triplicate of choice: DEP_1, DEP_2, and DEP_3. Let us also

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

presume that the involved MADM method ranked the DEPs as DEP$_3$, DEP$_1$, DEP$_2$. In the case, where DEP$_1$ is no longer available as an option (and it is supplanted by a comparable DEP$_4$), the expected outcome might be: DEP$_3$, DEP$_4$, and DEP$_2$. However, in the case of RRT#1, the actual outcome might be DEP$_2$, DEP$_4$, and DEP$_3$ (wherein the actual potentially optimal DEP$_3$ is displaced from first position). RRT#2 is similar to RRT#1; however, it differs in that DEP$_1$ and DEP$_4$ would not be comparable, such as for the case wherein DEP$_4$ is far less optimal than DEP$_1$ (expressed as DEP$_1$ > DEP$_4$). In the case of RRT#3, a comparison would be made between the overarching ranking against the sub-rankings; for example, taking the initial RRT#1 ranking of DEP$_3$, DEP$_1$, DEP$_2$ along with the replacement of DEP$_1$ with DEP$_4$, the sub-rankings might equate to DEP$_3$ > DEP$_4$ and DEP$_4$ > DEP$_2$. Yet, RRT#3 might manifest as having the sub-rankings of DEP$_3$ > DEP$_4$ and DEP$_2$ > DEP$_4$; DEP$_3$ and DEP$_2$ might be construed as being similar in that they are both > DEP$_4$ (expressed as DEP$_3$ ~ DEP$_2$), and an outcome could be DEP$_3$ ~ DEP$_2$ > DEP$_4$. RRT#4 is akin to RRT#3; however, it differs in that only sub-ranking inconsistencies are focused upon. For example, DEP$_3$ > DEP$_4$ and DEP$_4$ > DEP$_2$ could be construed as being consistent since DEP$_3$ > DEP$_4$ > DEP$_2$; if, however, the sub-ranking outcome was DEP$_3$ > DEP$_2$ and DEP$_2$ > DEP$_4$, which equates to DEP$_3$ > DEP$_2$ > DEP$_4$, then RRT#4 would have manifested itself. This progression continues for numerous other RRTs.

IV. EXPERIMENTATION FOR THE INVOLVED CASE STUDY

Zizovic et al. introduced the RAFSI method to mitigate against RR. Zizovic points out that a consistent/steady-state ranking across various scenarios (e.g., S0 through S5) constitutes mission success for the RR problem, such as exemplared in Zizovic's RAFSI Table 2 (exhibited as Table V) [64]. However, the anticipated results for the approach utilized in this paper would differ from Zizovic's RAFSI Table 2 (exhibited as Table V), as time is treated [64].

TABLE V. ZIZOVIC'S RAFSI "RANKING OF THE ALTERNATIVES IN SCENARIOS" [64]

Alternatives	***Scenarios***					
	S0	***S1***	***S2***	***S3***	***S4***	***S5***
A5	1	1	1	1	1	1
A1	2	2	2	2	2	
A4	3	3	3	3		
A2	4	4	4			
A3	5	5				
A6	6					

When considering just one of the EDQTs of Table III (e.g., EDQT#1, which centers upon the temporal aspect), the re-mapped (and simplified) table (using just the initial ranking of Table III) might resemble something like Table VI below.

TABLE VI. EDQT CONSIDERATIONS AND RE-MAPPING OF TABLE IV

Alternatives	***Scenarios***					
	UDC of LHM (EDQT#1)			***CDC of LHM (EDQT #1)***		
	S0	***S1***	***S2***	***S3***	***S4***	***S5***
DEP$_3$	1	1	1	3	3	3
DEP$_1$	2	2	2	1	1	1
DEP$_2$	3	3	3	2	2	2

Moreover, when considering EDQT#1 to 5, there are some significant reversals of findings when considering even simply UDC and CDC (of EDQT #1). For example, when comparing the medians of S1 (of the UDC scenarios) and S4 (of the CDC scenarios) of Table VI, whereas the initial ranking and expected ranking of DEP$_2$ were not in first position when treated generally, its ranking rose when considered against EDQT#1 to 5 (e.g., CDC), such as shown in Figures 1 and 2 below.

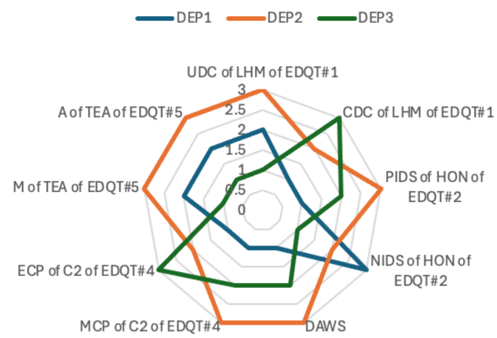

Figure 1. EDQT#1 to 5 for Scenario S1

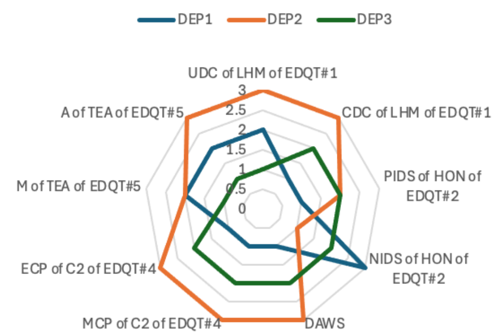

Figure 2. EDQT#1 to 5 for Scenario S4

Cognizant of the desired endstate so as to address RWS, a bespoke experimental architectural construct was further examined. Previously, the construct utilized was a plain vanilla MADM/MODM SM/OM counterpoising to comprise a C-MCDM. This is delineated in DQ#1 Bullet (B) 1 and B2, DQ#3 B2 and B3, DQ#4 B2, and DQ#5 B1 of Table III. For this paper, the construct was revised from that of Figure 3 (the TPMFS #s are from Table II) to Figure 4 so as to decrease the weighting of the MADM and to incorporate more apropos methods (that are more resistant/stable against RR); The BI is a Particle Swarm Optimization-centric Robust Convex Relaxation Framework (implementation details are delineated in DQ#1 B1, DQ#3 B1 and B3, and DQ#5 B1 through B5 of Table III), *is* equates to input set, and *ss* equates to solution set.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

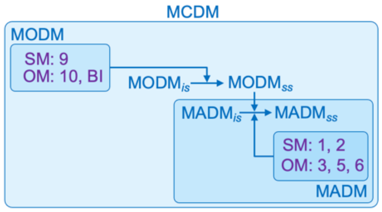

Figure 3. Prior Architectural Construct without RR Considerations

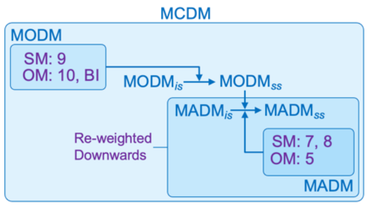

Figure 4. Current Architectural Construct with RR Consideration

Further experimentation was conducted to explore the TEA factor of the various methods employed. By way of example, PROMETHEE (TPMFS#7) was utilized as it is "easily… understood" [78][79]. Likewise, [fuzzy] VIKOR (TPMFS#6) was removed as it "less explainable than other more intuitive methods" [80]. These actions lend toward improving the System TEA (STEA). TPMFS#1 and 2 were removed for axiomatic RR reasons. TPMFS#5 remained in use and TPMFS#8 was added for their higher R/S with regards to RR.

This paper explored a particular facet of MCDM systems — the counterpoising of MADM/MODM SM/OM, wherein MADM would employ methods that exhibited higher R/S as pertains to RR and MADM, in general, would be re-weighted downwards. Given that the RR phenomenon greatly affects the most popular constituent component of MCDM — MADM — this constituted a non-trivial research goal. In addition, there was a constraint to select MADM methods that were more inclined towards the TEA aspiration (e.g., PROMETHEE is more intuitive and explainable). Among other advances, the research goal was approached from an EDQT vantage point, and the list of utilized methods was modified/winnowed from MADM SM 1,2 and OM 3, 5, 6 to MADM SM 7, 8 and OM 5. Two other non-trivial advancements should also be illuminated. First, the Abbas and Howard six fundamentals of DQ was extended for RWS via EDQT#1 through 5 and the foray explored within this paper — EDQT#6. The practicalities of EDQT#1 through 6 should not be underestimated. Second, the Zizovic RAFSI method to mitigate against RR was extended for RWS by considering the temporal element (from EDQT#1), such as that of UDC S0 to S2 and CDC S3 to S5. The aforementioned advancements were incorporated into the STEA advancement — the formulation of a bespoke architectural construct with RR considerations, such as reflected in Figure 4. The MODM OM BI was previously shown to have high efficacy in shaping an optimized selection of MADMs, so the new amalgam construct of Figure 4 constitutes an enhanced approach towards the treatment of RR. It should be noted that, depending upon the specific implementation, TPMFS#5 and 8 can exhibit drawbacks (when putting aside the TEA and R/S RR considerations) for factors, such as Flexibility (F) (for integration, hybridization, adaptation, etc.), Consistency (C), and Performance (P), as shown in Table VI below; TPMFS#6 is exhibited for comparison purposes only.

TABLE VII. EXEMPLAR BENCHMARKING FOR SELECT TPMFS

TPMFS #	*R/S RR*	*TEA*	*F*	*C*	*P*
5					
8					
6					

The range of MADM methods (e.g., ML, neural network, and other advanced computational methods) is constrained to those, for the purposes of this paper, deemed to exhibit higher practicality by way of being TEA-centric and suited for R/S RR.

V. CONCLUSION

This paper explores the challenges of RR in MCDM, specifically within MADM methods. Experimentation was provided through case studies that emphasize the temporal and control aspects. The paper integrates a variety of DQ dimensions (DQ#1 to DQ#6), which demonstrate: (1) how the model can be adapted to various DM contexts, and (2) how the overarching framework is well-suited for RWS applications. While various decision-making systems within the literature explore dynamic systems and/or describe time-sensitive DM, this paper differs via the unique amalgam treatment of EDQTs for the various DQs delineated in Table II. Planned future work includes a more granular comparison (the value of a quantitative comparison is still nebulous, as there are quantitative exactitude issues surrounding the involved benchmarking) as well as a prioritization list (e.g., a prioritization of TEA, R/S RR, C, F, and P), which will be informed by a survey to be conducted. Further MADM methods, pertaining to the listed factors, will also be explored.

REFERENCES

[1] R. Confalonieri, L. Coba, B. Wagner, and T. Besold, "A Historical Perspective of Explainable Artificial Intelligence," Wires Data Mining and Knowl. Discovery, vol. 11, pp. 1-21, Jan./Feb. 2021.

[2] M. Heder, "Explainable AI: A Brief History of the Concept," *Eur. Res. Consortium for Inform. and Math. (ERCIM) News*, vol. 134, Jun. 2023.

[3] M. Hosain, et al., "Path to Gain Functional Transparency In Arificial Intelligence With Meaningful Explainability," *J. of Metaverse,* vol. 3, pp. 166-180, Dec. 2023.

[4] M. Heder, "The Epistemic Opacity of Autonomous Systems and the Ethical Consequences," *AI & Society,* vol. 38, pp. 1819-1827, Jul. 2020.

[5] "IEEE P7001 Working Group," Meeting Draft Agenda July 31, 2017 10AM-11AM EDT Recorded by Natasha Alvarado, Secretary, Piscataway, NJ, Jul. 2017.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

[6] L. Methnani, M. Chiou, V. Dignum, and A. Theodorou, "Who's in Charge Here? A Survey on Trustworthy AI in Variable Autonomy Robotic Systems," *ACM Comput. Surv.*, vol. 56, pp. 1-32, Apr 2024.

[7] A. Winfield, et al., "IEEE P7001: A Proposed Standard on Transparency," *Front. Robot. AI*, vol. 8, pp. 1-11, Jul. 2021.

[8] "Statement on Algorithmic Transparency and Accountability," ACM US Public Policy Council, Washington, DC, Jan. 2017.

[9] B. Velden, "Explainable AI: Current Status and Future Potential," *Eur. Radiol.,* vol. 34, pp. 1187-1189, Aug. 2023.

[10] S. Nisanth, R. Manu, B. Sagar, T. Padmashree, and N. Cauvery, "Performance of CPUs and GPUs on Deep Learning Models for Heterogeneous Datasets," *2022 6th Int. Conf. on Electron., Commun. and Aerosp. Technol.,* Coimbatore, India, Dec. 2022, pp. 978-985.

[11] M. Dorrich, M. Fan, and A. Kist, "Impact of Mixed Precision Techniques on Training and Inference Efficiency of Deep Neural Networks," *IEEE Access*, vol. 11, pp. 57627-57634, Jun. 2023.

[12] "Innovation Through Artificial Intelligence Patents," USPTO [Online], Accessed: Feb. 20, 2025. Available: https://patentsview.org/ai-and-innovation.

[13] "USPTO releases new Artificial Intelligence Patent Dataset," USPTO [Online], Accessed: Feb. 20, 2025. Available: https://www.uspto.gov/about-us/news-updates/uspto-releases-new-artificial-intelligence-patent-dataset.

[14] N. Pairolero, A. Giczy, G. Torres, T. Erana, M. Finlayson, and A. Toole, "The Artificial Intelligence Patent Database (AIPD) 2023 Update," USPTO [Online]. Accessed: Feb. 20, 2025. Available: https://www.uspto.gov/sites/default/files/documents/oce-aipd-2023.pdf.

[15] "Vision for Competitiveness: Mid-Decade Opportunities for Strategic Victory," The Special Competitive Studies Project [Online]. May 2024. Available: https://www.scsp.ai/wp-content/uploads/2023/04/Vision-for-Competitiveness-1-1.pdf.

[16] E. Triantaphyllou, B. Shu, S. Sanchez, and T. Ray, "Multi-Criteria Decision Making: An Operations Research Approach," in *Encyclopedia of Elect. and Electron. Eng.,* J. G. Webster, Ed., New York, NY, USA: John Wiley & Sons, 1998, pp. 175-186.

[17] G. Fattoruso, "Multi-Criteria Decision Making in Production Fields: A Structured Content Analysis and Implications for Practice," *J. Risk Financial Manag.*, vol 15, pp. 1-21, Sep. 2022.

[18] S. Gebre, D. Cattrysse, E. Alemayehu, and J. Orshoven, "Multi-criteria decision making methods to address rural land allocation problems: A Systematic Review," *Int. Soil and Water Conservation Res.,* vol. 9, pp. 490-501, Dec. 2021.

