

Article

Not peer-reviewed version

---

# Great Minds Think Alike Entropy Artificial Intelligence and Knowledge Transfer

---

[Dr Ismail A Mageed](#) \*

Posted Date: 9 February 2024

doi: 10.20944/preprints202402.0563.v1

Keywords: Artificial intelligence (AI); Knowledge transfer (KT); intergenerational knowledge transfer (IGT); knowledge entropy (KE); Machine Learning (ML)



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# Great Minds Think Alike Entropy Artificial Intelligence and Knowledge Transfer

Ismail A Mageed, PhD, UK President of ISFSEA

drismail664@gmail.com

**Abstract:** An exposition of entropy applications to highlight some potential entropic applications to artificial intelligence and knowledge transfer. Entropy is a basic idea in artificial intelligence that is applied to many different tasks, such as reinforcement learning, data compression, and decision-making. By quantifying uncertainty and information content, it helps artificial intelligence models provide predictions and decisions that are well-informed. Therefore, this paper shines on spotlighting the significance of entropy to attract the attention of artificial intelligence research community to entropy as a powerful tool to advance artificial intelligence. The significance of knowledge transfer (KT), particularly intergenerational knowledge transfer (IGT), in knowledge management is also covered in this paper. To quantify the complexity of knowledge distribution inside an organisation and assess the efficacy of knowledge transfer (KT) initiatives, the notion of knowledge entropy (KE) is presented. Additionally, the KT model—which is based on the conceptions of tacitness and information content—is introduced. It combines personalisation and codification methodologies. Future research directions are offered alongside a few difficult open problems.

**Keywords:** Artificial Intelligence(AI); Knowledge Transfer (KT); Intergenerational Knowledge Transfer (IGT); Knowledge Entropy (KE); Machine Learning(ML)

