ERCIM
European Research Consortium
for Informatics and Mathematics

Connected Excellence in Research

Strategy Report

ERCIM/JST Joint Workshop 2024

Theme 2:
Extracting Actionable Knowledge in the Presence of Uncertainty

November 2025

JST/ERCIM Joint Workshop
Budapest, October 17-18, 2024

Working group #2: Extracting Actionable Knowledge in the Presence of Uncertainty

Co-ordinators:
Akira Uchiyama (Osaka University, Japan)
Nicolas Spyratos (University Paris-Saclay, France)

Contributors:
Dominique Laurent (CY Cergy Paris University, France)
Nicolas Spyratos (University Paris-Saclay, France)
George Tzagarakis (FORTH ICS, Greece)
Csaji Balasz Csanad (SZTAKI Institute for Computer Science, Hungary)
Noémi Friedman (SZTAKI Institute for Computer Science, Hungary)
Akira Uchiyama (Osaka University, Japan)
Takuya Yoshihiro (Wakayama University, Japan)
Yorie Nakahira (Carnegie Mellon University, Japan)
Masako Kishida (National Institute of Informatics, Japan)
Savong Bou (University of Tsukuba, Japan)

Summary:
The main objective of this group is to study the concept of uncertainty and its impact on actionable knowledge extraction. Our approach is to consider uncertainty in the context of data pipelines whose importance and complexity are greatly increasing nowadays. Good service provision by a data pipeline depends to a large extent on data quality, and uncertainty is one of the main factors influencing data quality. In this paper we focus on two important questions regarding uncertainty:

- What are the sources of uncertainty?
- How can uncertainty be managed?

We attempt to answer these questions in four important areas of research, namely databases, signal processing, control theory and AI.

1. Introduction

Good service provision by a data pipeline depends to a large extent on data quality. The importance of data quality has been recognized beyond the field of data engineering and database management systems. For example, in signal processing and AI applications, high data quality standards are crucial to ensure robust predictive performance and responsible usage of automated decision making [1]. Uncertainty is one of the main factors influencing data quality. Perhaps the simplest way to understand the concept of uncertainty is to consider how population surveys are performed. Indeed, we are often interested in the characteristics of a population of "objects", but usually we survey only a sample of the population rather than every object. This is timelier and more cost-effective and, if the sample is large enough and well designed, can lead to accurate statistics. Using a sample means that our statistics are usually accompanied by measures of uncertainty. Uncertainty relates to how the estimate might differ from the "true value" and these measures help users to understand the degree of confidence in the outputs. The measures of uncertainty include standard error, confidence interval, coefficient of variation and statistical significance [7].

Uncertainty may appear in each stage of a data pipeline, namely during data collection at the sources; during integration and storage of the collected data following the rules of a data model; and during knowledge extraction from the stored data (see also Figure 1).

Most systems today, and in particular signal processing systems, AI systems and database management systems, often operate in environments where uncertainty is a fundamental aspect. Representing and reasoning about knowledge in such uncertain domains is crucial for building robust and intelligent systems [8, 11].

An uncertain domain refers to a field or environment where the information available is incomplete, ambiguous, noisy, or inherently unpredictable. Unlike deterministic domains where outcomes can be predicted with certainty given the inputs, uncertain domains require systems to handle and reason about uncertainty in a structured manner. The main characteristics of uncertain domains are the following:

Incomplete Information: The system does not have access to all the data required to make a fully informed decision.

Ambiguity: Information might be unclear or open to multiple interpretations.

Noise: Data might be corrupted or imprecise due to measurement errors or external factors.

Stochastic Processes: The environment might involve random processes or events.

In this paper we focus on three areas of research, namely database systems, signal processing and AI, and we discuss two important questions regarding uncertainty in these areas:

- What are the sources of uncertainty during data collection, data integration and knowledge extraction?
- How can uncertainty be managed so as to extract actionable knowledge?

2. Database systems

Uncertainty in database systems has several causes, the most common of which are the following:

Uncertainty at the data sources

During data collection at the sources, data values may be erroneous or simply missing, due to incomplete data entry, equipment malfunctions, lost files, and many other reasons. In any dataset, there are usually some missing data. Incomplete datasets can break data pipelines and have devastating impact on downstream results when not detected. While a variety of approaches to impute missing values exist [13] comprehensive benchmarks comparing classical and modern imputation approaches under fair and realistic conditions are underrepresented [1].