[19] F. Sitorus, J. Cilliers, and P. Brito-Parada, "Multi-criteria decision making for the choice problem in mining and mineral processing: Applications and trends," *Expert Syst. with Appl.*, vol. 121, pp. 393-417, May 2019.

[20] S. Chan, "AI-Facilitated Dynamic Threshold-Tuning for a Maritime Domain Awareness Module," *2024 IEEE Int. Conf. on Ind. 4.0, Artif. Intell., and Commun. Technol. (IAICT)*, Bali, Indonesia, Jul. 2024, pp. 192-198.

[21] G. Conroy, "What's wrong with the h-index, according to its inventor," *Nature Index* [Online]. Accessed: Oct. 17, 2024. Available: https://www.nature.com/nature-index/news/whats-wrong-with-the-h-index-according-to-its-inventor.

[22] "WIPO Technology Trends 2019: Artificial Intelligence," World Intellectual Property Organization [Online]. Accessed: Oct. 17, 2024. Available: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf.

[23] V. Lenarduzzi and D. Taibi, "MVP Explained: A Systematic Mapping Study on the Definitions of Minimal Viable Product," *Conf. on Softw. Eng. and Adv. Appl. (SEAA)*, Limassol, Cyprus, Aug./Sep. 2016, pp. 112-119.

[24] S. Chan, "A Potentially Specious Cyber Security Offering for 5G/B5G/6G: Software Supply Chain Vulnerabilities within Certain Fuzzing Modules," *The Sixth Int. Conf. on Cyber-Technol. and Cyber-Syst.*, Barcelona, Spain, Oct. 2021, pp. 43-50.

[25] D. Boix-Cots, F. Pardo-Bosch, and P. Pujadas, "A systematic review on multi-criteria group decision-making methods based on weights: Analysis and classification scheme," Inf. Fusion, vol. 96, pp. 16-36, Aug. 2023.

[26] J. Girard, et al., "Reconsidering the Duchenne Smile: Indicator of Positive Emotion or Artifact of Smile Intensity?" *Int. Conf. Affect. Comput. Intell. Interact. Workshops*, Dec. 2019, pp. 594-599.

[27] E. Clark, et al., "The Facial Action Coding System for Characterization of Human Affective Response to Consumer Product-Based Stimuli: A Systematic Review," *Front Psychol*., vol 11, pp. 1-21, May 2020.

[28] Y. Bogodistov and F. Dost, "Proximity Begins with a Smile, But Which One? Associating Non-duchenne Smiles with Higher Psychological Distance," *Front. Psychol.*, vol. 8, pp. 1-9, Aug. 2017.

[29] S. Lyons-Padilla, et al, "Race Influences Professional Investors' Financial Judgments," *Proc. Natl. Acad. Sci*., vol. 116, pp. 17225-17230, Aug. 2019.

[30] "The History of 360 Degree Feedback," Organisation Development & Research Limited (ODRL) [Online]. Accessed: Oct. 17, 2024. Available: https://www.odrl.org/2019/12/27/360-degree-feedback-history/.

[31] "Katharine Cook Briggs: The Beginning of Type," Myers & Briggs Foundation [Online]. Accessed: Oct. 17, 2024. Available: https://www.myersbriggs.org/about-us/myers-briggs-jung-legacy/.

[32] "History and Validity of the Thomas-Kilmann Conflict Mode Instrument," The Myers-Briggs Company [Online]. Accessed: Oct. 17, 2024. Available: https://ap.themyersbriggs.com/themyersbriggs-history-validity-tki.aspx.

[33] H. Taherdoost and M. Madanchian, "Multi-Criteria Decision Making (MCDM) Methods and Concepts, *Encyclopedia 2023*, vol. 3, pp. 77-87, Jan. 2023.

[34] M. Tavana, M. Soltanifar, F. Santos-Arteaga, and H. Sharafi, "Analytic hierarchy process and data envelopment analysis: A match made in heaven," *Expert Syst. with Appl*., vol. 223, pp. 1-16, Aug. 2023.

[35] A. Ezzat and H. Hamoud, "Analytic hierarchy process as module for productivity evaluation and decision-making of the operation theater," *Avicenna J. Med*, vol. 6, pp. 3-7, Jan-Mar. 2016.

[36] A. Baykasoglu and E. Ercan, "Analysis of rank reversal problems in 'Weighted Aggregated Sum Product Assessment' method," *Softw. Comput*., vol. 25, pp. 15243-15254, Dec. 2021.

[37] V. Simic, D. Lazarevic, and M. Dobrodolac, "Picture fuzzy WASPAS method for selecting last -mile delivery mode: a case study of Belgrade," *Eur. Transp. Res. Rev*., vol. 43, pp. 1-22, Jul. 2021.

[38] A. Lubis, N. Khairina, and M. Riandra, "Comparison between Multiple Attribute Decision Making Methods through Objective Weighting Method in Determining Best Employee," *Int. J. of Innov. Res. in Comput. Sci. & Technol. (IJIRCST)*, vol. 11, pp. 2347-5552, Mar. 2023.

[39] E. Zavadskas, Z. Stevic, Z. Turskis, and M. Tomasevic, "A Novel Extended EDAS in Minkowski Space (EDAS-M) Method for Evaluating Autonomous Vehicles," *Stud. in Inform. and Control*, vol. 28, pp. 255-264, Oct. 2019.

[40] R. Kumar and S. Singal, "Penstock material selection in small hydropower plants using MADM methods," *Renewable and Sustain. Energy Rev.*, vol. 52, pp. 240-255, Dec. 2015.

[41] D. Kacprzak, "A Novel Extension of the Technique for Order Preference by Similarity to Ideal Solution Method with Objective Criteria Weights for Group Decision Making with Interval Numbers," *Entropy*, vol. 23, pp. 1-20, Sep. 2021.

[42] V. Thiagarasu and V. Rengaraj, "A MADM Model with VIKOR Method for Decision Making Support Systems," *Int. J. of Novel Res. in Comput. Sci. and Softw. Eng*., vol. 2, pp. 63-81, Jan.-Apr. 2015.

[43] Y. Suh, Y. Park, and D. Kang, "Evaluating mobile services using integrated weighting approach and fuzzy VIKOR," *PloS One*, vol. 14, pp. 1-28, Jun. 2019.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. **ISBN: 978-1-68558-259-3**

[44] A. Chaudhuri and T. Sahu, “PROMETHEE-based hybrid feature selection technique for high-dimensional biodmedical data: application to Parkinson’s disease classification,” *Electron. Lett.*, vol. 56, pp. 1403-1406, Oct. 2020.

[45] M. Basilio, V. Pereira, and F. Yigit, “New Hybrid EC-Promethee Method with Multiple Iterations of Random Weight Ranges: Step-by-Step Applicadtion in Python,” *MethodsX.*, vol. 13, pp. 1-16, Aug. 2024.

[46] H. Taherdoost and M. Madanchian, “Using PROMETHEE Method for Multi-Criteria Decision Making: Applications and Procedures,” *Iris J. of Econ. & Bus. Manage.*, vol. 1, pp. 1-7, May 2023.

[47] L. Vasto-Terrientes, A. Valls, R. Slowinski, and P. Zielniewicz, “ELECTRE-III-H: An outranking-based decision aiding method for hierarchically structured criteria,” *Expert Syst. with Appl*., vol. 42, pp. 4910-4926, Jul. 2015.

[48] M. Boujelben, “A unicriterion analysis based on the PROMETHEE principles for multicriteria ordered clustering,” *Omega,* vol. 69, pp. 126-140, Jun. 2017.

[49] E. Hoeijmakers et al., “How subjective CT image quality assessment becomes surprisingly reliable: pairwise comparisons instead of Likert scale,” *Eur. Radiol.* vol. 34, pp. 4494-4503, Jul. 2024.

[50] H. Taherdoost and M. Madanchian, “A Comprehensive Overview of the ELECTRE Method in Multi-Criteria Decision-Making,” *J. of Manage. Sci. & Eng. Res*., vol. 06, pp. 1-16, Sep. 2023.

[51] A. Hafezalkotob, A. Hafezalkotob, H. Liao, and F. Herrera, “An overview of MULTIMOORA for multi-criteria decision-making: Theory, developments, applications, and challenges,” *Inf. Fusion*, vol. 51, pp. 145-177, Nov. 2019.

[52] M. Mahjoob and P. Abbasian, “Designing a Cost-Time-Quality-Efficient Grinding Process Using MODM Methods,” *Eur. J. of Eng. and Technol. Res.,* vol. 7, pp. 12-17, Mar. 2022.

[53] A. Abbas, “Multi-objective Decision Making: Expected Utility vs. Some Widely Used (and Flawed) Methods,” in *Improving Homeland Secuirty Decisions*, A. Abbas, M. Tambe, and D. Winterfeldt, Eds., Cambridge, England, UK: Cambridge University Press, 2017, pp. 396- 426.

[54] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior*, 2nd rev. ed., Princeton, NJ, USA: Princeton University Press, 1947.

[55] C. Wu, E. Schulz, T. Pleskac, and M. Speekenbrink, “Time pressure changes how people explore and respond to uncertainty,” *Nature Sci. Rep.*, vol. 12, pp. 1-14, Mar. 2022.

[56] S. Chan, “The Triumvirate of an Adaptive Criteria Weighting Methodology, Isomorphic Comparator Similarity Measure, and Apropos High Dimensional Data Cluster Validation Index Measures for the Ascertainment of Bespoke Dynamic Fuzzy Lists,” *Future Comput. 2024*, pp. 1-10, Apr. 2024.

[57] L. Tian, X. Zhou, Y. Wu, W. Zhou, J. Zhang, and T. Zhang, “Knowledge graph and knowledge reasoning: A systematic review,” *J. of Lectr. Sci. and Technol.*, vol. 20, pp. 1-19, Jun. 2022.

[58] Y. Chen et al., “An Overview of Knowledge Graph Reasoning: Key Technologies and Applications,” *J. Sens. Actuator Netw*., vol. 11, pp. 1-26, Nov. 2022.

[59] S. Ji, S. Pan, E. Cambria, P. Martinen, and P. Yu, “A Survey on Knowledge Graphs: Representation, Acquisition, and Applications.” *IEEE Trans. Neural Netw. and Learn. Syst*., vol. 33, pp. 494–514, Feb. 2022.

[60] B. Cai, Y. Xiang, L. Gao, H. Zhang, Y. Li, and J. Li, “Temporal Knowledge Graph Completion: A Survey,” *Proc. of the Thirty-Second Int. Joint Conf. on Artif. Intell. (IJCAI-23),* Jan. 2022, pp. 6545-6553.

[61] M. Nayyeri, “Dihedron Algebraic Embeddings for Spatio-temporal Knowledge Graph Completion,” *European Semantic Web Conf. (ESWC)*, Hersonissos, Greece, May 2022, pp. 253-269.

[62] P. Pradhan, “Time-Dependent Properties of Sandpiles,” *Frontiers in Phys*., vol 9, pp. 1-13, May 2021.

[63] S. Chan, I. Oktavianti and P. Nopphawan, “Optimal Convex Relaxation-based Wavelet Covariance Transform for More Robust AOD-PM Characterization and Tracer Tracking of Biomass Burning Over Land/Sea Boundary Regions,” *2022 IEEE Ocean Eng. Technol. and Innov. Conf.: Manage. and Conservation for Sustain. and Resilient Marine and Coastal Resour. (OETIC)*, Surabaya, Indonesia, 2022, pp. 1-10.

[64] M. Zizovic, D. Pamucar, M. Albijanic, Prasenjit Chatterjee, and I. Pribicevic, “Eliminating Rank Reversal Problem Using a New Multi-Attribute Model-The RAFSI Method,” *Math.,* vol. 8, pp. 1-16, Jun 2020.

[65] M. Garcia-Cascales and M. Lamata, “On rank reversal and TOPSIS method,” *Math. and Comput. Model.*, vol. 56, pp. 123-132, Sep. 2012.

[66] R. Aires and L. Ferreira, “The Rank Reversal Problem in Multi-Criteria Decision Making: A Literature Review,” *Pesquisa Oper.*, vol. 38, pp. 331-362, May 2018.

[67] A. Majumdar, M. Tiwari, A. Agarwal, and K. Prajapat, “A new case of rank reversal in analytic hierarchy process due to aggregation of cost and benefit criteria,” *Oper. Res. Perspectives*, vol. 8, pp. 1-10, Apr. 2021.

[68] N. Munier, “A New Approach to the Rank Reversal Phenomenon in MCDM with the SIMUS Method,” *Multiple Criteria Decis. Making*, vol. 11, pp. 137-152, 2016.

[69] M. Zizovic, “Eliminating Ranking Reversal Problem Using a New Multi-Attribute Model-The RAFSI Method,” *Math.*, vol. 8, pp. 1-18, Jun. 2020.

[70] B. Kizielewicz, A. Shekhovtsov, and W. Salabun, “A New Approach to Eliminate Rank Reversal in the MCDA problems,” *Comput. Sci. - The Int. Conf. on Comput. Sci. (ICCS),* Jun. 2021, pp. 338-351.

[71] J. Dezert, A. Tchamova, D. Han, J. Tacnet, “The SPOTIS Rank Reversal Free Method for Multi-Criteria Decision-Making Support,” *IEEE 23rd Int. Conf. on Inf. Fusion (FUSION)*, Jul. 2020, pp., 1-8.

[72] J. Wieckowski, R. Krol, J. Watrobski, “Towards robust results in Multi-Criteria Decision Analysis: ranking reversal free methods case study,” *Int. Conf.e on Knowledge-Based and Intell. Inf. & Eng. Syst. (KES)*, 2022, pp. 4584-4592.

[73] H. Maleki and S. Zahir, “A Comprehensive Literature Review of the Rank Reversal Phenomenon in the Analytic Hierarchy Process,” *J. of Multi-Criteria Decis. Anal.: Optim., Learning, and Decis. Support*, vol. 20, pp. 141-155, Aug. 2012.

[74] X. Yu, W. Yang, and S. Suntrayuth, “A New Approach to Solving TOPSIS Rank Reversal Based on S-Type Utility Function,” *Int. J. of Innov. Comput., Inf. and Control*, vol. 19, pp. 1501-1516, Oct. 2023.

[75] M. Aazadfallah, “A New Feature of Rank Reversal in Some of MADM Models,” J. of Appl. Inf. Sci., vol. 4, pp. 1-11, 2015.

[76] Y. Wang, and Y. Luo, “On rank reversal in decision analysis,” *Math. And Comput. Modelling*, vol. 49, pp. 1221-1229, Mar. 2009.

[77] Y. Wang and T. Elhag, “An approach to avoiding rank reversal in AHP,” Decision Support Syst., vol 42, pp. 1474-1480, Dec. 2006.