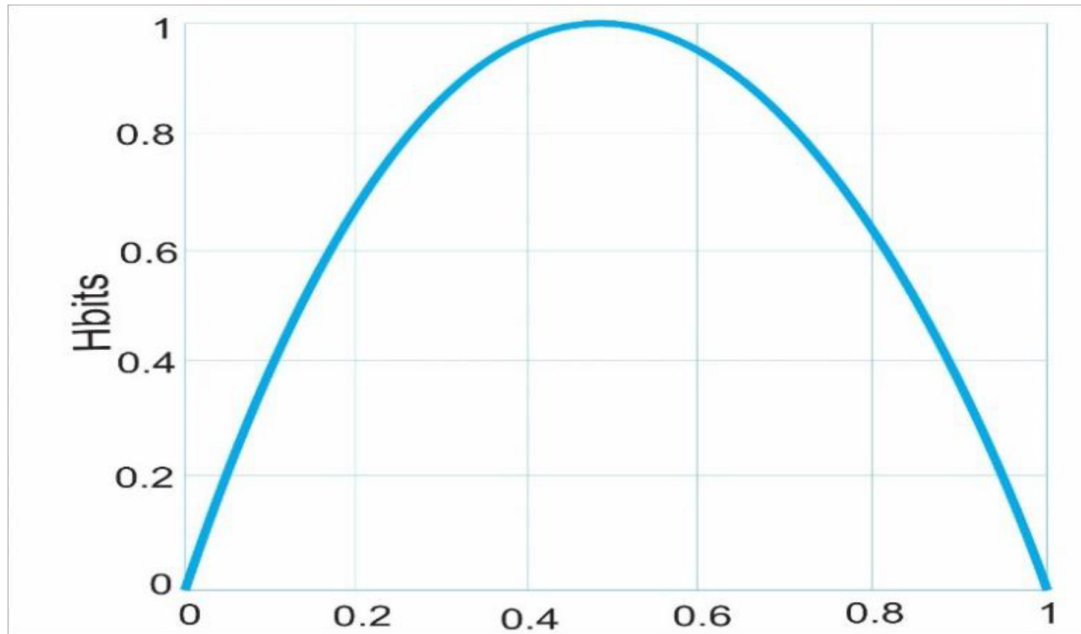
## 1. Introduction

Within the field of information theory, entropy [1] is a notion that has become useful in several applications, including secure data transmission, lossless data compression, data mining, machine learning (ML), and classification. In [1], information and entropy in a fair and fraudulent coin flip game are calculated using Shannon's entropy formula, a key measurement tool in information theory. The findings demonstrate that entropy quantifies system uncertainty and that, by using fewer bits, its application has enhanced efficient communication and data storage. Furthermore, the entropy formula helps to clarify information measurement, quality, and relevance by calculating the minimal number of bits needed to encode characters in a data group.

A mathematical concept, called entropy [1] is used to quantify the degree of disorder or uncertainty in every data source. It is computed using the entropy formula, which considers both the total amount of data and the probability that each piece of data is important and is denoted by the letter  $H$  in information theory. Entropy reaches its maximum when the probability of two states is equal, signifying the highest degree of uncertainty: as likelihood increases, uncertainty decreases.  $H$  reads as:

$$H = \sum_{i=1}^n p_i \ln(p_i) \quad (1)$$

Entropy, as defined in information theory, is highest when the probabilities of two states are equal, meaning  $p = 0.5$ . This indicates the highest level of uncertainty or disorder in the system. As the likelihood of a situation increases, the level of uncertainty decreases, leading to a reduction in entropy. See Figure 1(c.f., [1]).



**Figure 1.** The  $H$ -bit data amount entropic variation illustrates how uncertainty peaks when the probability of two states is equal ( $p = 0.5$ ). As the likelihood of a situation increases, uncertainty starts to diminish. This example helps demonstrate the concept of entropy, which is used to measure disorder or uncertainty in information theory.

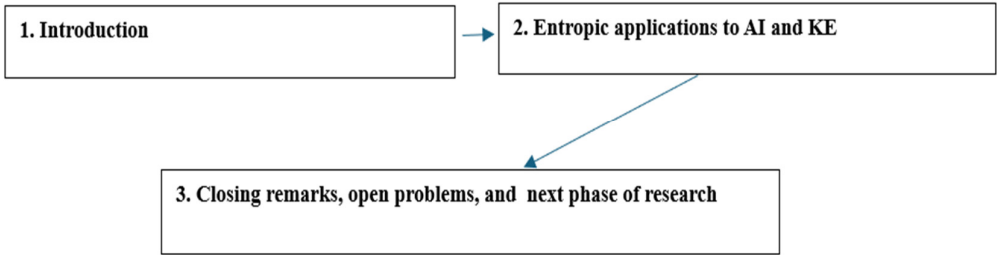
In the context of information criterion approaches, entropy plays a unique role, and one commonly used type of entropy in artificial intelligence is Cross-Entropy. When utilising an improper encoding scheme, Cross-Entropy is utilised in a variety of AI techniques to determine the average amount of bits needed to encode an event [1,2].

Knowledge transfer (KT) is widely regarded as a crucial phase in knowledge management. IGT (intergenerational knowledge transfer) is one of these essential forms of KT [5–9]. More thought should be given to the fact that the KT process entails complex relationships between the three fundamental types of knowledge—rational, emotional, and spiritual knowledge [10–13]. The notion of knowledge entropy (KE) is introduced in this text, which explains how it can be applied to information management and understanding knowledge transfer. KE is a measure of complexity that can be related to thermodynamic phenomena, providing insights into the intricacies of knowledge distribution and transfer within an organization. We define KE [13] as follows:

$$KE = -C \sum_i p_i \ln p_i \quad (2)$$

Here, a random positive constant named  $C$  is used to calibrate the measurement to a specific scale and environment. If every employee either possesses the information or serves as a source of it, we might assume that  $p_1, \dots, p_n$  represents the distribution of knowledge inside any organisation.