Uncertainty at the data model

Integration of data collected at the sources follows a data model and the dominant model for databases today is the relational model. Following this model, a database consists of a set of tables where data values are inserted. The most common sources of uncertainty at model level are the following:
(a) *Missing values*, where it is not possible to record a data value in a table. For example, in a table containing the attributes Employee, Salary and Maiden name, it is impossible to record the maiden name of a male employee (because a male employee has no maiden name). As a result, there is an empty cell in the table causing uncertainty during data processing. This is an example of missing value of the type "value does not exist". However, there are other types of missing values as well, for example "value exists but is currently unknown" or "there is no information as to the existence or non-existence of the value". The presence of any of these various types of missing values in the database causes uncertainty in data processing.

(b) *Integrity constraint violations*, when the same key-value in a table is associated with two or more values in non-key columns of the table. For example, consider a table T(Emp, Mgr) in which Emp is the key. In this example, there is a key dependency Emp→Mgr. If we have two tuples, (e, m) and (e, m') in the table with $m \neq m'$ then the table violates the key and this will cause uncertainty as to the semantics of the table.

Uncertainty during knowledge extraction

The main tool for knowledge extraction in a database management system is the query language. Knowledge is extracted from the database in the form of answers to queries. In order for this knowledge to be actionable, the query answers must contain no uncertainty. For example, consider a relational database with three tables, T1(<u>Emp</u>, Dep), T2(<u>Dep</u>, Mgr) and T3(<u>Emp</u>, Mgr), where the underlined attributes are the keys. Now, to find all employees and their managers one might ask the following relational algebra query: $Q= \pi_{Emp,\ Mgr}(T1 \bowtie T2) \cup T3$, where "$\bowtie$" stands for the join operation of the relational model [37]. The result of this query may violate the key dependency Emp→Mgr. This is due to an ambiguity of the relational model: key-dependencies are declared over the set of all attributes appearing in the database while database consistency is verified at table level. In other words, local satisfaction of integrity constraints at table level does not imply satisfaction of the integrity constraints by the database as a whole. In our example, it is not clear whether an employee's manager coincide with the manager of the employee's department. A data table violating semantic constraints can break data pipelines and cause important damage in critical applications. Designing algorithms that "repair" such tables before their use is a hot research topic in the area of databases [2]. An approach to the problem of repairing tables with missing values *and* integrity constraint violations can be found in [3, 36], while a unifying framework for uncertainty, inconsistency and incompleteness is proposed in [40].

3. Signal processing

Signal processing [22, 23] retrieves knowledge from spatial and/or time-series data that involves modifying, synthesizing, and analyzing signals, such as seismic signals, altimetry processing, images, sound, potential fields, and scientific measurements. Signal processing is useful in various optimizations and recognitions, such as the detection of motifs [14], storage efficiency [15], transmission acceleration [16], distortion correction [17], object tracking [18], medical applications [19], sound/video processing [20], etc. However, the performance can be greatly impacted if there are high uncertainties [21]. Real-time systems, such as audio, video, sensor streams, and communication devices, can perform poorly if there are errors or information shortages in signal processing. Handling signal processing errors is crucial in improving the reliability and efficiency of many applications.

Sources of Uncertainty

There are various causes contributing to the uncertainty in signal processing, including data shortage, measurement or environmental error, coding, processing algorithms, editing, model reliability, or tabulating data. Some common sources of uncertainty in signal processing results can be caused by various types of noises or limitations of processing methods:

- Environmental errors: acoustic or optical reflections, movements of objects, etc.
- Measurement errors: electromagnetic interference, thermal noise, clock jitter, etc.
- Quantization: finite resolution or precision of the system, etc.
- Data shortage: lacking key data, information shortage in data, etc.
- Model capacity: limitation of model or algorithm ability, insufficient parameter optimization, over-simplification of models, etc.
- Computing resources: model limitations caused by computing resources, etc.

One of the major causes of uncertainty in signal processing systems comes from numerical noises. Numerical errors can be in different forms. Some common types of errors include:

- Random errors: Such as white noise or Gaussian noise, which have no correlation, no pattern, and vary unpredictably.
- Systematic errors: Such as harmonic distortion or offset errors with a consistent pattern or bias.

- Transient errors: Such as impulsive noise or glitches, which occurs sporadically or temporarily due to sudden changes in the system or environment.