[78] M. Brans and B. Mareschal, “Promethee Methods,” *Multiple Criteria Decision Analysis: State of the Art Surveys*, pp. 163-186, Jan. 2005.

[79] J. Brans and P. Vincke, “A Preference Ranking Organisation Method: (The PROMETHEE Method for Multiple Criteria Decision-Making), *Manage. Sci*., vol. 31, pp. 647-656, Jun. 1985.

[80] Y. Lin and T. Chen, “Type-II fuzzy approach with explainable artificial intelligence for nature-based leisure travel destination selection amid the COVID-19 pandemic,” *Digit. Health,* vol. 8., pp. 1-15, Jun. 2022.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

Fugacity Phase Transition and Hyper-Heuristic Convergence for AI-centric Conceptual Estimating

Steve Chan
VTIRL, VT/DE-STEA
Orlando, USA
Email: stevec@de-stea.tech

***Abstract*—Levels of effort and timetable posits for the development and operationalization of System Transparency, Explainability, and Accountability (STEA)-centric Artificial Intelligence (AI) Systems (AIS) are beset by underestimation in often overlooked areas, such as the "Optimizing" facet of the "Deploying and Optimizing" phase of the AI Development Life Cycle, among others. This is a high derailment factor in conceptual estimating, particularly for those mission-critical AIS that do not well consider biases stemming from the broader Socio-Technical System (STS), which impact Interpretability & Explainability (I&E). In furtherance of bias mitigation and AIS whitening — STS-STEA-I&E (SSI) — an amalgam construct for facilitating/discerning a Fugacity Phase Transition (FPT) and Hyper-Heuristics (HH) convergence, segueing to an enhanced SSI contribution, is delineated.**

Keywords-AI Development Life Cycle; interpretability; explainability; justification logic; decision engineering.

I. Introduction

The development and deployment of Artificial Intelligence (AI) Systems (AIS) is on the rise, and the rapid growth in market size for AIS and related supply chains have been abundantly memorialized; Compound Annual Growth Rates (CAGR), such as 36.6% from 2024 to 2030, have been reported [1]. Cisco's 2024 AI Readiness Index asserts that "nearly all companies (98%) report that the urgency to deploy AI has increased in the last year" [2]. Rough Order of Magnitude (ROM) cost estimates for Levels of Effort (LOEs) and the associated timetables for the design/development and operationalization of these AIS are being requested in a torrential fashion to keep pace with the escalating demand/adoption rate [3][4]. This is buttressed by Stanford University's AI Index Report 2024, which notes a dramatic increase of interest in GitHub AI projects (more than doubling between 2022 to 2023) [5]. Simply, AIS are in high demand.

Despite the $184 billion market size for AI as of November 2024, the anticipated $826 billion market size by 2030, and the rising price tags for AIS deployments, conceptual estimating (e.g., positing ROMs prior to the substantial completion of the involved architecture/design) has not yet become sufficiently mature and/or robust; these ROMs are often far off target with a plethora of cost/schedule overruns and project failures populating the landscape [6][7][8][9]. Generally speaking, cost estimates are typically predicated upon the historical costs of successfully completed projects, and since the corpus of historical data is still quite limited in this arena, a myriad of conceptual estimating and cost estimator issues have arisen; ROMs are often erroneous.

To aggravate matters, not all AIS are equal. By way of example, a number of the earlier AIS had been withdrawn from the market due to their problematic "black box" architectures and prospective biases (e.g., algorithmic), which had not been well accounted for during their architectural/design phases [10]. Since that time, the AI ecosystem has progressively moved toward a paradigm of System Transparency, Explainability, and Accountability (STEA) for the prototypical stages/phases of the AI Development Life Cycle (ADLC) (as pertains to the development and operationalization of an AIS). The number of phases varies depending upon organizational preference and model selection — e.g., 3, 5, 6, 8, etc.; for simplicity, 3 phases will be considered herein; of the 3 basic phases — (1) Planning & Collection, (2) Designing & Training, and (3) Deploying & Optimizing — the "Optimizing" facet (a substantive contributor towards the success of the AIS) of (3) constitutes a formidable STEA challenge. *Without careful consideration, the STEA treatment for "Optimizing" can dramatically increase the required LOEs and potentially derail any posited ADLC timetable for the STEA-centric AIS.*

Yet, without even considering the STEA complexities and requisite mitigations against biases stemming from the larger Socio-Technical System (STS) rubric, which includes the ecosystem of "humans, technology, and the environs," there are a variety of staggering statistics to consider: (1) the Project Management Institute has reported that "almost half of business projects fall behind schedule, and up to a third are not completed at all," (2) a Boston Consulting Group (BCG) survey reports that "nearly half of all respondents said that more than 30% of their organization's technology development projects were over budget and late," (3) McKinsey & Company (McK), in collaboration with the [BT Group plc, formerly British Telecom] BT Centre for Major Programme Management at the University of Oxford, reports that "on average, large [Information Technology] IT projects run 45 percent over budget and 7 percent over time, while delivering 56 percent less value than predicted" while McKinsey further reports that "software projects run the highest risk of cost and schedule overruns," (4) [Research & Development] RAND Corporation notes that, "by some estimates, more than 80 percent of AI projects fail — twice the rate of failure for information technology projects," and

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

(5), The Computing Technology Industry Association (CompTIA) notes that “nearly 80% of the AI projects typically don’t scale beyond a [Proof of Concept] PoC or lab environment” [11]-[16]. Against this backdrop, when the complexities of IT/AI projects are conjoined with the cited STEA and STS complexities, it becomes clear that the devising of a robust STS/STEA-centric AIS architecture is non-trivial. Accordingly, four central aspects, among others, need to be well considered for an STEA-centric AIS architecture prior to providing a ROM.

The first is the desired level of transparency. The literature describes the principal variations in AIS architecture — “black-box,” “gray-box,” and “white-box — as being distinguished by gradations in transparency (most opaque to most transparent). The second is the desired level of interpretability, which centers upon the AIS’s Decision Engineering/Decision-Making (DE/DM) processes. The third is the desired level of explainability, which centers upon the rationale/underlying logic employed to arrive at the, hopefully, non-biased and reasonable outcomes [17]; the University of Toronto’s Schwartz Reisman Institute for Technology & Society and others further distinguish between Explainable AI (which centers upon “fact”) and Justifiable AI (a.k.a., justifiability) (which centers upon “judgment”) [18]. The fourth is the degree of accuracy (ACC) desired. For the second and third aspects, *interpretability* describes *how* the AIS formulates certain posits (e.g., the DE/DM processes), and *explainability* describes *why* the AIS made certain posits (e.g., the justification logic). These (i.e., Interpretability and Explainability) are often referred to as I&E, and along with the fourth aspect, there is an ongoing dialectic in the literature regarding the trade-off between *ACC* and *I&E*. Some argue that reduced ACC AIS are more readily interpreted; along this vein, some argue that enhanced ACC AIS are less able to be interpreted in an intuitive fashion [19][20]. A similar argument has been made regarding explainability [21][22]. Amidst this backdrop, researchers have endeavored to achieve high-performance AIS that still have high I&E [23]. Suffice it to say, this arena constitutes a challenging study space.

In the interim, research forays have trended towards more transparent white-box (a.k.a., glass-box) architectures, which reputedly have better I&E-by-design [24]. However, the performance tends to, as reported by some, lag behind the more translucent/opaque black-box architectures [25]. Accordingly, researchers have actively investigated the feasibility of middle-ground gray-box architectures. Along the vein of the previously discussed AIS project cost/schedule overruns, the initial *development time* for a high-performance **S**TS/**S**TEA/**I**&E (SSI)-centric AIS architecture can vary greatly (e.g., from months to years), and Gartner notes that, generally, “organizations” take about “7 months to develop AI initiatives, with 47% of the surveyed companies taking between 6 to 24 months from prototype to production” [12][26][27]; some AIS implementers assert that SSI-centric AIS architectures can take several years to devise and realize. The *testing times* can also vary greatly. Generally speaking, black-box testing can require less time than white-box testing since the latter would require additional LOEs (i.e., an increased amount of time) to comprehend the DE/DM pathways and logic employed. The AIS *model training time*/cost is also highly variable, as the training data needs to be refreshed in an ongoing fashion, particularly for Real World Scenario (RWS) AIS applications. With regards to the “Optimizing” facet of (3) of the ADLC, the degree of ACC (versus I&E) needs to be specified, and the various involved optimizations (e.g., pertaining to the involved computational resources, quantity/quality of the training data, heuristics/algorithms employed, tuning/fine-tuning efficacy for [e.g., Deep Neural Network or DNN] weights/hyperparameters, complexity of the AIS model and AIS architecture/design, etc.) is central. Stanford University and Epoch AI (a multidisciplinary research institute that investigates the arc of AI) reviewed AI model training cloud compute times/costs, and MIT Technology Review noted that “the process used to build most of the… [AI] models we use today can’t tell if they will work in” RWS, “and that’s a problem” [28][29][30]. Some AIS implementers argue that the greater the desired level of SSI, the “more time-consuming and resource-intensive” the processes can be — with an ensuing increase to Capital Expenditures (CAPEX). Over time, the seeming CAPEX advantage of “black-box” over “white-box” architectures may potentially be offset by ever-escalating Operational Expenditures (OPEX) related to brittleness and obsolescence issues (e.g., undetected issues, such as data drift may result in dramatic performance degradation) that often beset black-box architectures; in other words, the downstream OPEX-related disadvantages may offset the initial CAPEX advantages of the earlier developmental and testing phases. Gray-box architectures seem to constitute a middle-ground.

Certain SSI challenges that beset the ADLC are illuminated within this paper, such as at the “Optimization” facet of (3) of the ADLC. To assist the reader, a table of acronyms is provided in Table I below.

TABLE I. TABLE OF ACRONYMS

Acronym	***Full Form***
ACC	Accuracy
ACM	Association for Computing Machinery
AdapHH	Adaptive selection Hyper-Heuristics
ADLC	AI Development Life Cycle
AI	Artificial Intelligence
AIS	Artificial Intelligence System
ALGB-WG	Algorithmic Bias Working Group
BCG	Boston Consulting Group
CAGR	Compound Annual Growth Rate
CAPEX	Capital Expenditure
CompTIA	Computing Technology Industry Association
CRITIC	CRiteria Importance through Intercriteria Correlation
CWA	Connection Weights Algorithm
DE	Decision Engineering
DM	Decision-Making
DNN	Deep Neural Network
EO	Expert Opinion
FPT	Fugacity Phase Transition
GA	Garson’s Algorithm
GI	Gini Importance
HH	Hyper-Heuristic
HH-CF	Choice-Function-based Hyper-Heuristic
HH-R	Reward-based Hyper-Heuristic
HH-SF	Statistical Frequency-based Hyper-Heuristic

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

I&E	Interpretability & Explainability
IEEE	Institute of Electrical and Electronics Engineers
LOE	Level of Effort
MADM	Multi-Attribute Decision-Making
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MCDM	Multi-Criteria Decision-Making
McK	McKinsey & Company
MDA	Mean Decrease in Accuracy
MDI	Mean Decrease in Impurity
MLR	Multiple Linear Regression
MODM	Multi-Objective Decision-Making
NIST	National Institute of Standards and Technology
OA	Olden's Algorithm
OM	Objective Measure
OPEX	Operational Expenditure
OPH	Operator/Procedure/Heuristic
PLSR	Partial Least Squares Regression
POC	Proof of Concept
PROMETHEE	Preference Ranking Organization Method for Enrichment Evaluation
PSO	Particle Swarm Optimization
QoS	Quality of Service
QR	Quantile Regression
RAND Corp.	Research & Development Corporation
RMSE	Root Mean Square Error
ROM	Rough Order of Magnitude
RR	Ridge Regression
RWS	Real World Scenario
SM	Subjective Measure
SSHH	Sequence-based Selection Hyper-Heuristic
SSI	Socio-Technical System-System Transparency, Explainability, and Accountability-Interpretability & Explainability
STEA	System Transparency, Explainability, and Accountability
STS	Socio-Technical System
TransE	Translating Embeddings
WIP	Work-in-Progress
XAI	Explainability in AI

Section I delineates the impetus of the paper — *the illumination and consideration of certain SSI-related derailment facets that may dramatically increase the required LOEs and potentially derail posited ADLC timetables for the SSI-centric AIS.* Section II provides pertinent background information regarding: (1) **S**TS/**S**TEA (in general) and **I**&E (in particular) (collectively, "SSI") for certain facets of the ADLC for an AIS, (2) the "Fugacity Phase Transition" (FPT) (e.g., the series of deviations between the "ideal" training data and the "actual" observed data), and (3) certain other key considerations (e.g., the SSI aspects of the utilized Hyper-Heuristics or HH) that are critical to consider prior to putting forth conceptual estimating ROMs for an SSI-centric AIS. Section III delineates the presets & theoretical foundations as well as benchmarking & insights related to the involved HH/FPT experimentation. Section IV concludes and presents some prospective future work.

II. Background

A. The Import of SSI for AIS

In these contemporary times, there is a heightened expectation for SSI-centric AIS, particularly with regards to I&E. Winfield and others have remarked on various STEA-centric Work-in-Progress (WIP) standards, as well as actual standards that have buttressed the Explainability in AI (XAI) movement; these WIPs/standards include, among others, the U.S. National Institute of Standards and Technology (NIST) Special Publication 1270 "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence," the Association for Computing Machinery (ACM) "Principles for Algorithmic Transparency and Accountability," and the Institute of Electrical and Electronics Engineers (IEEE) Standard for Transparency of Autonomous Systems (P7001), among others. There are also a range of engaged working groups, such as the IEEE Algorithmic Bias Working Group (ALGB-WG) (P7003). On the topic of bias, NIST has opined that certain AI biases (e.g., "human biases and systemic, institutional biases as well") may stem from the larger STS rubric [31]. This includes the involved corpus of data, which may, potentially, derive from problematic "facts," "assessment surveys," and other bias-related problems from the "Collection of Data" facet of (1) of the ADLC [32][33].