The flowchart of this paper is as follows:



2. Entropic applications to AI and KE

The current section investigates some entropic applications tom AI and KE.

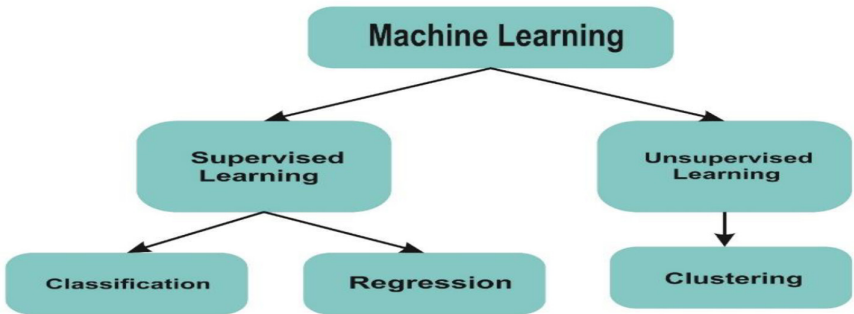
2.1. Entropy AI

In AI systems [1], the concept of entropy is used to measure uncertainty and information disorder. Entropy plays a crucial role in various techniques [1], such as Bayesian inference and maximum entropy, which aim to find the best distribution or model for a given dataset. These approaches [1] utilize conditional probability and information disorder to compare occurrences and optimize the distribution’s entropy.

Decision trees [1] are a classification technique that uses entropy to construct a tree structure. Information Gain (IG) is a measure of the difference in entropy between different states and is used in decision tree algorithms like ID3 and C4.5. Loss functions [1], such as Mean Squared Error, are used in classification techniques like Artificial Neural Networks to evaluate the performance of the model by measuring the average irregularity of the predicted outcomes compared to the actual ones. Logistic Regression [1] is a classification method that calculates the logarithm of the ratio between the occurrence and non-occurrence of an event, using categorical dependent variables.

In the field of AI [1], when working with datasets to address problems, it is important to utilize information from previous attempts or problems. Depending on the structure and content of the dataset, various preprocessing techniques are applied to ensure the data is suitable for analysis and modelling to achieve effective AI solutions.

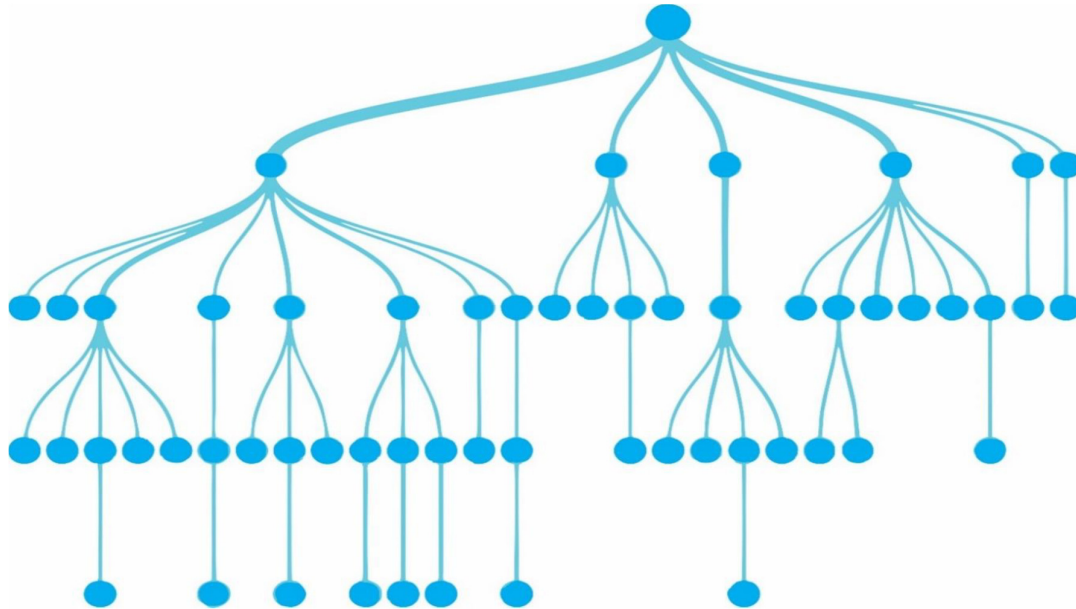
Figure 2 (c.f., [1]), illustrates the two main categories of supervised and unsupervised learning in the field of Artificial Intelligence, specifically in the context of ML. From these categories, three subcategories are derived: classification, clustering, and regression. These subcategories represent different approaches used to address various challenges in AI.