Uncertainty management

Uncertainty management [23] is a key technology to improve and guarantee the quality of signal processing systems. The management methodologies can be classified as follows.

- Noise management: suitable preprocessing and filtering, stochastic error estimation, inconsistency recognition, super-resolution, etc.
- Multi-modal processing: required data identification, multi-modal data synthesis, sensor fusion, etc.
- Model improvement: improving algorithms, stochastic or deep-learning models, data encoding, parameter optimization, etc.
- Accuracy analysis: uncertainty/risk quantification in the application area, feedback from reality, etc.

Uncertainty quantification

Uncertainty quantification [35] refers to the process of identifying, characterizing, and measuring uncertainty in signal processing systems. It aims to model the sources of uncertainty and represent them mathematically. The key techniques for uncertainty quantification are categorized based on statistical, probabilistic, and learning-based methods, including the following:

- *Probabilistic & Bayesian Methods*: These approaches model uncertainty explicitly using probability distributions; e.g. Bayesian Inference, Bayesian Model Averaging, Gaussian Processes, Monte Carlo Methods (MC, MCMC, SMC), and Hidden Markov Models & Variants.
- *Surrogate Modeling & Polynomial Methods*: These methods approximate system behavior to reduce computational costs while quantifying uncertainty; e.g. Polynomial Chaos Expansion, Stochastic Collocation, Kriging (Gaussian Process Regression), Radial Basis Function Surrogates, and Sparse Grid Methods.
- *Interval, Set, & Evidence-Based Methods*: These methods quantify uncertainty when probability distributions are unknown or hard to estimate; e.g. Interval Analysis, Dempster-Shafer Theory, Fuzzy Set Theory, and Convex Modeling & Robust Optimization.
- *Information-Theoretic & Entropy-Based Methods*: These methods quantify uncertainty using information measures; e.g. Shannon Entropy & Mutual Information, Fisher Information, Kullback-Leibler Divergence, and Optimal Transport for Uncertainty Quantification.
- *Machine Learning & Data-Driven Methods*: Used for complex, multi-modal data, these methods leverage deep learning and AI for uncertainty quantification; e.g. Deep Ensemble Methods, Bayesian Neural Networks, Variational Inference, and Adversarial Training for Uncertainty Quantification.

- *Hybrid & Domain-Specific Techniques*: These techniques combine different strategies for robust uncertainty handling; e.g. Kalman Filter & Variants, Wavelet-Based Uncertainty Analysis, and Surrogate Models & Reduced Order Modeling.

3. Control Theory

Control theory provides a mathematical framework for regulating dynamical systems, ensuring stability, performance, and robustness in applications ranging from robotics and aerospace to industrial automation and economic systems [24]. It helps extract actionable knowledge from systems affected by uncertainties in modeling, measurements, and external disturbances, enabling reliable decision-making. Techniques such as robust, adaptive, and stochastic control allow systems to function effectively even with incomplete or imprecise information. As systems face new challenges – growing complexity [25], AI integration [26], and stricter safety requirements [27] – control theory continues to evolve to tackle them while ensuring reliable performance.

Sources of uncertainties

Control systems face three main types of uncertainties:

- Model uncertainties: Occur when complex systems are simplified into mathematical models, leading to missing dynamics or inaccurate approximations
- External uncertainties: Include disturbances from environmental factors, changing conditions, system parameter variations, or communication delays
- Measurement uncertainties: Result from sensor noise, incomplete data, quantization errors, or sensor drift

As systems become larger and more complex, and with the increasing integration of AI, additional uncertainty factors are becoming increasingly significant:

- Large-scale systems: Complex interactions and computational challenges in interconnected systems
- AI integration: Errors in training data, model limitations, and unpredictable AI behavior
- Human interaction: Unpredictability from human behavior in semi-autonomous systems

These uncertainties often interact, amplifying their effects and making it more challenging to extract actionable knowledge for reliable control strategies, especially in safety-critical applications where formal guarantees are required.

Uncertainty management

To handle these uncertainties, several approaches have been developed, including, but not limited to:

- Robust control: Ensures stability and performance despite bounded uncertainties in system parameters and external disturbances
- Adaptive control: Dynamically adjusts controller parameters in response to time-varying or initially unknown system properties

- Stochastic control: Models uncertainties probabilistically to optimize decision-making under random disturbances

With emerging challenges from larger systems and stricter safety requirements, these techniques are evolving to incorporate new tools such as machine learning while maintaining rigorous performance guarantees. These approaches help extract actionable knowledge from uncertain environments, enabling robust and adaptive decision-making.