Traditionally, it has been opined that, for the ADLC, approximately "80% time" is spent on (1) [34]. For the "Collection of Data" facet of (1) of the ADLC, Westland has noted that the "bias and informativeness" of Subjective Measures (SMs) (e.g., Likert-type measurements) "have been the center of recent" dialectic [35][36]. From an SSI perspective, STS-related biases, such as from a variety of assessment data utilized as input to the AIS (e.g., from surveys) has recently been illuminated as a prospective Achilles heel for AIS. For example, McLeod informs us that "prior research has shown that using Likert scales can be problematic," via a variety of biases (e.g., "social desirability bias, acquiescence bias," central tendency bias, etc.) [37]. Taherdoost affirms this by noting that Likert "scale validity may be difficult to demonstrate[,] and there is a lack of reproducibility" [38]. To further underscore the aforementioned, Louangrath's experimentation reports on the higher reliability levels of non-Likert scales (e.g., "92%") over Likert-type scales (e.g., "90, 89, and 88% reliability") as well as higher validity levels of non-Likert scales (e.g., "93%") over Likert-type scales (e.g., "89, 61, and 57%") [39]. Hence, *the formulation/implementation of enhanced assessments (e.g., STS-related surveys) for the "Collection of Data" facet of (1) of the ADLC, which is a key part of the STS rubric, will likely increase the time needed for formulating and instantiating SSI-centric AIS architectures.*

B. The Fugacity Phase Transition (FPT) between Phases (2) and (3) of the ADLC

With regards to the AIS architecture's DE/DM apparatus, Fattoruso depicts Multi-Criteria Decision-Making (MCDM) as being comprised of Multi-Attribute Decision-Making (MADM) and Multi-Objective Decision-Making (MODM) [40]. Generally speaking, while MODM concurrently addresses a range of objectives ("and endeavors to determine an optimal solution set among "undetermined continuous alternatives"), MADM addresses a single objective and "organizes/sorts/ranks" (in the endeavor to ascertain the optimal solution among "a finite set of discrete alternatives") [41]. For the "Collection of Data" facet of (1) of the ADLC, a more robustly counterpoised MADM/MODM

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

SM/Objective Measures (OM) construct is crucial for facilitating SSI robustness, as it can better contend with the issue of AIS model drift (a.k.a., model decay) (i.e., shifts in the involved data/relationships that can result in AIS model performance degradation, wherein the posits become increasingly less effective), particularly in situations for which the RWS data encountered is far different "from the data it was trained to recognize or handle" [42]. Generally speaking, it can be easier to discern this drift within a higher SSI-centric than a lower SSI-centric AIS architecture. To assist in contextualizing/delineating this paradigm, the term "fugacity" (an apropos term utilized by Dreyfus-Schmidt-DuPhan-Desfontaines that nicely references the "tendency… to escape from one phase to another") is utilized; "'fugacity' measures the difference between the expected 'ideal' data… [that the AIS] model was trained on and the observed 'real' data" that the AIS model encounters (i.e., the distinction between the "reference distribution" and the "prediction distribution") [43][44][45]. The indicators of low drift and low fugacity can be utilized in ascertaining when a transitioning from phase (2) to (3) of the ADLC (i.e., FPT) is prudent. It should be noted that the FPT is not a singular punctuating event/milestone; rather, it denotes a fairly steady-state paradigm, wherein the fugacities for the successive states of dynamically updated AIS heuristics (acting in conjunction with the involved AIS algorithms) are low enough to be of satisfactory utility for the involved RWS AIS application. *The monitoring of the involved AIS model (and encompassing AIS architecture) will require a sufficient temporal span given the SSI-centric AIS architectural requirement.*

C. *Potential ADLC pitfalls (e.g., HH) and I&E Robustness for the "Optimization" Facet of the ADLC*

As alluded to in Section I and Section IIA, the requisite time to develop a sufficiently robust performance SSI-centric AIS architecture can vary greatly. It consists of the (1), (2), and (3) phases referenced in Section I, as well as various facets, such as that of "Optimization." Within the phases of a 3-phase ADLC, the *"Collection of Data"* and "*Training/Inferencing"* (e.g., which might be subject to the prospective inversion of the classical training:inferencing ratios) facets, as discussed in [46], are noteworthy, for they need to be well considered prior to positing LOEs and their associated timetables for an SSI-centric AIS (i.e., conceptual estimating).

The counterpoising of SM and OM (for MADM and MODM), such as for the *"Collection of Data"* facet of (1) of the ADLC is non-trivial. This is further complicated with the need to appropriately weight and "organize/sort/rank," which may be accomplished via the utilization of various OM combinatorials; this includes the leveraging of OM methods, such as the CRiteria Importance through Intercriteria Correlation (CRITIC) OM for the ascertainment of apropos weights and the Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) OM for the ensuing ranking. *The apropos selection and testing of more SSI-oriented OM (as well as SM) combinatorials will also likely increase the time needed for SSI-centric AIS architectures.*

Moreover, there is a fundamental distinction between the paradigm of static weights and that of dynamically updated weights. The literature is abundant with regards to the criticality of a dynamic weighting strategy for the "*Training/Inferencing"* facet [47]. Along this vein, oftentimes, heuristic approaches are leveraged to complement algorithmic approaches, particularly for RWS AIS applications. After all, the amalgam of heuristics and algorithms lend to numerical methods implementations of higher efficacy, and a dynamically updated heuristic model lends to more optimal convergence for a "better-fit" or "best-fit" approximation, etc. (e.g., the robust convex relaxation discussed in [48]). *The SSI-related issue is that while algorithms have received increasing SSI attention, the myriad of static/brittle heuristics populating the AIS landscape has not received comparable SSI attention; this is an area that can increase the ADLC time needed.*

Beyond the *"Collection of Data"* and the "*Training/Inferencing"* facets, the *"Optimization"* facet of the ADLC (e.g., optimizing the involved AIS model) is critical, for it facilitates more accurate and efficient predictions, which segues to enhanced performance, decreased OPEX, and higher practicality/applicability for RWS. In particular, optimization (e.g., such as with regards to AIS model size, complexity, etc.) can facilitate more rapid inferencing with less computational resources (e.g., energy consumption) and lend toward scalability (e.g., optimized AIS models are more readily deployed). By way of context, the heuristic problem-solving approach is geared for ascertaining a "good enough" solution within a bounded period of time, but there is no certainty that it will provide an optimal solution; in contrast, certain algorithmic approaches are favored for ascertaining an optimal solution, but the "runtimes" may vary greatly. To date, "research in the explainability of optimisation techniques has largely focused on meta-heuristics" (which "directly search the solution space of a problem") [49][50]. There has been far less research on HH (higher-level Operator/Procedure/Heuristic (OPH) methods that "operate on a search space of low[er]-level heuristics…rather than solutions directly"), which can pose herculean SSI challenges due to the use of a plethora of lower-level OPHs, which complicates matters [51]. There have been some notable SSI-related explorations that have shown promise with regards to SSI, such as Misir's Adaptive selection HH (AdapHH) and Drake's Sequence-based Selection HH (SSHH) (which leverages probability matrices to facilitate I&E) [49][51]. By leveraging these lessons learned as well as the presets delineated in Section IIIA, a more SSI-centric HH paradigm can be leveraged.

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

III. Experimentation

A. Presets & Theoretical Foundations

Ali, Piccialli, and others have noted that a substantive portion of AI researchers opine that "a deeper network is better for decision-making than a shallow network" [24]. Yet, the prototypical DNNs are increasingly more difficult to examine at the deeper layers given the increasingly complex patterns/abstractness (as contrasted to the more straightforward patterns residing at the more shallow/earlier layers). From an SSI perspective, this makes certain reportage of DNN usage for mission-critical HH of even greater import [52]. For the experimentation herein, a preset (i.e., a precursor experimental construct) leveraged was in the form of an RWS-oriented Particle Swarm Optimization (PSO)-based Meta-Heuristic approach, as depicted in [53]. Another preset centered upon the selection of MODM OMs, such as CRITIC and PROMETHEE as well as those delineated in [54]. These presets are reflected in Figure 1 in bright red; the critical counterpoisings shown in lavender are supported by these presets. The focus of the experimentation is at the "Optimization" nexus of FPT/HH (denoted in brick red).

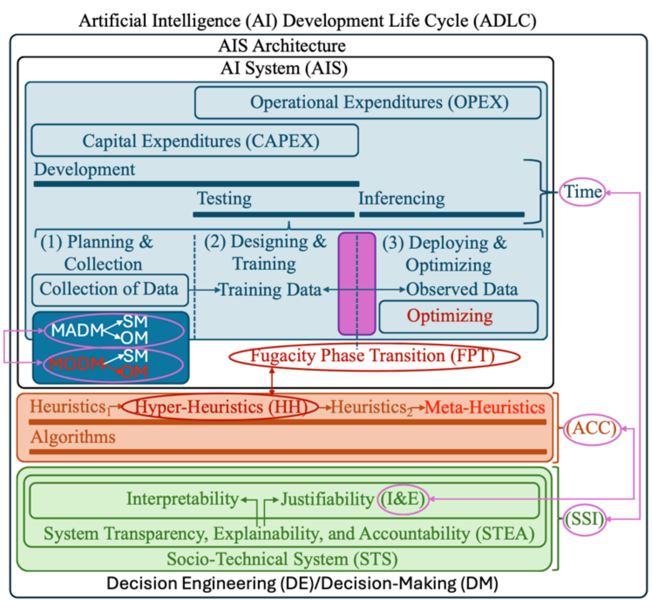

Figure 1. ADLC with the experimental focus, as indicated in brick red.

Drake asserts that there are two types of HHs: (1) selection HH that select a sequence of Low[er]-Level Heuristics (LLHs), and (2) generation HHs that spawn LLHs [51]. From an SSI perspective, the degree of I&E depends upon the involved mechanism; for example, Maashi's Choice-Function-based HH (HH-CF) facilitates the examination of LLHs, as LLHs are designated with a score/normalized score based upon prior performance and chosen accordingly. Qu's Statistical Frequency-based HH (HH-SF) can reveal LLH sequences that relate to the more optimal solutions (thereby making I&E more self-evident). Kheiri's Reward-based HH (HH-R) leverages LLH usage and transitions among LLH to yield transition probabilities for enhanced I&E [49].

B. Benchmarking & Insights

For the purposes herein, operators will be construed as: (1) diversification, (2) intensification, and (3) perturbation, Typically, (1) will leverage randomness to induce a substantive variation (e.g., to avoid stagnation at local optima) to expand the search space (e.g., progress to unexplored areas), (2) will spawn solution variations in high potential areas of the search space, and (3) will induce minute variations (e.g., to facilitate the gauging of LLH performance). In some cases, the sequencing of (1), (2), and (3) is effective; in other cases, (3), (1), and (2) may have efficacy. Our experimentation finds that the (1), (2), (3), (2) sequence has high efficacy; our findings are consistent with Drake's reportage that LLHs/LLH sequences "which are ineffective at the start of the search process prove to be highly effective at the end, and vice versa" [51]. In essence, the efficacy of LLHs/LLH sequences and their concomitant HHs need to be gauged *over time*. For this temporal consideration, the assessment of the LLHs/HHs also needs to include consideration of the long-tail (part of the [statistical distribution], which is far afield from the head and centroid) phenomena prevalent in RWS; Samuel reports that "strongly unbalanced data with a long-tail is ubiquitous in numerous domains and problems" and "learning [*over time*] with unbalanced data causes models to favor head classes" [55][56]. Various techniques (e.g., based upon Wang's Translating Embeddings or TransE) for better balancing across both head and tail classes are discussed in [57]. There is also the matter of AIS model drift *over time*. Along this vein, HHs can be leveraged to avoid a high drift paradigm (i.e., a drift score closer to 1), such as for the case where the features underlying the AIS model drift are of low significance; HHs can also be leveraged to lower the drift paradigm (e.g., moving the drift score closer to 0) by recognizing features of high significance, whose removal would dramatically degrade the AIS model performance. Interestingly, the challenge of feature significance determination centers upon the fact that features are not independent; actually, a substantive portion of features are highly correlated (a.k.a., collinear features). Spearman's and Pearson's correlation [coefficient] (R) can be used to gauge collinearity (e.g., a high R indicates collinearity), and given the plethora of collinear features, the notion of feature families becomes quite useful. Given a high R^2 (a value closer to 1, which implies a perfect fit), wherein R^2 = 1 – Sum Squares of Error or SSE/Total Sum of Squares or SST, the removal of a high dependency feature will likely not have a significant impact upon ACC for the feature family; on the other hand, a lower Root Mean Square Error (RMSE) (square root of the average squared differences between the measured values and actual values) and Mean Absolute Error (MAE) (average of the absolute differences) implies a

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

better fit. As Matel notes, "the larger the drop in R^2 when a variable [/feature] is removed…, the more important it is assumed to be" [58]. This is affirmed by Gini Importance (GI), Mean Decrease in Impurity (MDI), and Mean Decrease in Accuracy (MDA) (a higher GI, MDI, and MDA indicates higher variable/feature significance). In essence, the involved RWS AIS evaluation was conducted *over time* (i.e., the FPT).

Matel's experimentation was utilized for benchmarking purposes, as Matel had reported that his conceptual estimating model exhibited "a 14.5% improvement in the accuracy" over Hyari's model when considering Mean Absolute Percentage Error (MAPE) [58]. Matel's findings are as follows: (1) for the Connection Weights Algorithm (CWA), "the lowest MAPE with all 16 variables was 50.36%," but the MAPE dropped "to 27.41%" "when only the top 5 variables were used," (2) for Multiple Linear Regression (MLR), "when [only] the top 5 to 7 variables" were used, the MAPE was "42.47%," and (3) for Expert Opinion (EO), when only "the top 5 variables" were used, the "MAPE was 93.25%" [58]. Hence, in terms of efficacy, CWA >> MLR >>> EO; this should be no surprise, for while CWA can accommodate non-linear relationships, MLR is not able to. For the case herein, Matel's experimentation was reiterated with HH utilized for determining the top variable/features, and the results were somewhat comparable. The results are shown in Figure 2, which also incorporates Garson's Algorithm (GA) and Olden's Algorithm (OA) as alternatives to CWA as well as Partial Least Squares Regression (PLSR), Quantile Regression (QR), and Ridge Regression (RR) as alternatives to MLR.

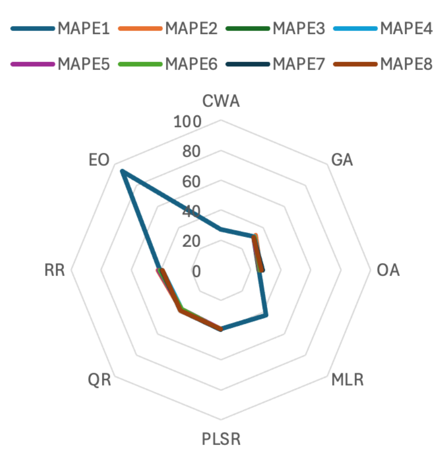

Figure 2. Benchmarking of Section III with Matel's Experimentation

CWA tends to outperform GA, and OA (as an implementation of CWA) is more nuanced than the plain vanilla CWA. PLSR is better suited for multi-collinearity than MLR, and QR can better handle outliers than MLR. Apart from that, the principal distinction was that of a steady-state convergence that was obtained with the amalgam of: (1) low drift, (2) low RMSE and MAE reflecting low fugacities/a more narrow FPT, (3) high GI, MDI, and MDA affirming variable/feature significance, (4) high R (reflecting collinearity) and a high R^2 (wherein the removal of high dependency features did not have a substantive ACC impact), and (5) high efficacy HH ascertainment at 8 variables/features. This logical progression through the amalgam composition and FPT/HH convergence should make clear the FPT/HH SSI contribution.