**Figure 2.** The ML categorical structure in the context of supervised learning involves training a system to categorize data based on specific outcomes. The model is generated using provided data and parameters, and subsequent questions or challenges are addressed using the knowledge gained

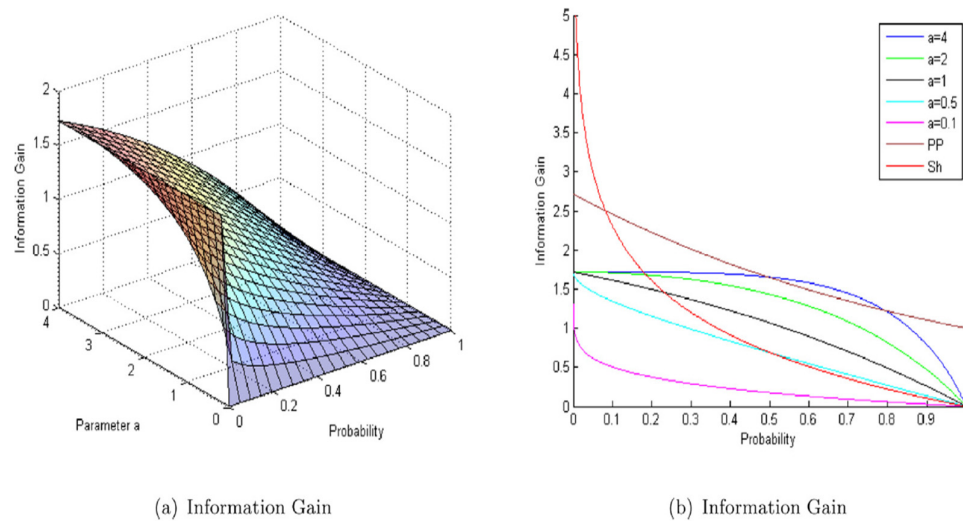
during training. This structure is one of the fundamental categories in artificial intelligence approaches, particularly in ML.

Decision trees [1] are a commonly used technique for grouping and estimation problems. During the training phase, these algorithms follow a top-down or general-to-specific approach, creating branches based on the attribute values of each node in the tree structure. This structure resembles a flowchart and helps in making decisions or predictions based on the input data. See Figure 3(c.f., [1]).



**Figure 3.** The example decision tree is a visual representation of a classification model that helps make decisions based on certain attributes. It shows how the model divides the data based on different criteria, such as age, foreign language proficiency, and salary expectation, to determine the hiring status of potential employees. By following the branches of the decision tree, one can assess whether an individual is likely to be hired or not based on their qualifications.

In [3], the concept of information gain in decision-making has been thoroughly explained, where it has been stated that the degree of uncertainty reduction depends on both the probability of an outcome and the decision-maker's attitude towards risk. The proposed information [3] gain function, characterized by probability ( $p_i$ ) and the agent's conservatism level ( $a$ ), shows that as  $p_i$  increases, the information gain decreases, with the rate of decrease determined by  $a$ , as shown by Figure 4 (c.f., [3]).



**Figure 4.** The function shows that as  $a$  decreases, the information gain decreases, indicating that the function is more sensitive to lower probability values for a more adventurous decision-maker. So, In the real world [3], individuals have their own unique attitudes, and the proposed function in this context represents a decision maker's attitude through a parameter called  $a$ .

