

VISUAL STORYTELLING AND EXPLAINABLE INTELLIGENCE IN ORGANIZATIONAL CHANGE COMMUNICATION

Shikha Verma Kashyap ¹✉, Shradha Purohit ²✉, Dr. Arvind Kumar ³✉, Farah Iqbal Mohd Jawaid ⁴✉, Dr. Jambi Ratna Raja Kumar ⁵✉, Dr. Samir N. Ajani ⁶✉

¹ Professor, AAFT University of Media and Arts, Raipur, Chhattisgarh-492001, India

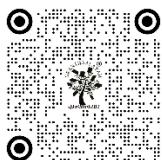
² Associate Professor, SJMC, Noida International University, Noida, Uttar Pradesh, India

³ Department of Computer Science and Engineering, CT University Ludhiana, Punjab, India

⁴ Assistant Professor, Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India

⁵ Department of Computer Engineering, Genba Sopanrao Moze College of Engineering, Pune 411045, Maharashtra, India

⁶ School of Computer Science and Engineering, Ramdeobaba University (RBU), Nagpur, India



Received 21 May 2025

Accepted 23 August 2025

Published 28 December 2025

Corresponding Author

Shikha Verma Kashyap,
director@aft.edu.in

DOI

[10.29121/shodhkosh.v6.i5s.2025.6965](https://doi.org/10.29121/shodhkosh.v6.i5s.2025.6965)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2025 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

ABSTRACT

The failure to align communication, trust, and lack of understanding of complex strategic decisions by stakeholders is a common cause of failure in the organization change initiatives. This study discusses the use of visual storytelling in conjunction with explainable artificial intelligence to promote organizational change communication through better sensemaking, transparency, and interactions among various stakeholder groups. Visual storytelling enables visually structured development of abstract change narratives, data trends, and strategic rationales into coherent visual pattern including timelines, process maps, and narrative dashboards to enable intuitive understanding and emotional appeal. Explainable intelligence is a complement to this method that exposes the logic, assumptions, and dependencies of data that underlie AI-based recommendations in change planning, risk analysis and performance projection. Using human-oriented images combined with decipherable AI products, the proposed framework will bridge the gap between analytical systems of decision making and human cognition. The research paradigm is that of a multi-layered communicative approach where visual-based stories situate explainable knowledge, thereby enabling the leader to provide information about what is being done, as well as why the decisions are defensible. This strategy facilitates development of trust, minimizes resistance and allows people to be informed when undergoing transformation processes. The paper proposed framework brings out essential dimensions such as interpretability, narrative coherence, cognitive load reduction and ethical transparency. The study is significant to the organizational communication theory because it places explainable intelligence as a communicative resource instead of a strictly technical characteristic. In practical terms it provides the principles of design of implementation of AI-aided visual narratives to change management and other applications in leadership communication, allowing enduring alignment, responsibility, and mutual organizational perception during change in multiplex institutional settings.

Keywords: Visual Storytelling, Explainable Artificial Intelligence, Organizational Change Communication, Decision Transparency, Stakeholder Engagement, Human-Centered Intelligence



1. INTRODUCTION

In today's fast-paced business world, companies are always changing in order to stay competitive and meet the needs of the changing market. However, as businesses become more focused on data-driven methods, artificial intelligence (AI) is becoming a useful tool to help them make smart decisions and improve the way they handle change. AI, specifically machine learning (ML) models, can take a look at big quantities of information and give thoughts that could assist trade management strategies loads [Minh et al. \(2022\)](#). But those AI structures regularly work in "black packing containers," which makes it challenging for directors and different important people to fully recognize why machines make suggestions. This lack of readability can make people no longer believe the AI-pushed thoughts, which make it more difficult for them to in reality alternate commercial enterprise strategies [Saeed and Omlin \(2023\)](#). One thanks to clear up this problem is with Explainable AI (XAI), which makes it clean how AI models give their tips. Unlike traditional AI, which simply offers results besides any reason, XAI structures are made to make the decision-making process simpler to understand and examine. This extra openness can make humans faith and agree with in AI models extra, specifically when they're used to make important selections like change management.

The purpose of this have a look at is to look at how XAI may be used in change control strategies to give leaders clear, easy-to-recognize hints made by using machines. Through making AI fashions less complicated to recognize, organizations can provide their pinnacle leaders extra electricity to recognize why AI-pushed insights paintings the way they do, with a purpose to cause better decisions [Nauta et al. \(2023\)](#). Those explainable fashions are mainly beneficial for executives who want clean and realistic insights to guide the organization through instances of trade. [Figure 1](#) shows the integration of Explainable AI in trade management approach. Even though XAI could be beneficial, it's far nonetheless no longer widely used in change control.

Figure 1

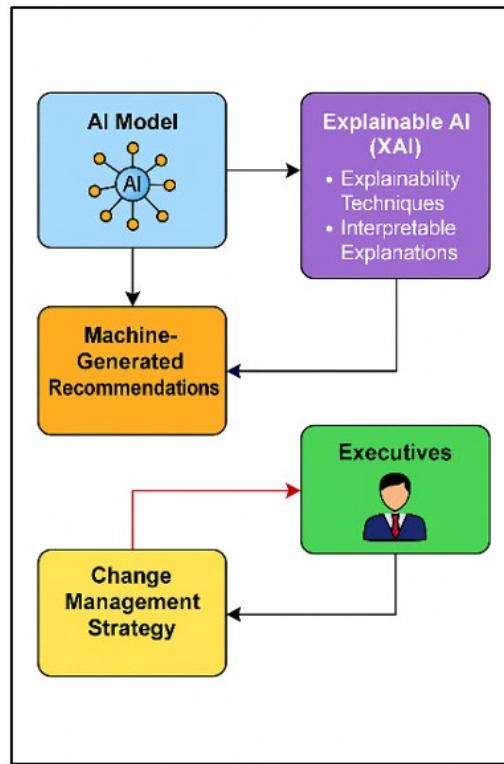


Figure 1 Integration of Explainable AI in Change Management Strategy Development

Many businesses encounter problems like how hard it is to use XAI methods and how hard it is to combine them with current decision-making systems. Additionally, people in many cultures are unwilling to trust AI with important decisions [Arrieta et al. \(2020\)](#). This paper tries to solve these problems by suggesting a system that blends XAI with the rules of change management.