IV. CONCLUSION

The use of heuristics, to assist with algorithmic convergence for RWS AIS applications, is on the rise. These applications are likely to have specific stringent RWS timing requirements (e.g., pursuant to the involved Quality of Service or QoS). The adherence to these stringent RWS timing requirements constitutes a key facet of why the dynamically updated heuristic model (e.g., via HH) tangibly contributes towards the utility/practicality expected for RWS applications. *Hence, HHs become critical to the equation, and their SSI orientation becomes central;* it should be noted that HH has gained traction "in addressing NP-hard optimisation problems because it generalises well across problem domains" [59]. This paper presented an FPT/HH convergence approach (i.e., low drift, narrow FPT, and high efficacy HH) that would lend to a more SSI-centric optimization facet of the ADLC; accordingly, conceptual estimating and cost estimator ROMs can be made more robust. To conclude, this paper explores the development and implementation of improved assessment methods, such as STS-oriented surveys, for the "Data Collection" process within ADLC, a key component of the STS framework. The study highlights how these enhancements may impact the creation and deployment timelines of AIS architectures focused on SSI. By refining evaluation approaches, the research aims to improve the efficiency and effectiveness of data-driven decision-making within STS-based systems. Future work will involve more quantitative experimentation and benchmarking.

REFERENCES

[1] "Global Artificial Intelligence Market Size & Outlook," *Grand View Research*. [Online]. Accessed: Mar. 1, 2025. Available: https://www.grandviewresearch.com/horizon/outlook/artificial-intelligence-market-size/global.

[2] "Cisco's 2024 AI Readiness Index: Urgency Rises, Readiness Falls," *Cisco*, Nov. 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2024/m11/cisco-2024-ai-readiness-index-urgency-rises-readiness-falls.html#:~:text=Nearly%20all%20companies%20(98%25),from%2014%25%20a%20year%20ago.

[3] "Artificial Intelligence's Use and Rapid Growth Highlight Its Possibilities and Perils," *U.S. Gov. Accountability Office*, Sep. 2023. [Online]. Accessed: Mar. 1, 2025. Available: https://www.gao.gov/blog/artificial-intelligences-use-and-rapid-growth-highlight-its-possibilities-and-perils.

[4] "Groundbreaking Framework for the Safe and Secure Deployment of AI in Critical Infrastructure Unveiled by Department of Homeland Security," *U.S. Department of*

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Homeland Security, Nov. 2024. [Online]. Accessed: Mar. 1, 2025. Available from: https://www.dhs.gov/news/2024/11/14/groundbreaking-framework-safe-and-secure-deployment-ai-critical-infrastructure#:~:text=These%20sectors%20are%20increasingly%20deploying,distribute%20mail%20to%20American%20households.

[5] "Artificial Intelligence Index Report 2024," *Stanford University*. [Online]. Accessed: Mar. 1, 2025. Available: https://aiindex.stanford.edu/wp-content/uploads/ 2024/05/ HAI_AI-Index-Report-2024.pdf.

[6] "Artificial Intelligence – Worldwide," *Statista*, Mar. 2024. [Online]. Accessed: Mar. 1, 2025. Available from: https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide.

[7] "Artificial Intelligence (AI) market size worldwide from 2020 to 2030," *Statista*, Mar. 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide# market-size.

[8] B. Cottier, R. Rahman, L. Fattorini, N. Maslej, and D. Owen, "The rising costs of training frontier AI models," *arXiv*, May 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://arxiv.org/abs/2405.21015.

[9] "AI exposes challenges to cost estimation," Acquisition Talk, Jul. 2019. [Online]. Accessed: Mar. 1, 2025. Available from: https://acquisitiontalk.com/2019/07/ ai-exposes-challenges-to-cost-estimation/.

[10] K. Hao, "The two-year fight to stop Amazon from selling face recognition to the police," *MIT Technology Review*, Jun. 2020. [Online]. Accessed: Mar. 1, 2025. Available: https://www.technologyreview.com/2020/06/12/ 1003482/ amazon-stopped-selling-police-face-recognition-fight/.

[11] B. Velden, "Explainable AI: Current Status and Future Potential," *Eur. Radiol.,* vol. 34, pp. 1187-1189, Aug. 2023.

[12] M. Lorko, M. Servatka, and L. Zhang, "A Better Way to Avoid Project Delays," *MIT Sloan Manage. Rev.*, Nov. 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://sloanreview.mit.edu/article/a-better-way-to-avoid-project-delays/.

[13] S. Palumbo, B. Rehberg, and H. Li, "Software Projects Don't Have to Be Late, Costly, and Irrelevant," *Boston Consulting Group*, Apr. 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://www.bcg.com/publications/2024/software-projects-dont-have-to-be-late-costly-and-irrelevant.

[14] M. Bloch, S. Blumberg, and J. Laartz, "Delivering large-scale IT projects on time, on budget, and on value," *McKinsey Digital*, Oct. 2012. [Online]. Accessed: Mar. 1, 2025. Available: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/delivering-large-scale-it-projects-on-time-on-budget-and-on-value.

[15] J. Ryseff, B. Bruhl, and S. Newberry, "The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed," *RAND*, Aug. 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://www.rand.org/pubs/research_reports/RRA2680-1.html.

[16] "Artificial Intelligence in Business: Top Considerations Before Implementing AI," *CompTia*. [Online]. Accessed: Mar. 1, 2025. Available: https://connect.comptia.org/content/guides/business-considerations-before-implementing-ai.

[17] "Logic & Reasoning for XAI," *The 2nd World Conf. on Explainable Artif. Intell.* [Online]. Accessed: Mar. 1, 2025. Available from: https://xaiworldconference.com/2024/logic-reasoning-for-xai/

[18] G. Hadfield, "Explanation and justification: AI decision-making, law, and the rights of citizens," *Schwartz Reisman Inst. for Technol. and Soc., Univ. of Toronto*, May 2021. [Online]. Accessed: Mar. 1, 2025. Available from: https://srinstitute.utoronto.ca/news/hadfield-justifiable-ai.

[19] B. Seo and J. Li, "Explainable machine learning by SEE-Net: closing the gap between interpretable models and DNNs," *Sci. Rep.*, vol. 14, pp. 1-16, Nov. 2024.

[20] W. Esser-Skala and N. Fortelny, "Reliable interpretability of biology-inspired deep neural network," *npj Systems Biology and Appl.*, vol. 9, pp. 1-8, Oct. 2023.

[21] A. Nussberger, L. Luo, L. Celis, and M. Crockett, "Public attitudes value interpretability but prioritize accuracy in Artificial Intelligence," *Nature Commun.*, vol. 13, pp. 1-13, Oct. 2022.

[22] V. Hassija, et al., "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence," *Cognitive Computation*, vol. 16, pp. 45-74, Aug. 2024.

[23] F. Candelon, T. Evgeniou, and D. Martens, "AI Can Be Both Accurate and Transparent," *Harvard Bus. Rev.*, May 2023. [Online]. Accessed: Mar. 1, 2025. Available: https://hbr.org/2023/05/ai-can-be-both-accurate-and-transparent.

[24] S. Ali, et al., "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence," *Inf. Fusion*, vol. 99, pp. 1-52, Nov. 2023.

[25] O. Loyola-Gonzalez, "Black-Box vs. White Box: Understanding Their Advantages and Weaknesses From a Pracitcal Point of View," *IEEE Access*, vol 7., pp. 154096-154113, Oct 2019.

[26] D. Lambert, "How To Architect and Deliver AI At Scale," *Enterprise Architecture Professional J. (EAPJ)*, Oct. 2023. [Online]. Accessed: Mar. 1, 2025. Available: https://eapj.org/how-to-architect-and-deliver-ai-at-scale/#:~:text=According%20to%20Gartner%20it%20takes,months%20from%20prototype%20to%20production.

[27] "Gartney Survey Finds Generative AI Is Now the Most Frequently Deployed AI Solution in Organizations," *Gartner*. May 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://www.gartner.com/en/newsroom/press-releases/ 2024-05-07-gartner-survey-finds-generative-ai-is-now-the-most-frequently-deployed-ai-solution-in-organizations.

[28] "The Training Costs of AI Models Over Time," *Voronoi*, May 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://www.voronoiapp.com/technology/The-Training-Costs-of-AI-Models-Over-Time-1334.

[29] *"How Much Does It Cost to Train Frontier AI Models," Epoch AI.* [Online]. Accessed: Mar. 1, 2025. Available from: https://epoch.ai/blog/how-much-does-it-cost-to-train-frontier-ai-models.

[30] W. Heaven, "The way we train AI is fundamentally flawed," *MIT Technology Review*, Nov. 2020. [Online]. Accessed: Mar. 1, 2025. Available: https://www.technologyreview.com/2020/11/18/1012234/training-machine-learning-broken-real-world-heath-nlp-computer-vision/.

[31] R. Schwartz, et al., "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence," *NIST Special Publication 1270*, 2022.

[32] M. Gruet, " 'That's Just Common Sense'. USC researchers find bias in up to 38.6% of 'facts' used by AI," *USC Viterbi School of Engineering*, May 2022. [Online]. Accessed: Mar. 1, 2025. Available: https://viterbischool.usc.edu/news/2022/05/thats-just-common-sense-usc-researchers-find-bias-in-up-to-38-6-of-facts-used-by-ai/.

[33] E. Ntoutsi et al., "Bias in data-driven artificial intelligence systems-An introductory survey," *WIREs Data Mining and Knowledge Discovery*, vol. 10, pp. 1-14, May/June 2020.

[34] S. Sagiraju, "Managing the Data For The AI Lifecycle," *Forbes*, Apr. 2022. [Online]. Accessed: Mar. 1, 2025.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Available: https://www.forbes.com/councils/ forbestechcouncil/ 2022/04/20/managing-the-data-for-the-ai-lifecycle/

[35] B. Choi and A. Pak, "A Catalog of Biases in Questionnaires," *Prev Chronic Dis*., vol. 2, pp. 1-13, Dec 2004.

[36] J. Westland, "Information loss and bias in likert survey responses," *PLOS One,* vol. 17, pp. 1-17, Jul. 2022.

[37] A. McLeod, S. Pipp, and J. Wong, "Revisiting the Likert scale: can the fast form approach improve survey research?" *Int. J. of Behav. Accounting and Finance*, vol. 2, pp. 310-327, 2011.

[38] H. Taherdoost, "What Is the Best Response Scale for Survey and Questionnaire Design; Review of Different Lengths of Rating Scale/ Attitude Scale/Likert Scale," *Int J. of Acad. Res. in Manage. (IJARM)*, vol. 8, pp.1-12, Mar. 2019.

[39] P. Louangrath, "Reliability and Validity of Survey Scales," *Int. J. of Res. & Methodology in Social Sci*., vol. 4, p. 50-62, Mar. 2018.

[40] G. Fattoruso, "Multi-Criteria Decision Making in Production Fields: A Structured Content Analysis and Implications for Practice," *J. Risk Financial Manag.*, vol. 15, pp. 1-21, Sep. 2022.

[41] S. Chan, "AI-Facilitated Dynamic Threshold-Tuning for a Maritime Domain Awareness Module," *2024 IEEE Int. Conf. on Ind. 4.0, Artif. Intell., and Commun. Technol. (IAICT)*, Bali, Indonesia, Jul. 2024, pp. 192-198.

[42] S. Bigelow, "How to identify and manage AI model drift," *TechTarget*, May 2024. [Online]. Accessed: Mar. 1, 2025. Available from: https://www.techtarget.com/ searchenterpriseai/tip/How-to-identify-and-manage-AI-model-drift

[43] "Model Drift Monitoring," *Dataiku*. [Online]. Accessed: Mar. 1, 2025. Available: https://www.dataiku.com/product/plugins/model-drift/.

[44] M. Khan and M. Islam, "Sustainable Waste Management," in *The Petroleum Eng. Handbook: Sustainable Oper.*, Gulf Publishing Company, Houston, Texas, 2007, pp. 135-188.

[45] "Dataiku 11.3 Is Hot Off the Presses," *Dataiku*. [Online]. Accessed: Mar. 1, 2025. Available:https://community.dataiku.com/discussion/31960/new-year-new-features-dataiku-11-3-is-hot-off-the-presses.

[46] S. Chan, "Inference-Optimized Metaheuristic Approach for a Prospective AI Training/Inference Inversion Paradigm in Optimized Energy-Aware Computing," *2024 IEEE World AI IoT Congr. (AIIoT)*, Seattle, WA, USA, Jul. 2024, pp. 370-379.

[47] L. Liu, B. Zhou, Z. Zhao, and Z. Liu, "Active Dynamic Weighting for multi-domain adaptation," *Neural Netw.*, vol. 177, Sep. 2024.

[48] S. Chan, "Enhanced Robust Convex Relaxation Framework for Optimal Controllability of Certain Large Complex Networked Systems," *The 2022 IARIA Annual Congr. on Frontiers in Sci., Technol., Services, and Appl.*, Nice, France, May 2022, pp. 87-96.

[49] W. Yates, E. Keedwell, and A. Kheiri, "Explainable Optimisation through Online and Offline Hyper-heuristics," *ACM Trans. on Evol. Learn.*, pp. 1-29, Oct. 2024.

[50] M. Blocho, "Heuristics, metaheuristics, and hyperheuristics for rich vehicle routing problems," *Smart Delivery Syst.*, Jan. 2020, pp. 101-156.

[51] J. H. Drake, E. Özcan, and E. K. Burke, "A modified choice function hyper-heuristic controlling unary and binary operators," *2015 IEEE Congr. on Evolutionary Comput. (CEC)*, Sendai, Japan, May 2015, pp. 3389-3396.

[52] M. Panzer, B. Bender, and N. Gronau, "A deep reinforcement learning based hyper-heuristic for modular production control," *Int. J. of Prod. Res.h*, pp. 2747-2768,, Jul. 2023.