## 2.2. Entropy KE

One can conceptualise that distribution both in terms of time and space. However, as every individual is different from the next and has their own experiences, feelings, and spirituality, this kind of situation is not possible. This suggests that it is not possible to quantify absolute knowledge for everyone using a knowledge management metric. Instead, relative values of knowledge can be used in relation to a specific degree of knowledge. This approach acknowledges that individuals have different experiences and perspectives, making it more practical to assess knowledge in a relative sense rather than an absolute one.

Here, a random positive constant named  $C$  is used to calibrate the measurement to a specific scale and environment. If every employee either possesses the information or serves as a source of it, we might assume that  $p_1, \dots, p_n$  represents the distribution of knowledge inside any organisation. That distribution can be thought of in terms of both space and time. However, as every individual is different from the next and has their own experiences, feelings, and spirituality, this kind of situation is not possible. The concept of Knowledge Entropy (KE) can be valuable in evaluating the effectiveness of Knowledge Transfer (KT) in training programs or information-sharing activities within businesses or communities of practice. By assessing the distribution of information within a department or organization, managers can determine how to enhance knowledge entropy, which promotes innovation. Researchers can also utilize knowledge entropy to develop intellectual capital and establish intelligent organizations, as it provides insights into the expected distribution of knowledge within a specific company at a given time [11]. By assessing the information distribution inside a certain department or organisation, managers can determine how best to increase knowledge entropy, which fosters innovation. By measuring knowledge entropy, researchers may also show how to develop intellectual capital and create intelligent organisations [12,13].

As stated in [9], KE illustrates how knowledge is expected to be distributed within a specific company at a given time. Although we think of organisational knowledge as being like a field, it resides with particular people and produces the distribution of individual knowledge for a given amount of time. KE and information entropy are similar from the standpoint of mathematical modelling, but they are entirely distinct from a semantic one.

Due to the current study trend's exclusive focus on KE and lack of empirical research, KE has become limited. When examining the knowledge probability distribution function theoretically, it



can be challenging to interpret. However, there are useful methods available to gather relevant data and calculate the knowledge entropy indicator for different scenarios and time points. Future research should focus on developing practical approaches to compute knowledge entropy and probability sets, as well as conducting empirical studies to gain a better understanding of how knowledge affects the performance of organizations.

In [18], the subject of assessing recognition knowledge—specifically, categorization knowledge—and its evolution was covered. In the given context, the discussion is based on three principles. A model was proposed to understand how knowledge affects uncertainty, leading to the creation of two formulas for evaluating recognition knowledge levels in different scenarios. Additionally, the concept of knowledge entropy was introduced, and its formula was derived by examining how ignorance changes when there is uncertainty. We examined how it differed from Shannon's entropy and how similar it was to Boltzmann's entropy. A mathematical analysis revealed evidence to support the following conclusions:

- As a result of learning, knowledge entropy decreases.
- As knowledge entropy decreases, the people's rating order becomes more distinctive.
- The total knowledge level of a group's members does not always equal the group's collective knowledge level.
- A person's knowledge entropy will never rise if their thirst for information never grows.

Personalization and codification are the two-knowledge transfer (KT) strategies that are most frequently employed in companies and organisational networks [1]. A theoretical model of Knowledge transfer (KT) has been presented by [18] to assess how organisations (KT) convey tacit knowledge, or knowledge gained without going through the experience, and the associated information content. In the context of knowledge transfer strategies, Shannon's entropy from information theory is used to explain the concepts of tacitness (implicit knowledge) and information content, and how they influence the selection of knowledge transfer approaches. This utilization of Shannon's entropy helps in understanding and making informed decisions about the most suitable methods for conveying tacit knowledge within organizations. Specifically, [18] has helped with:

- Making predictions about the KT mechanism selection based on information content.
- The creation of a tacitness expression and an intuitive justification for the tacit-explicit continuum.
- Creating a theoretical KT model that may be used to predict which KT mechanisms will be used in real-world situations and characterizing the information content of different product varieties.