4. Generative AI

Generative AI techniques, such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs) [28, 29], and Large Language Models (LLMs) [30], have emerged as transformative tools for creating synthetic data [31]. These data are increasingly used in domains where real-world data are scarce, expensive, or fraught with privacy concerns—examples include medical imaging [32], anomaly detection [33], and human activity recognition [34]. However, the use of synthetic data introduces new challenges in the extraction of actionable knowledge, stemming from the inherent uncertainty in the data's fidelity and representativeness.

Sources of Uncertainty in Synthetic Data

(a) Fidelity to Real Data**:** The synthetic data's realism and utility often depend on how well the generative models capture the distribution and nuances of the original data. This fidelity is influenced by:

- Model Limitations: Insufficient model capacity or training data can lead to artifacts or biases.
- Domain Variability**:** For example, synthetic Wi-Fi signal data used for activity recognition may fail to generalize across different environments without significant preprocessing.

(b) Measurement Dependencies: When synthetic data represent physical phenomena (e.g., sensor outputs), the transformations required to ensure environment-independent features (e.g., vectorization) introduce additional layers of uncertainty.

(c) Ambiguity in Reference Data**:** Domain knowledge or real data distributions, often used as benchmarks, may themselves be incomplete or imprecise.

Quantification and Management of Uncertainty

To extract actionable knowledge from synthetic data, it is essential to quantify and address uncertainty. The following approaches are pertinent:

(a) Reference-Based Validation: Synthetic data can be evaluated against real-world benchmarks or domain-specific distributions. For instance:

- Comparing statistical properties (e.g., means and variances).
- Using domain-specific tools like vectorization for meaningful similarity measurements.

(b) Stochastic Modeling: Estimating the uncertainty bounds of generated data using probabilistic methods can provide confidence levels for downstream applications.
(c) Digital Twin Integration: In industrial and IoT applications, digital twins can help gauge discrepancies between simulated synthetic data and observed real-world data, offering a pathway to refine models.

Challenges in Knowledge Extraction

The unique challenges of extracting actionable knowledge from synthetic data include:
(a) Feature Alignment: Transformations such as environment-independent feature extraction (e.g., using vectorization techniques akin to word2vec for time-series or spatial data) are necessary but complex.
(b) Bias Propagation: Generative models trained on biased or incomplete datasets can propagate and even amplify these biases, impacting decision-making.
(c) Robustness: Ensuring that models trained on synthetic data perform consistently under real-world conditions remains an open challenge.

Research Directions and Implications

- Domain-Specific Methods: Developing tailored techniques for synthetic data validation, such as leveraging control theory to design robust data correction algorithms.
- Collaborative Validation: Integrating insights from multiple domains, such as databases, AI, and signal processing, to improve synthetic data quality.
- Ethical and Practical Considerations: Addressing privacy and ethical implications, especially in sensitive applications like healthcare or surveillance.

Generative AI holds immense promise for addressing data scarcity and enhancing analytical capabilities. However, systematically managing and mitigating uncertainty is essential to ensure that the knowledge derived from synthetic data is not only actionable but also trustworthy and robust.

5. Conclusions

In this paper, we have discussed the concept of uncertainty and its impact in four specific domains of high current interest: databases, signal processing, control theory and generative AI. Our analysis highlights that uncertainty is present at multiple levels within each of these domains: at the level of data sources, at the level of data modeling and finally at the level of knowledge extraction from data. Given that the effectiveness of a data pipeline heavily depends on data quality, and uncertainty is a key factor influencing this quality, further research is essential to mitigate its impact [38, 39].

We therefore encourage the EU research funding agencies, as well as the funding bodies of the EU member states, and the Japan Science and Technology Agency to support research activities aiming to understand, quantify, and manage uncertainty. Advancing these efforts will enable more reliable knowledge extraction from uncertain data sources, ultimately enhancing the robustness of data-driven decision-making.

References

[1] Jäger, S., Allhorn, A., Bießmann, F., A benchmark for data imputation methods. Frontiers in Big Data, Vol.4, Article 693674, 2021. https://doi.org/10.3389/fdata.2021.693674

[2] Arenas, M., Bertossi, L., Chomicki, J., Consistent query answers in inconsistent databases. In PODS '99: Proceedings of the ACM SIGMOD/PODS International Conference on Principles of Database Systems, 1999.