[53] S. Chan, "Inference-Optimized Metaheuristic Approach for a Prospective AI Training/Inference Inversion Paradigm in Optimized Energy-Aware Computing," *2024 IEEE World AI IoT Congr. (AIIoT)*, Seattle, WA, USA, Jul. 2024, pp. 370-379.

[54] S. Chan, "The Triumvirate of an Adaptive Criteria Weighting Methodology, Isomorphic Comparator Similarity Measure, and Apropos High Dimensional Data Cluster Validation Index Measures for the Ascertainment of Bespoke Dynamic Fuzzy Lists," *The 16th Int. Conf. on Future Comput. Technol. and Appl. (Future Comput. 2024)*, Venice, Italy, Apr. 2024, pp. 8-17.

[55] D. Samuel, Y. Atzmon, G. Chechik, "From generalized zero-shot larning to long-tail with class descriptors," *2021 IEEE Winter Conf. of Appl. of Comp. Vision (WACV)*, Waikoloa, Hawaii, 2021, pp. 286-295.

[56] M. Buda, A. Maki, and M. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Netw*., vol. 106, pp. 249-259, Oct 2018.

[57] S. Chan, "Cascading Succession of Models for an Enhanced Long-Tail Discernment AI System," *2024 IEEE World AI IoT Congr. (AIIoT)*, Seattle, WA, USA, 2024, pp. 393-401.

[58] E. Matel, "An artificial neural network apporach for cost estimation of engineering services," *Int. J. of Construction Manage*., vol. 22, pp. 1274-1287, Nov. 2019.

[59] E. Burke, M. Hyde, G. Kendall, G. Ochoa, E. Ozcan, and J. Woodward, "A Classification of Hyper-heuristic Approaches," in *Handbook of Metaheuristics*: *Int. Ser. in Oper. Res. & Manage. Sci.*, M. Gendreau and J. Potvin, Eds. Boston: Springer, 2010, pp. 449-46.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

Interstitial b-SHAP-centric Amalgam for the Enhancement of an AI-centric Construct Validity Approach

Steve Chan
VTIRL, VT/DE-STEA
Orlando, USA
Email: stevec@de-stea.tech

***Abstract*—This paper describes an Artificial Intelligence (AI)-based Construct Validity Verification Methodology (CVVM) being advanced. The proposed methodology includes an amalgam utilization of temporal-centric Finite-Change Shapley-Owen values along with, among others, Generic Shapley-Owen values and Variance-Based Shapley-Owen values (i.e., a bespoke SHAP amalgam or b-SHAP implementation), CRiteria Importance through Intercriteria Correlation (CRITIC), and Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) for enhancing the interpretability of not only the machine learning constituent components of an AI system, but also the interstices (e.g., between/among individual components as well as amalgams/clusters of components locally/globally). This approach extrapolates upon and furthers current proposals for the utilization of SHAP in local, global, and glocal (a hybridized intermediary of local and global) contexts. It turns out that this Interstitial SHAP-centric Amalgam (ISA), by better correlating features with each interim intended construct, potentially segues to better interpretability and construct validity at the component, interstitial, and overall system level, particularly when ISA is conjoined with a well-counterpoised Multi-Attribute Decision-Making (MADM)/Multi-Objective Decision-Making (MODM) Subjective Measures (SM)/Objective Measures (OM) paradigm and a modified Constriction Factor (CF)-Particle Swarm Optimization (PSO)-Robust Convex Relaxation (RCR)-Long Short-Term Memory (LSTM)-Deep Convolutional Neural Network (DCNN) (CPRLD) metaheuristic architectural construct.**

Keywords-artificial intelligence systems; machine learning; construct validity; explainability; interpretability.

I. Introduction

The impact of AI within the industrial sector and business, in general, should not be underestimated. Subhadra and others underscore the "rise of AI in business and industry" [1]. As AI is a transformative technology, it is envisioned to spur innovation and revolutionize various industries [2]. Honeywell's *Industrial AI Insights* report notes that, for the majority of cases, the "C-Suite has already decided to expand AI use," and in 91% of the cases, new use cases are brought to light "during AI implementation" [3]. Hence, AI forays are begetting further AI forays. These implementations involve AI software engineering, which leverages Machine Learning (ML) models and techniques to automate various tasks. The ML models of these AI Systems (AIS) are being increasingly relied upon to process/interpret *Big Data* so as to put forth meaningful forecasts/posits, thereby enhancing and illuminating certain Decision Engineering (DE) pathways so as to inform Decision-Making (DM).

A. The Criticality of Construct Validity

To ensure that the AIS ML models are robustly depicting the Real-World Scenarios (RWS), which they are tasked to emulate, the notion of *construct validity* becomes central. Sjoberg depicts construct validity as being "concerned with whether one can justifiably make claims at the conceptual level that are supported by results at the operational level" [4]; Sjoberg had conducted a Software Engineering (SE)-centric Systematic Literature Review (SLR) for the years 2000 through 2019 and determined that over this period of time, the prominence of the construct validity term rose by "sevenfold" [4]. Zhou affirms the criticality of validity within the SE sector and noted, comparatively speaking, the lack of research regarding the challenges related to construct validity [5]. Hence, despite the "sevenfold" increase, Deets and others find that the notion of construct validity is still "underdiscussed" [6]. As the ML models for AIS evolve, construct validity becomes particularly important to ensure that the involved progression leads to the intended construct. For example, construct validity can help ensure that the feature set aligns with the intended construct (i.e., feature alignment); also, given the understood constraints of the Shannon-Weaver model in communications theory, consideration of construct validity can help to avoid misinterpretation of the AIS ML model's posits (i.e., more robust interpretation). In essence, failure modes/blindspots and bias can be more readily identified and mitigated against.

B. Transparency, Explainability, and Accountability (TEA) Evaluation & Testing for Enhanced Construct Validity

Evaluation/testing (which ensures that the ML model well handles unseen data) and fine-tuning (which ensures that the ML model is optimized for a winnowed subset of data or particular task) are both integral for the enhancement of the involved AIS. The evaluation/testing of ML models involves both *construct validity,* as well as *performance metrics* to capture the intended construct and generalize well upon unseen data, respectively. The distinction is often not made, but *evaluation* and *testing* are quite marked and

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. **ISBN: 978-1-68558-259-3**

disparate. For example, with regards to performance metrics, evaluation tends to encompass accuracy, precision, recall, F1 score (determined by the precision and recall scores), Area Under the Receiver Operating Characteristic (AUC-ROC), cross-validation, etc. However, these types of evaluation do not provide insight into particular behaviors and/or potential Root Cause Analysis (RCA), which resides more in the realm of testing; while evaluation tends to focus upon performance of the model in its entirety, testing tends to focus upon the performance intricacies of the constituent components of the ML model. In the case of this paper, it is posited that the testing paradigm should also be extended to the interstices (e.g., interstitial areas between/among individual/amalgam of components, particularly in a glocal context). In any case, the evaluation/testing and fine-tuning paradigms are complicated enough for a single AIS, but in a System-of-Systems (SoS) (wherein constituent systems support the overarching function of the larger system) paradigm (wherein the incorrect testing and/or fine-tuning of one AIS may adversely impact another AIS), the notion of construct validity is crucial. The improving of AIS TEA at the component/interstitial areas can lead to enhanced construct validity, as feature alignment, more robust interpretation, etc. can likely be more readily achieved.

C. Enhancing TEA for Enhanced Construct Validity

Pathways for the advancement of System TEA (STEA) include a better understanding of the influence of Higher-Order Network (HONs), a finer-tuned Dynamic Assessment and Weighting System (DAWS) (wherein more apropos weights can be derived), as well as a more understandable/interpretable corpus of experience such that it can be better leveraged in a Lower Ambiguity (wherein the repertoire of experience suffices) Higher Uncertainty (LAHU) situation (given a sufficient repertoire of experience, the tolerance for uncertainty is higher, such that a decision can be made without, necessarily, the need for more *Big Data*) when time is of the essence. In addition, STEA-related SoS boundary areas also need to be taken into consideration as ML of ML becomes increasingly prevalent. After all, ML algorithms have a propensity to spawn “non-monotonic, non-polynomial [unable to be captured as a summation of terms], and even non-continuous functions” [7]. This is not dissimilar to the paradigm, wherein the transformation of “non-convex Mixed Integer Non-Linear Programming (MINLP) to convex problems, often spawn[ed] further non-convex MINLP problems” that necessitated further handling [8]. The enhancement of STEA can lead to better discernment of problematic constituent components (e.g., those exhibiting issues with *feature alignment*, *robust interpretation*, selection bias, etc.); this segues to enhanced *construct validity*.

Accordingly, this paper describes an AI-based Construct Validity Verification Methodology (CVVM) (i.e., the extent to which the AIS is accurately gauging the actual underlying concept/intended theoretical construct) being advanced. To assist the reader, a table of acronyms is provided in Table I as follows.

TABLE I. TABLE OF ACRONYMS

Acronym	***Full Form***
ACM	Association for Computing Machinery
ADMB	Automatic Differentiation Model Builder
AI	Artificial Intelligence
AIS	Artificial Intelligence System
AUC-ROC	Area Under the Receiver Operating Characteristic
c-SHAP	Classical Shapley Additive exPlanation
C2	Command and Control
CF	Constriction Factor
CNN	Convolutional Neural Networks
CPRLD	Constriction Factor-Particle Swarm Optimization-Robust Convex Relaxation-Long Short-Term Memory-Deep Convolutional Neural Network
CRITIC	CRiteria Importance through Intercriteria Correlation
CVVM	Construct Validity Verification Methodology
CWT	Continuous Wavelet Transform
DAWS	Dynamic Assessment and Weighting System
DCGAN	Deep Learning Convolutional Generative Adversarial Network
DCNN	Deep Convolutional Neural Network
DE	Decision Engineering
DeepLIFT	Deep Learning Important FeaTures
DL	Deep Learning
DM	Decision-Making
E	Execution Time
ELECTRE	ÉLimination Et Choix Traduisant la REalité
FCSO	Finite-Change Shapley-Owen
GAN	Generative Adversarial Network
GNU	GNU’s Not Unix
GPL	General Public License
Grad-CAM	Gradient-weighted Class Activation Mapping
GSO	Generic Shapley-Owen
HON	Higher-Order Network
I	Interpretability
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IPOPT	Interior Point OPTimizer
ISA	Interstitial SHAP-centric Amalgam
ISO	International Organization for Standardization
LAHU	Lower Ambiguity Higher Uncertainty
LIME	Local Interpretable Model Agnostic Explanations
LSTM	Long Short-Term Memory
MA	Model Agnostic
MADM	Multi-Attribute Decision-Making
MINLP	Mixed Integer Non-Linear Programming
ML	Machine Learning
MODM	Multi-Objective Decision-Making
MS	Model Specific
NP-hard	Non-deterministic Polynomial-time Hardness
OM	Objective Measure
OSNS	Optimal Shapley-Nondominated Solution
OSONS	Optimal Shapley-Owen-Nondominated Solution
PROMETHEE	Preference Ranking Organization Method for Enrichment Evaluation
PSO	Particle Swarm Optimization
RCA	Root Cause Analysis
RCR	Robust Convex Relaxation
RR	Rank Reversal
RWS	Real-World Scenarios
S	Sensitivity
SDP	Semi-Definite Programming
SE	Software Engineering
SHAP	Shapley Additive exPlanation
SLR	Systematic Literature Review
SM	Subjective Measure
SNOPT	Sparse Nonlinear OPTimizer
SoS	System-of-Systems
SQP	Sequential Quadratic Programming
STEA	System Transparency, Explainability, and

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

	Accountability
TEA	Transparency, Explainability, and Accountability
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
U	Performance under Uncertainty
V	Validity
VBSO	Variance-Based Shapley-Owen
VC-dim	Vapnik-Chervonenkis dimension
XAI	Explainable AI

Section I provided an overview, which underscored the criticality of the notion of *construct validity*. The remainder of this paper is organized as follows. Section II reviews the notion of AI SoS ML on ML and the need for STEA (particularly *interpretability* as an actualizing agent for enhanced STEA) to facilitate viable ML of ML. Section III presents theoretical foundations, the experimental testbed, and the experimental construct for addressing the challenge of AI-based CVVM. Section IV provides some concluding remarks and puts forth some future work.

II. BACKGROUND

A. AI System of Systems (SoS)

The notion of SoS is well-known; it is then axiomatic that an AI-related SoS is comprised of subordinate AIS. In theory, the involved ML at the top-tier AIS should be able to leverage the experiential base (e.g., lessons learned) of the lower-tier ML; in essence, the upper echelon ML should be able to enhance its efficacy by “learning” from the “successes” and “failures” of the lower echelon ML systems. This “learning” is effectuated by way of, among other types: (1) Collaborative learning (wherein ML systems collectively address a problem, such as in an ensemble and/or federated fashion, via *learning* from each other’s discernments and approaches), (2) Multi-agent reinforcement learning (wherein ML system *learnings* can inform the subsequent pathways undertaken by other AIS to achieve more optimal results), (3) Coopetition (a portmanteau of “cooperation” and “competition”) *learning*, such as in the case of Generative Adversarial Networks (GANs), wherein two AIS (e.g., generator and discriminator) engage in an “adversarial process” that segues to a win-win cooperative paradigm, (4) Transfer learning (e.g., wherein a pre-trained ML, with certain *learnings* already incorporated, can be fine-tuned and leveraged to undertake other tasks or wherein a distillation ML can transfer knowledge in a condensed form, thereby quickly enhancing efficacy and efficiency). However, to ascertain whether the *learnings* (e.g., employed approaches) are “effective” (or not) necessitates an AIS SoS ML on ML architecture that is more “white box” (e.g., wherein there is a higher degree of interpretability, such that the influencing variables are readily identifiable and the process — the involved model by which posits are generated — is more readily discernable) than “black box” (e.g., wherein opaqueness and/or translucency abounds); in other words, the desired “white box” AIS SoS ML on ML architectures need to have higher STEA (particularly interpretability).

B. AI-centric STEA and its Criticality for ML on ML

Along this vein, International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) 42001 focus upon AIS STEA; likewise, the Association for Computing Machinery (ACM) “Principles for Algorithmic Transparency and Accountability,” Institute of Electrical and Electronics Engineers (IEEE) Standard for Transparency of Autonomous Systems (P7001), and others follow suit. Addressing the “T,” a key factor for AIS architecture (e.g., “black-box,” “gray-box,” and “white-box”) is in the form of transparency (e.g., opaque, translucent, and fully transparent). Addressing the “E,” McKinsey portrays it as the “capacity to express why an AIS reached a particular decision, recommendation, or prediction” [9]; this tracks with prevailing definitions within the Explainable AI (XAI) field. Addressing the “A,” it involves the prior “T” and “E,” as the justification logic employed needs to be articulated; on this point, there is a nuance. While *explainability* and *interpretability* are often treated synonymously within the literature, perhaps they should be better distinguished. While *explainability* focuses upon *why* the AIS made certain posits, *interpretability* focuses upon *how* the AIS formulated its posits; restated, the latter delves into the AIS’s DE/DM processes to derive insights into the pathways for the justification logic involved. Together, *interpretability* & *explainability* are referred to as I&E, and I&E is a lynchpin for operationalizing effective ML of ML.