The KT model [18], which combines personalization and codification techniques, is shown in Figure 5. It is evident that tacitness and information content volume are two crucial characteristics that influence the KT process. When individuals and businesses operate across longer distances, individualized interactions become more expensive, and standardization approaches become the norm. However, there are three main barriers that could prevent individualization from being applied. Standardization tactics take hold when individuals and organizations operate in larger geographic regions and individualized interactions become more expensive.

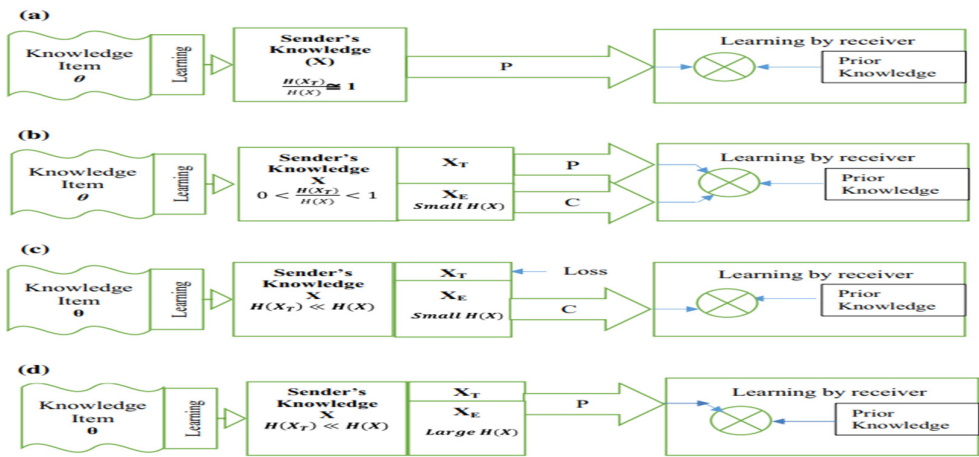


Figure 5. The KT model schematically illustrated [18].

On a different note, KT [19] refers to the process of exchanging knowledge within an organization. It is crucial for improving the organization's knowledge stock and competence. The capacity of knowledge senders to effectively disseminate knowledge and the absorptive capacity of knowledge recipients significantly impact the success of knowledge transfer. Additionally, individuals' social benefits, such as status and respect, play a role in stimulating knowledge senders and recipients, ultimately enhancing the dissemination of knowledge, and accelerating the transfer process within the organization.

The authors [19] constructed a KT conceptual model, which considers the relationships between knowledge senders, knowledge recipients, their disseminative capacity, absorptive capacity, and benefits. This model aims [19] to understand how knowledge is exchanged within an organization and how factors like disseminative capacity, absorptive capacity, and benefits influence this process. The model provides a framework for analysing and improving knowledge transfer within organizations. Figure 6 provides a schematic for this model (c.f., [19]).