[3] Laurent, D., Spyratos, N., Handling inconsistencies in tables with nulls and functional dependencies, Journal of Intelligent Information Systems, Vol.59, pp.285–317, 2022. https://doi.org/10.1007/s10844-022-00700-0

[4] Representing knowledge in an uncertain domain in AI, GeeksforGeeks, Retrieved June 13, 2024 from https://www.geeksforgeeks.org/representing-knowledge-in-an-uncertain-domain-in-ai/

[5] Uncertainty, Wikipedia, Retrieved March 9, 2025 from https://en.wikipedia.org/wiki/Uncertainty

[6] Payong, A., Introduction to uncertainty in machine learning models: Concepts and methods, Paperspace, Retrieved October 15, 2023 from https://blog.paperspace.com/aleatoric-and-epistemic-uncertainty-in-machine-learning/

[7] Office for National Statistics, Uncertainty and how we measure it for our surveys, Retrieved March 9, 2025 from https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/uncertaintyandhowwemeasureit

[8] Jeevanandam, N., The importance of probabilistic reasoning in AI, IndiaAI, Retrieved October 15, 2021 from https://indiaai.gov.in/article/the-importance-of-probabilistic-reasoning-in-ai

[9] California Institute of Technology, What is the uncertainty principle and why is it important? Caltech Science Exchange, Retrieved March 9, 2025 from https://scienceexchange.caltech.edu/topics/quantum-science-explained/uncertainty-principle

[10] ScienceDirect, Uncertainty quantification, Retrieved March 9, 2025, from https://www.sciencedirect.com/topics/engineering/uncertainty-quantification

[11] Wu, J., Shang, S., Managing uncertainty in AI-enabled decision making and achieving sustainability. Sustainability, Vol.12, No.21, 8758, 2020. https://doi.org/10.3390/su12218758

[12] Catak, F. O., & Kuzlu, M., Uncertainty quantification in large language models through convex hull analysis, Discovery of Artificial Intelligence, Vol.4, No.90, 2024. https://doi.org/10.1007/s44163-024-00200-w

[13] Pham, T. M., Pandis, N., & White, I. R., Missing data: Issues, concepts, methods. Seminars in Orthodontics, Vol.30, No.1, pp.37–44, 2024. https://doi.org/10.1053/j.sodo.2024.01.007

[14] Schäfer, P., Leser, U., Motiflets – Simple and accurate detection of motifs in time series, PVLDB, Vol.16, No.4, pp.725–737, 2022. https://doi.org/10.14778/3574245.3574257

[15] Kalkha, H., Khiat, A., Bahnasse, A., Ouajji, H., Enhancing warehouse efficiency with time series clustering: A hybrid storage location assignment strategy, IEEE Access, Vol.12, pp.52110–52126, 2024. https://doi.org/10.1109/ACCESS.2024.3386887

[16] Zhang, S., Feng, R., Wu, Y., Yu, N., Adaptive compressed sensing for acceleration data transmission in human motion capture, In Proceedings of the 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, pp. 1–6, 2017. https://doi.org/10.1109/CISP-BMEI.2017.8302268

[17] Yin, W., Zang, X., Wu, L., Zhang, X., Zhao, J., A distortion correction method based on actual camera imaging principles, Sensors, Vol.24, No.8, 2406, 2024. https://doi.org/10.3390/s24082406

[18] Faseeh, M., Bibi, M., Khan, M. A., Kim, D.-H.,Real-time long-range object tracking based on ensembled model, IEEE Access, Vol.13, pp.2679–2693, 2025. https://doi.org/10.1109/ACCESS.2024.3517711

[19] Byeon, H., Seno, M. E., Yajid, M. S. A., et al., Wearable sensor technology and medical robotics for fatigue assessment using electromyography signal processing, SIViP, Vol.18, pp.8767–8780, 2024. https://doi.org/10.1007/s11760-024-03505-6

[20] Wang, K., Yan, X., Retraction note: Performance analysis of ethylene-propylene diene monomer sound-absorbing materials based on image processing recognition, Journal of Image Video Processing, Vol.17, 2022. https://doi.org/10.1186/s13640-022-00598-2