C. Interpretability and AIS SoS ML on ML Architecture

For the dual pillars of I&E, interpretability turns out to be paramount. Yet, despite its criticality, interpretability tends to be challenged by the degree of complexity of the involved AIS architecture. For example, Table I presents degrees of interpretability (wherein green denotes high, yellow denotes medium, orange denotes medium/low and red denotes low) for various complexities; there is a column “Monotonic” denoting when the ML model is monotonically constrained (wherein a change at the input variable segues to a change at the response function output), and there is a row “Linear” to indicate when the output is proportional to the input as well as a row “Non-linear” to denote when the relationship is more complex (e.g., convoluted interplays among features, ambiguous boundary areas, intricate sequences of local, glocal, and global transformations, etc.). Table I is rudimentary since, as noted in Section I, the spawning of “non-monotonic, non-polynomial, and even non-continuous functions” is not infrequent [7]; this greatly complicates matters, and gauges for interpretability are often tied to “measure[s] of model complexity,” such as “the Vapnik-Chervonenkis dimension (VC-dim)” [10]; the VC-dim can, by way of example, be indicative of the number of weights, rules, etc., (but does not equate to them).

TABLE II. EXEMPLAR ML MODEL PROCESS INTERPRETABILITY

	Monotonic	**Non-monotonic**
Linear		
Non-linear		

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

To date, STEA Efforts have tended to be on the *post*-side (e.g., Model Agnostic or MA), and those on the *pre*-post-side (e.g., Model Specific or MS) have had varied limitations. Exemplars of MA (e.g., Local Interpretable MA Explanations or LIME, Shapley Additive exPlanations or SHAP, etc.) and MS approaches (e.g., Gradient-weighted Class Activation Mapping or Grad-CAM, which is geared more for Convolutional Neural Networks or CNNs; Deep Learning Important FeaTures or DeepLIFT, which is geared for Keras and TensorFlow implementations; etc.) — the latter being constrained to a more limited set of ML models — are shown in Table II.

TABLE III. ML MODEL TYPES WITH EXEMPLAR I&E TOOLS

Model Specific (MS)	**Exemplar I&E Tools**	***Model Agnostic (MA)***
Linear Regression (LR)	e.g., InterpretML	e.g., LIME; SHAP
Decision Tree (DT)	e.g., GPTree	
Neural Network (NN)	e.g., Grad-CAM	
Deep Learning (DL)	e.g., DeepLift	

On the MS side, since the LR coefficients "directly represent the influence of each feature on the prediction," LR is construed as green when compared to the yellow of DT (which may have a complicated branching structure), the orange of NN (which may have complex internal workings, as contrasted to the more simplistic rules of DT), and the red of DL (which typically has a far greater number of layers than NN) [11]. On the MA side, LIME is oriented for more localized and individualized instances while SHAP capabilities extend beyond local and can well contribute towards a more global perspicacity across a gamut of instances; SHAP is well-suited to ascertain the more impactful features (i.e., as each feature will have a SHAP value to signify the impact on the posit, the features of import can be ascertained, and feature combinations that are able to maintain posit accuracy can be formulated while also considering the non-dominance principle, wherein no other feature combinations can provide posits without a degradation of efficacy in another facet) at the local, glocal, and global levels.

D. Optimal Shapley-Owen-Nondominated Solution (OSONS) for Enhanced STEA and Construct Validity

The Optimal Shapley-Nondominated Solution (OSNS) paradigm of Section IIC was explored as shown in Table III.

TABLE IV. EXEMPLAR DIGITAL OBJECT IDENTIFIERS (DOI) FOR VARIOUS FACETS OF OSNS

OSNS context	***Facet***	***DOI***
STEA	In general:	• 10.1109/AIIoT61789.2024.10579033 • 10.1109/OETIC57156.2022.10176215
	HON	• 10.1109/AIIoT61789.2024.10579029 • 10.1109/IBDAP62940.2024.10689701
	DAWS	• 10.1109/ICPEA56918.2023.10093212 • 10.1109/ICSGTEIS60500.2023.10424230
	LAHU	• 10.1109/GEM61861.2024.10585580
STEA-related SOS boundary areas	C2 of C2	• 10.1109/IEMCON.2019.8936241 • 10.1109/IAICT62357.2024.10617473
	ML of ML	**This paper**

In essence, it delineates prior work in the context of: (1) enhanced STEA, which facilitates a better understanding of the influence of HON-related drivers, a finer-tuned and more robust DAWS, and a more readily interpretable/leverageable repertoire of experience for a LAHU situation, as well as (2) STEA-related SoS boundary areas, such as those related to Command and Control (C2) of C2 (i.e., now ML of ML). For this paper, the notion of OSNS is expounded upon, as varied SHAP approaches differ in their local and global efficacies. By way of background, Borgonovo had referred to this hybridized efficacy as "glocal" (a portmanteau of "global" and "local"). Among other contributions, as a gauge of feature import (a key tasking of construct validity), SHAP values can be invaluable; Lundberg had advocated for SHAP to "explain various machine learning [ML] algorithms" [12]. With regards to the previously discussed (1) of this Section IID, Balog affirms the import of STEA-related HON-related drivers, and Sundararajan reinforces this perspective [13][14]. Kwon addresses the import of STEA-related DAWS, introduces "WeightedSHAP," and distinguishes it from the standard SHAP, which "uses the same weight for all marginal contributions;" Kwon also "demonstrates that the influential features identified by WeightedSHAP are better able to recapitulate the model's predictions compared to the features identified by the [classical] Shapley value" [15]. Addressing the matter from a different vantage point, Kotthoff raises the significance of utilizing the temporal-sensitive/temporal-centric (as contrasted with the classical) Shapley value, and the temporal-centric LAHU notion is delineated by the associated DOI shown in Table III [16]. With regards to the previously discussed (2) of this Section IID, Guidotti affirms the importance of ML model inspection at the margins (e.g., STEA-related SoS boundary areas) [17]. These SoS boundary areas refer to, among others, regions between/among individual/amalgam constituent components as well as local/glocal/global interstices. With regards to the former, Dhamdhere affirms the notion of "Shapley-Owen values" "for the quantification of joint contributions" [18]. With regards to the latter, Borgonovo advocates the use of *Finite-Change Shapley-Owen or FCSO values*, such as articulated by Dhamdhere), which are well suited for the *testing* facet (e.g., the discussed aspect of Section IB is more focused upon local/hyper-local scrutinization of the ML model) [18]; in conjunction with this, the Shapley-Owen values (generally, the *Generic Shapley-Owen or GSO valu*es, such as articulated by Grabisch, and more granularly, Borgonovo's suggested *Variance-Based Shapley-Owen or VBSO values*) can well serve in a generalized fashion — globally — across the model in its entirety [19][20]. Specifically, Borgonovo underscores the fact that *FCSO values* have equivalence to what Mase deemed to be the Baseline Shapley (i.e., the average of the *FCSO values* function under uncertainty) [20][21]; this Baseline Shapley also relates to the *VBSO*, since the upstream local finite-changes for the *FCSO values* segues to the Glocal Partial Dependence Function (which segues to the Conditional

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Regression Function and what Mase deemed to be the "Squared Cohorts" value function) [20][21]. Borgonovo notes that by averaging the *"Squared Cohorts" Shapley-Owen or SCSO values*, the VBSO values can be obtained [20]. This reflects one of the many interplays among local, glocal, and global, and is also indicative of how "additional insights into the [ML] model behavior" are possible [20]; these supplemental insights segue to enhanced construct validity, which provides the basis for more robust ML of ML.

III. Experimentation

ML of ML is a central tenet of this paper. To improve upon the ML model and the involved SoS, the need for interpretability (and STEA) is paramount. After all, constituent component and interstitial analyses is vital for determining whether the prospective ML learnings are of potential benefit; in some cases, RCA will be needed to discern and mitigate against problematic areas affecting performance. Borgonovo's glocal notion can help bridge the gap, and the significance of the OSNS segueing to an Optimal Shapley-Owen-Nondominated Solution (OSONS) paradigm is well articulated by Casajus, Lopez, Beal, and others [22][23][24]. In essence, the Owen value (which well captures the nuanced interactions between/among the members of the feature set) extends the Shapley value (which well captures the individual feature contributions) in a consistent fashion. However, OSONS is also just a precursor, and the utilization of the b-SHAP amalgam (e.g., temporal-centric FCSO values, SCSO values, and GSO values/VBSP values) is central. In turn, the b-SHAP amalgam needs to be leveraged in conjunction with a well-counterpoised MADM/MODM SM/OM paradigm. Wu, Wang and others have advocated for the use of the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) OM in conjunction with SHAP [25][26]. Meanwhile, Hua and others have advocated for the use of the PROMETHEE OM with SHAP (there is a dearth of research for SHAP with other OMs, such as ÉLimination Et Choix Traduisant la REalité or ELECTRE) [27]. The experimentation evaluated both of the former cases, and a finding, among others, is that of utilizing an OM (e.g., CRITIC) to first, derive the criteria weights and second, use a complementary pairing for the ensuing ranking (e.g., TOPSIS, PROMETHEE).

A. *Theoretical Foundations*

As described in the last paragraph of Section I, the issue of Non-deterministic Polynomial-time Hardness (NP-hard) problem spawning is problematic, such that *Spawn Reduction* becomes critical [8]. The involved optimization problem transformation pathways, such as those shown in Figure 1, strive to effectuate the non-convex to convex transmogrification.

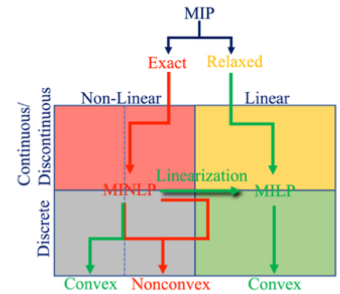

Figure 1. Non-convex to convex Transformation Pathways (e.g., non-convex discontinuous non-linear MINLPs to convex form)

A similar phenomenon is shown in Figure 2; after all, ML algorithms have a propensity to spawn "non-monotonic, non-polynomial, and even non-continuous (i.e., discontinuous) functions" [7]. Of note, the transformation of non-convex to convex can often inadvertently spawn further NP-hard problems. However, once in a convex form, a variety of Semi-Definite Programming (SDP) solvers can be employed to resolve the optimization problems in polynomial time [28].

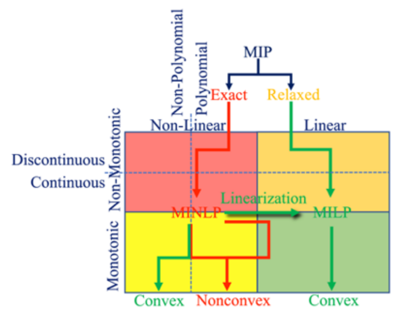

Figure 2. Non-convex to convex Transformation Pathways (e.g., non-convex [non-monotonic, discontinuous] non-polynomial MINLPs to convex form)

B. *Experimental Testbed*

Taking the case of NN, as depicted in Table II of Section II, the interpretability is in the orange (medium/low interpretability), as NN is more complex that DT and LR. However, for an enhanced STEA/construct validity-centric paradigm, a tasked ML can well learn atop the other MLs, adjust the involved ML model[s], and ascertain ways to mitigate against/lower the inadvertent spawning (i.e., *Spawn Reduction*). For this reason, the *testing* facet (at the constituent component level and interstices) of the *performance metrics* conjoined with *construct validity* considerations become central to the ML of ML task for the reduction of the spawning of further non-convex MINLP (e.g., from the transformation pathways of non-convex MINLP to convex MILP). In this case, the *testing* facet mechanisms and the utilized SDP solvers were implemented aboard GNU's Not Unix (GNU) Octave (a "numerical computation platform" that is "under the GNU [General Public License] (GPL) v3 license" and is generally "compatible with the likes of MATLAB") along with a

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

myriad of Octave Forge packages [28]. As noted in [28], "the source code was modified in the lab environment" so as to implement accelerants for the referenced SDP solvers to quickly address the various involved convex optimization problems described herein. Also, as noted in [28], "GPLv3 avoids the issue of tivoization (the instantiation of a system that incorporates software under the terms of a copyleft software license but leverages hardware restrictions or digital rights management to prevent users from running modified versions of the software on the involved hardware)" [28]. Testing was conducted using a variety of open-source software packages, such as Automatic Differentiation Model Builder (ADMB) (for non-linear statistical modeling) and Interior Point OPTimizer (IPOPT) (for large-scale nonlinear optimization) [28]; other promising software packages, such as LOQO (like IPOPT, it is based upon the interior-point method) and Sparse Nonlinear OPTimizer (SNOPT) (it leverages Sequential Quadratic Programming or SQP for resolving large-scale non-linear optimization problems) were examined, but they were not utilized given their licensing caveats.

It had been discussed in [8] that a particular numerical implementation of Continuous Wavelet Transforms (CWTs), aboard a CPRLD architectural paradigm, well contributes to STEA by way of the intrinsic "successive convolutional layers (which contain the cascading of ever smaller 'CWT-like' convolutional filters)" [8]. The referenced CPRLD construct handled the various transformation pathways delineated in Figures 1 and 2 (e.g., convex approximations, series of convex relaxations, etc.), and the architectural implementation for this paper was unique in that a ML of ML paradigm was implemented for *Spawn Reduction* (SR_2 on SR_1), such as shown in Figure 3.

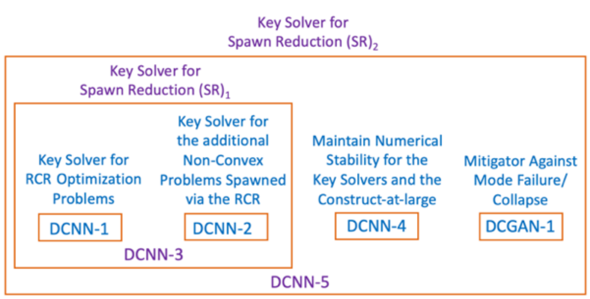

Figure 3. CPRLD Architectural Construct with a ML of ML (SR_2 on SR_1) Spawn Reduction paradigm

In terms of implementation details, a DCNN-centric instantiation was chosen for the requisite sufficient balance of reduced computational complexity along with sufficient robustness to be fit for purpose. The assigned tasks of the various DCNN are labeled accordingly in Figure 3. For example, as DCNN-1 was tasked with being the key solver for the involved convex optimization problems, it required a high degree of numerical stability, and PyTorch version 0.4.1 was selected; DCGAN-1 leveraged a "forward stable" TensorFlow-based DL Convolutional GAN (DCGAN) implementation to be able to well address the potentiality of mode collapse/mode failure (a phenomenon that may occur when adversarial GANs, which are being trained in tandem, are either unable to converge or undergo an anomalous convergence) [8].