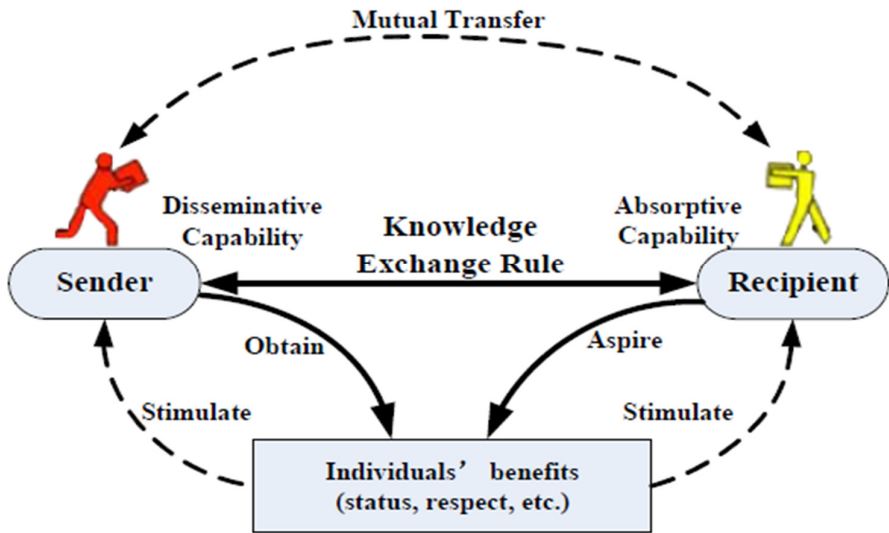


Figure 6. Organizational KT model.

In the given conceptual model of KT [19], individuals within an organization can act as both knowledge senders and recipients. The knowledge disseminative capacity (D) represents the ability of a sender to articulate and teach knowledge [19], while the knowledge absorptive capacity (A) refers



to the recipient's ability to identify, learn, and apply knowledge from senders. Additionally [19], the model considers the influence of social benefits, such as status and respect, which can motivate senders to transfer knowledge and recipients to absorb and become senders themselves, leading to a continuous knowledge spiral process that enhances the overall knowledge level of the organization.

### 3. Closing remarks, open problems, and next phase of research

This paper discusses the applications of entropy in artificial intelligence and knowledge transfer. Entropy, a concept used in various AI tasks like reinforcement learning, data compression, and decision-making, helps quantify uncertainty and information content. The paper emphasizes the significance of entropy as a powerful tool to advance AI and covers the importance of knowledge transfer, particularly intergenerational knowledge transfer, in knowledge management. It introduces the notion of knowledge entropy to assess the complexity of knowledge distribution and presents a knowledge transfer model based on tacitness and information content.

Some new, unresolved issues are suggested in the current paper:

- Can we overcome the three main obstacles that could prevent personalization from being implemented by using Ismail's entropies in place of Shannonian entropic formulas (c.f., [20–23])?
- If the  $p_i$ 's (c.f., Equation (2)) are dependent on time, how would this affect the investigation's level of complexity? It's still up for debate.

The next phase of research includes investigations to solve the proposed open problem as well as widening the search for more entropic applications in multi-interdisciplinary fields of human knowledge.

### References

1. [1] Ali, A., Anam, S., & Ahmed, M. M. (2023). Shannon Entropy in Artificial Intelligence and Its Applications Based on Information Theory. *Journal of Applied and Emerging Sciences*, 13(1), 09-17.
2. [2] Barron, J. T. (2019). A general and adaptive robust loss function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4331-4339).
3. [3] Aggarwal, M. (2021). Attitude-based entropy function and applications in decision-making. *Engineering Applications of Artificial Intelligence*, 104, 104290.
4. [4] Ali, A.; Naeem, S.; Anam, S.; Zubair, M. Agile Software Development Processes Implementing Issues and Challenges with Scrum, in *Proceedings of the MOL2NET'22, Conference on Molecular, Biomedical & Computational Sciences and Engineering*, 8th ed., 1–31 January 2023, MDPI: Basel, Switzerland, doi:10.3390/mol2net-08-13907.
5. [5] Nguyen, T. M., Siri, N. S., & Malik, A. (2021). Multilevel influences on individual knowledge sharing behaviours: the moderating effects of knowledge sharing opportunity and collectivism. *Journal of Knowledge Management*, 26(1), 70-87.
6. [6] Liu, S. (2020). *Knowledge management: an interdisciplinary approach for business decisions*. Kogan Page Publishers.
7. [7] Abdillah, A., Widianingsih, I., Buchari, R. A., & Nurasa, H. (2023). The knowledge-creating company: How Japanese companies create the dynamics of innovation: by Nonaka, Ikujiro & Takeuchi, Hirotaka, New York, Oxford University Press, 1995, 284 pp., 19.39(Hardcover)& 7.40 (paperback), ISBN: 0199879923, 9780199879922.
8. [8] Rupčić, N. (2021). Interorganizational learning: a context-dependent process. *The Learning Organization*, 28(2), 222-232.
9. [9] Kyriakopoulos, G. L. (2021). Globalized inclination to acquire knowledge and skills toward economic development. *WSEAS Trans. Bus. Econ*, 18, 1349-1369.
10. [10] Bratianu, C.(2018). "Intellectual capital research and practice: 7 myths and one golden rule. *Management & Marketing*." *Challenges for the Knowledge Society* 13(2), 859-879.
11. [11] Bratianu, C. and R.Bejinaru. (2019). "The theory of knowledge fields: a thermodynamic approach." *Systems* 7(2), 20, 1-12.
12. [12] Bratianu, C. (2019). "Exploring knowledge entropy in organizations." *Management Dynamics in the Knowledge Economy* 7(3):353-66.
13. [13] Bratianu, C. and R.Bejinaru. (2020). " Knowledge dynamics: a thermodynamic approach.", *Kybernetes* 49(1), 6-21.