[21] Wang, B., Jiang, Z., Sun, Y., et al., An intelligent signal processing method against impulsive noise interference in AIoT, EURASIP Journal on Advances in Signal Processing, Vol.104, 2023. https://doi.org/10.1186/s13634-023-01061-8

[22] Chen, X., Zhao, L., Xu, J., Liu, Z., Zhang, C., Xu, L., & Guo, N., A self-supervised contrastive denoising autoencoder-based noise suppression method for micro thrust measurement signals processing, IEEE Transactions on Instrumentation and Measurement, Vol.73, pp.1–17, 2024. https://doi.org/10.1109/TIM.2023.3341134

[23] Li, S., Signal processing and noise compensation algorithm for carrier communication system considering electric load fluctuations, In Proceedings of the International Conference on Big Data Mining and Information Processing (BDMIP ’23), pp.94–98, 2023. https://doi.org/10.1145/3645279.3645296

[24] Åström, K. J., Murray, R., Feedback systems: An introduction for scientists and engineers, Princeton University Press, 2021.

[25] Baggio, G., Bassett, D. S., Pasqualetti, F., Data-driven control of complex networks. Nature Communications, Vol.12, No.1, pp.1–13, 2021.

[26] 6th Annual Learning for Dynamics & Control Conference (L4DC), Oxford, 2024.

[27] Hobbs, K. L., Mote, M. L., Abate, M. C. L., Coogan, S. D., Feron, E. M., Runtime assurance for safety-critical systems: An introduction to safety filtering approaches for complex control systems, IEEE Control Systems Magazine, Vol.43, No.2, pp.28–65, 2024.

[28] Carbonera, M., Ciavotta, M., Messina, E., Variational autoencoders and generative adversarial networks for multivariate scenario generation, Data Science & Transportation, Vol.6, No.23, 2024. https://doi.org/10.1007/s42421-024-00097-y

[29] Devyatkin, D., Trenev, I., Data generation with variational autoencoders and generative adversarial networks, Engineering Proceedings, Vol.33, No.1, 37, 2023. https://doi.org/10.3390/engproc2023033037

[30] Choenni, S., Busker, T., Bargh, M. S., Generating synthetic data from large language models, In Proceedings of the 2023 15th International Conference on Innovations in Information Technology (IIT), pp. 73–78, 2023. https://doi.org/10.1109/IIT59782.2023.10366424

[31] Goyal, M., Mahmoud, Q. H., A systematic review of synthetic data generation techniques using generative AI, Electronics, Vol.13, No.17. 3509, 2024. https://doi.org/10.3390/electronics13173509

[32] Ali, M., Ali, M., Hussain, M., et al., Generative adversarial networks (GANs) for medical image processing: Recent advancements, Archives of Computational Methods in Engineering, 2024. https://doi.org/10.1007/s11831-024-10174-8

[33] Lin, W., Gao, J., Wang, Q., Li, X., Learning to detect anomaly events in crowd scenes from synthetic data. Neurocomputing, Vol.436, pp.248–259, 2021. https://doi.org/10.1016/j.neucom.2021.01.031

[34] Joshi, I., Grimmer, M., Rathgeb, C., Busch, C., Bremond, F., Dantcheva, A., Synthetic data in human analysis: A survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.46, No.7, pp.4957–4976, 2024. https://doi.org/10.1109/TPAMI.2024.3362821

[35] Ghanem, R., Higdon, D., Owhadi, H., Handbook of uncertainty quantification, Springer Cham, 2017. https://doi.org/10.1007/978-3-319-12385-1

[36] Laurent, D., Spyratos, N., Consistent query answering in multi-relation databases. Information and Computation, Vol.303, 2025. https://doi.org/10.1016/j.ic.2025.105279

[37] Ullman, J. D., Principles of databases and knowledge-base systems (Vols. 1–2). Computer Science Press, 1988.

[38] Scholes, M. S., Artificial intelligence and uncertainty, Risk Sciences Journal, 2025.

[39] Salimzadeh, S., He, G., Gadiraju, U., Dealing with uncertainty: Understanding the impact of prognostic versus diagnostic tasks on trust and reliance in human-AI decision-making. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24), Article No.25, pp.1–17, 2024. https://doi.org/10.1145/3613904.3641905

[40] Benny Kimelfeld, Phokion G. Kolaitis, A Unifying framework for incompleteness, Inconsistency, and uncertainty in Databases, Communications of the ACM, March 2024 No 3