C. *Experimental Construct*

With regards to the involved experimental construct, as can be seen in Figure 4, prior experimentation aspects used as presets are reflected in blue font while current experimental elements are shown in purple font. The "t-" elements (e.g., f-FCSO, t-SCSO, t-GSO, t-VBSO) of b-SHAP are extrapolations of Borgonovo's work (previously discussed in Section IID) that more fully consider Kotthoff's emphasis on temporal-sensitive/temporal-centric Shapley values [20]. STEA-related experimental forays for various OM were conducted. The OM of CRITIC was utilized as a preset for deriving the criteria weights, and the OMs of PROMETHEE, TOPSIS, and ELECTRE were utilized for the subsequent rankings. Initial selections and avoidances, among others, were based upon the following rationale. For example, PROMETHEE was known to be "easily… understood" and interpretable, so it was selected for testing [29][30]. Along this vein, [fuzzy] VIKOR was not selected, as it was known to be less interpretable and "less explainable than other more intuitive methods" [31].

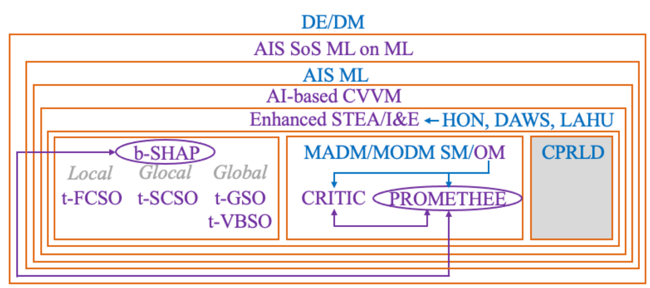

Figure 4. AI-based CVVM (ISA) Experimentation Aspects

Overall, selections were made to improve STEA/I&E. Yet, there were other technical considerations as well. A number of methodologies are subject to a phenomenon known as "Rank Reversal" (RR), wherein ranking results might change when the method changes or when the set of alternatives changes (leading to inconsistent and/or inaccurate results). The select OMs experimented with were known to be the most resistant to RR (yet are still subject to the phenomenon), and preliminary results are shown in Figure 5 below [32]. The key for the chart is as follows. First, the referenced "select OMs" of this Section IIIC are self-evident: ELECTRE, TOPSIS, and PROMETHEE. Second, these "select OMs" were benchmarked by execution time (E), sensitivity (S), performance under uncertainty (U), validity (V), and interpretability (I). Third, the aforementioned were benchmarked against classical SHAP (c-SHAP), as well as the b-SHAP approach described within this paper. Using the CPRLD as a preset,

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. ISBN: 978-1-68558-259-3

collectively, this forms the basis of the ISA described herein. The relative values were normalized against a scale of one to ten for ease of comparison.

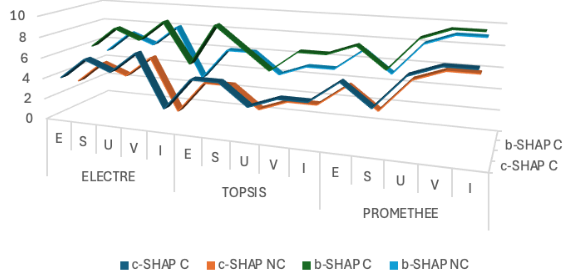

Figure 5. Preliminary Results from b-SHAP/select OM Benchmarking

The V and I were higher for PROMETHEE than for TOPSIS or ELECTRE. The E for TOPSIS was notably higher than that of the others, but the computational complexity is known to be less, and the performance under conditions of U was weaker than that of the others; the performance of PROMETHEE under conditions of U were seemingly better than ELECTRE and TOPSIS, in that order. Overall, the performance of b-SHAP was better than that of c-SHAP across the board for the range of E, S, U, V, I (for all the "select OMs" of ELECTRE, TOPSIS, and PROMETHEE). Hence, the b-SHAP-PROMETHEE amalgam (along with the CRITIC, CPRLD, etc. presets) exhibits promise.

IV. Conclusion

In consideration of Abraham Maslow's notion regarding the predilection that follows when there is only one tool to utilize, Section IIIC depicted some of the metrics underpinning the selection of a variety of methods and the comparative performance. For example, with regards to I&E, PROMETHEE was initially chosen over [fuzzy] VIKOR. As another example, PROMETHEE, TOPSIS, and ELECTRE were selected for testing, as they were reported to be more resistant to RR than certain other methods. As yet another example, Figure 5 depicted the relative performance of the methods for E, S, U, V, I; TOPSIS had a comparatively better E when E was considered in isolation, but it did not fare well under U, and along this vein, PROMETHEE did fare reasonably well under conditions of U when compared to ELECTRE and TOPSIS, etc. This brings us to the primary impetus of this paper, which centered upon enhancing robustness of the *testing* facet (with more granularity) at the interstices (e.g., *interstitial areas* between/among individual/amalgam component at the local, glocal, and global levels), better illuminating *I&E*/STEA DE/DM pathways, and operationalizing *AI-based CVVM* for the purposes of achieving higher efficacy AI SoS ML on ML for RWS. The hitherto lack of methodologies in this regard have led to RWS paradigms, wherein AIS adversely impact other AIS with the potentiality of cascading failure for the involved AI SoS (a.k.a., "near misses"). Moreover, the *testing* facet involves *performance metrics* conjoined with *construct validity* considerations. On the performance metrics front, OSONS was found to have greater efficacy than OSNS. Similarly, the b-SHAP (which involves various temporal-centric SHAP instantiations for local, glocal, and global) and PROMETHEE (along with CRITIC) amalgam was found to be more robust than the b-SHAP/TOPSIS or b-SHAP/ELECTRE amalgams on the *interpretability* front. Also on the performance front, *spawn reduction* turns out to be central, for once in the convex form, a myriad of SDF solvers can be leveraged to handle the involved optimization problems in polynomial time; otherwise, NP-hard spawn can congest matters with an indefinite impasse. The advancement of STEA/I&E necessarily involves HONs, DAWS, and LAHU, and these presets were discussed; the enhanced STEA/I&E discernment segues to more robust feature alignment, robust interpretation, etc., which constitutes enhanced *construct validity*. For this reason, it seems apropos to have the "Enhancement of an AI-based Construct Validity Approach" be the overarching descriptor of this paper. Future work will involve more quantitative and qualitative experimentation in the aforementioned areas.

References

[1] D. Subhadra, et al., "Rise of Artificial Intelligence in Business and Industry," *J. of Inform. Educ. and Res.*, pp. 869-875, May 2024.

[2] J. McKendrick, "The Many, Many Ways AI Spurs Innovation," *Forbes*, Dec 2023. [Online]. Accessed: Mar. 1, 2025. Available: https://www.forbes.com/sites/joemckendrick/2023/12/10/ai-spurs-innovation-heres-how/.

[3] "Industrial AI Insights," *Honeywell*, Jul. 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://www.honeywell.com/us/en/ai/research.

[4] D. Sjoberg and G. Bergersen, "Construct Validity in Software Engineering," *IEEE Trans. on Softw. Eng.*, vol. 49, pp. 1374-1396, Mar 2023.

[5] X. Zhou, Y Jin, H. Zhang, S. Li, and X. Huang, "A Map of Threats to Validity of Systematic Literature Reviews on Software Engineering," *23rd Asia-Pacific Softw. Eng. Conf. (APSEC)*, Apr 2017, pp. 153-160.

[6] S. Deets, C. Baulch, A. Obright, and D. Card, "Content Analysis, Construct Validity, and Artificial Intelligence: Implications for Technical and Professional Communication and Graduate Research Preparation," *J. of Bus. and Tech. Commun.*, vol. 38, pp. 303-315, Mar 2024.

[7] P. Hall, S. Ambati, and W. Phan, "Ideas on interpreting machine learning," *O'Reilly*, Mar. 1, 2017. [Online]. Accessed: Mar. 1, 2025. Available: https://www.oreilly.com/radar/ideas-on-interpreting-machine-learning/.

[8] S. Chan, "AI-Facilitated Dynamic Threshold-Tuning for a Maritime Domain Awareness Module," *IEEE Int. Conf. on Industry 4.0, Artif. Intell., and Commun. Technol. (IAICT)*, Aug. 2024, pp. 192-198.

[9] L. Grennan, A. Kremer, A. Singal, and P. Zipparo, "Why businesses need explainable AI – and how to deliver it," *McKinsey*, Sep. 2022. [Online]. Accessed: Mar. 1, 2025. Available:https://www.mckinsey.com/capabilities/quantumblack/our-insights/why-businesses-need-explainable-ai-and-how-to-deliver-it

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

[10] F. Scarelli, A. Tsoi, and M. Hagenbuchner, “The Vapnik-Chervonenkis dimension of graph and recursive neural networks,” *Neural Netw.*, vol. 108, pp. 248-259, Dec. 2018.

[11] G. Vilone and L. Longo, “Notions of explainability and evaluation approaches for explainable artificial intelligence,” *Inf. Fusion*, vol. 76, pp. 89-106, May 2021.

[12] S. Lundberg and S. Lee, “A Unified Approach to Interpreting Model Predictions,” *Proc. of the Adv. in Neural Inf. Proc. Syst.*, Dec. 2021, pp. 4765–4774.

[13] D. Balog, T. Bátyi, P. Csóka, and M. Pintér, “Properties and comparison of risk capital allocation methods,” *European J. of Oper. Res.*, vol. 259, pp. 614-625, 2017.

[14] M. Sundararajan and A. Najmi, “The Many Shapley Values for Model Explanation,” *Proc. of the 37th Int. Conf. on Mach. Learn.*, vol. 119, pp. 9269-9278, 2020.

[15] Y. Kwon, and J. Zou, “WeightedSHAP: analyzing and improving Shapley based feature attributions,” *Proc. of the 36th Int. Conf. on Neural Inf. Process. Syst.*, pp. 34363-34376, Oct 2022.

[16] L. Kotthoff, et al., “Quantifying Algorithmic Improvements Over Time,” *Proc. of the Twenty-Seventh Int. J. Conf. on Artif. Intell. (IJCAI-18)*,” pp. 5165-5171, Jul. 2018.

[17] R. Guidotti, et al., “A Survey of Methods for Explaining Black Box Models,” *ACM Comput. Surv. (CSUR)*, vol. 51, pp. 1-42, Aug. 2018.

[18] K. Dhamdhere, A. Agarwal, and M. Sundararajan, “The Shapley Taylor Interaction Index,” *Proc. of the 37th Int. Conf. on Mach. Learn.*, pp. 9259–9268, 2020.

[19] M. Grabisch and M. Roubens, “An axiomatic approach to the concept of interaction among players in cooperative games,” *Int. J. of Game Theor.*, vol. 28, pp. 547-565, Nov. 1999.

[20] E. Borgonovo, E. Plischke, and G. Rabitti, “The many Shapley values for explainable artificial intelligence: A sensitivity analysis perspective,” *European J. of Oper. Res.*, vol. 318, pp. 911-926, Nov. 2024.

[21] M. Mase, A. Owen, and B. Seilver, “Explaining black box decision by Shapley cohort refinement,” *Arxiv.org*, Oct. 2020. [Online]. Accessed: Mar. 1, 2025. Available: https://arxiv.org/pdf/1911.00467.

[22] A. Casjus, “The shapley value, the owen value, and the veil of ignorance,” *Int. Game Theor. Rev*., vol. 11, pp. 453-457, Dec. 2009.

[23] S. Lopez, “On the relationship between Shapley and Owen values,” *Central European J. of Oper. Res*., vol. 17, pp. 415-423, Dec. 2009.

[24] S. Beal, M. Diss, and R. Takeng, “New axiomatisations of the Diversity Owen and Shapley values,” *CRESE Working Paper 2024-09*, Feb. 2024. [Online]. Accessed: Mar. 1, 2025. Available: https://hal.science/hal-04502031v1/document.

[25] Z. Wu, “Evaluation of Provincial Economic Resilience in China Based on the TOPSIS-XGBoost-SHAP Model,” *J. of Math.*, pp. 1-12, Oct. 2023.

[26] X. Wang and Z. Piao, “An interpretable dynamic risk assessment approach in insurance: integrating TOPSIS, GM (1,1) and SHAP,” *Proc. of the 2024 Guangdong-Hong Kong-Macao Greater Bay Area Int. Conf. on Digit. Econ. and Artif. Intell. (DEAI)*, pp. 697-701, Jul. 2024.

[27] Z. Hua and Xiaochuan Jin, “A generalized Shapley index-based interval-valued Pythagorean fuzzy PROMETHEE method for group decision-making,” *Soft Comput.*, vol. 27, pp. 6629-6652, Feb. 023.

[28] S. Chan, “Mitigation Factors for Multi-domain Resilient Networked Distributed Tessellation Communications,” *Fifth Int. Conf. on Cyber-Technol. and Cyber-Syst. (CYBER 2020)*, Aug. 2020, p. 66-73.

[29] M. Brans and B. Mareschal, “Promethee Methods,” *Multiple Criteria Decision Analysis: State of the Art Surveys*, pp. 163-186, Jan. 2005.

[30] J. Brans and P. Vincke, “A Preference Ranking Organisation Method: (The PROMETHEE Method for Multiple Criteria Decision-Making), *Manage. Sci*., vol. 31, pp. 647-656, Jun. 1985.

[31] Y. Lin and T. Chen, “Type-II fuzzy approach with explainable artificial intelligence for nature-based leisure travel destination selection amid the COVID-19 pandemic,” *Digit. Health,* vol. 8., pp. 1-15, Jun. 2022.

[32] M. Aazadfallah, “A New Feature of Rank Reversal in Some of MADM Models,” *J. of Appl. Inf. Sci*., vol. 4, pp. 1-11, 2015.

Courtesy of IARIA Board and IARIA Press. Original source: ThinkMind Digital Library https://www.thinkmind.org

Copyright (c) IARIA, 2025. **ISBN: 978-1-68558-259-3**