14. [14] A Mageed, D.I. Maximum Ismail's Second Entropy Formalism of Heavy-Tailed Queues with Hurst Exponent Heuristic Mean Queue Length Combined with Potential Applications of Hurst Exponent to Social Computing and Connected Health. Preprints 2024, 2024012055. <https://doi.org/10.20944/preprints202401.2055.v1>
15. [15] Lentjušenkova, O., & Lapiņa, I. (2020). An integrated process-based approach to intellectual capital management. *Business Process Management Journal*, 26(7), 1833-1850.
16. [16] Wenger, E. (1998). "Communities of practice: learning, meaning, and identity." Cambridge, UK: Cambridge University Press.
17. [17] Hou, F. (2018). "Measuring knowledge for recognition and knowledge entropy.", arXiv preprint arXiv:1811.06135.
18. [18] Mageed, I. A. (2022). Extended Entropy Maximisation and Queueing Systems with Heavy-Tailed Distributions (Doctoral dissertation, University of Bradford). <https://bradscholars.brad.ac.uk/handle/10454/19756>
19. [19] Liu, N., Wu, J., & Xuan, Z. 2015. Knowledge Transfer within the Organization for the Consideration of Disseminative Capacity, Absorptive Capacity and Individuals' Benefits.
20. [20] I. A. Mageed and Q. Zhang, "An Information Theoretic Unified Global Theory For a Stable M/G/1 Queue With Potential Maximum Entropy Applications to Energy Works," *2022 Global Energy Conference (GEC)*, Batman, Turkey, 2022, pp. 300-305, doi: 10.1109/GEC55014.2022.9986719.
21. [21] Mageed, I.A, Fractal Dimension of Ismail's Third Entropy with Fractal Applications to CubeSat Technologies and Education, The 2nd International Conference on Applied Mathematics, Informatics, and Computing Sciences (AMICS 2023), Ghent University, Belgium.
22. [22] I. A. Mageed, "Fractal Dimension( $D_f$ ) Theory of Ismail's Second Entropy( $H_1^q$ ) with Potential Fractal Applications to ChatGPT, Distributed Ledger Technologies(DLTs) and Image Processing(IP)," *2023 International Conference on Computer and Applications (ICCA)*, Cairo, Egypt, 2023, pp. 1-6, doi: 10.1109/ICCA59364.2023.10401817.
23. [23] I. A. Mageed, "Fractal Dimension ( $D_f$ ) of Ismail's Fourth Entropy  $H_{IV}^{(q,a_1,a_2,\dots,a_k)}$  with Fractal Applications to Algorithms, Haptics, and Transportation," *2023 International Conference on Computer and Applications (ICCA)*, Cairo, Egypt, 2023, pp. 1-6, doi: 10.1109/ICCA59364.2023.10401780.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.