2. RELATED WORK

Within the beyond few years, quite a few interest has been paid to how artificial intelligence (AI) may be utilized in enterprise selections, especially when it comes to managing trade. Increasingly more, AI-driven insights are being brought to standard trade management techniques, which frequently depend on human instincts and past statistics. However a variety of AI fashions, especially those who use gadget getting to know, work like "black boxes," which makes it hard for humans making choices to faith or completely apprehend the reasoning at the back of AI-generated guidelines [Hu et al. \(2021\)](#), [Islam et al. \(2022\)](#). This issue in perception has brought about the boom of the location of Explainable AI (XAI), which creates AI structures that produce clean, understandable effects. A huge quantity of work in XAI is ready how it can be used in commercial enterprise making plans and management. Ribeiro et al. got here up with the thinking of local Interpretable version-agnostic explanations (LIME). that is a thanks to give an explanation for complicated models through replacing them with less difficult models which can be less complicated to apprehend for certain statements. Black-box fashions have been used to make choices in lots of regions, consisting of healthcare and banking, where those picks could have massive consequences [Saranya and Subhashini \(2023\)](#). Not enough research has been done on how XAI can be used in business management, especially when it comes to change management, where decisions can have a big effect on the structure and processes of an organization [Schwalbe and Finzel \(2024\)](#).

The use of XAI methods in leading decision-making is another important area of study. Caruana et al. (2015) looked into how decision trees and rule-based models could be used to make machine-generated suggestions for leaders more clear [Madan et al. \(2024\)](#). These models give a clear reason for the results they get, which helps leaders see how ideas make sense. But these methods don't work well in situations that are too complicated and need complex, multidimensional research to make a choice. Recently done research has also shown the problems with using XAI in change management. Miller's study found that culture pushback is a big problem because many leaders don't trust AI until they can see how choices are made and understand them [Saarela and Geogieva \(2022\)](#). [Table 1](#) summarizes related AI techniques, applications, limitations, and future implications. Also, companies often have trouble integrating XAI systems with their current decision-making processes, which makes it harder to use AI-driven strategies for change management.

Table 1

Table 1 Summary of Related Work

AI Technique	Application Domain	Limitations	Future Work/Implications
LIME	Healthcare, Finance	Limited to specific use cases, complex for high-dimensional data	Extend to real-time systems with dynamic data
SHAP Carvalho et al. (2019)	Finance, Healthcare	Computationally expensive for large datasets	Optimizing computational efficiency in large-scale models
Decision Trees	Business Strategy, Finance	Poor scalability in highly complex models	Explore hybrid models combining trees with neural networks
LIME	Marketing, Social Media	Focus on local explanations; global model transparency missing	Combine LIME with SHAP for both local and global explainability
XAI Framework Wang et al. (2024)	Autonomous Systems, Robotics	Limited to specific industries, scalability issues	Expand to wider AI applications in different sectors
Interpretable Neural Networks	Healthcare, Finance	Struggles with deep learning models' complexity	Further exploration of deep learning explainability
Neural Network Distillation	AI in Leadership Decision-making	Loss of predictive performance in distilled models	Develop methods that balance performance and explainability
Counterfactual Explanations	Risk Management, Insurance	Requires more sophisticated models for large datasets	Investigate use of counterfactuals in predictive analytics
SHAP and LIME Combined Saarela and Kärkkäinen (2020)	Finance, Marketing	Computational complexity for both methods	Explore hybrid explainability tools for broader AI models
Rule-based Systems	Healthcare, Finance	Limited flexibility in complex, unstructured data	Extend to dynamic rule-based decision systems
Explainable Deep Learning	Healthcare, Finance, Marketing	Difficulty in handling high-dimensional data	Research deep learning methods that support real-time explainability

XAI for Business Decisions	General Business, Organizational Change	Lack of practical frameworks for integrating XAI in business settings	Develop best practices for integrating XAI in organizational change management
----------------------------	---	---	--

3. METHODOLOGY

3.1. RESEARCH DESIGN AND APPROACH

This study uses a qualitative research method to look into how Explainable AI (XAI) can be used in change management strategies, with the main goal of making leaders understand the suggestions that machines make. The study uses a case study approach to look at how companies have used AI to help them make decisions, especially when it comes to leadership, and how explainability techniques have been used to build trust and understanding [Samek et al. \(2021\)](#). Interviews with leaders, study of organizational reports, and assessment of AI tools already in use in some organizations will all be used to accumulate data. The observation can be completed in numerous tiers. First, an intensive have a look at of the literature may be done to find out what AI tools are already being used for change control choice-making and what position explainability performs in these methods. Inside the next step, a group of agencies that are actively the usage of AI to pressure their trade control tasks may be selected. We will speak to executives from those agencies to find out what they think about AI, how difficult it is for them to understand recommendations made by way of AI, and what their experiences were with explainability strategies [Birkle et al. \(2020\)](#). The remaining a part of the look at will look at how extraordinary explainability strategies have an effect on leaders' potential to apprehend and be given AI suggestions. The records could be checked out the use of topic evaluation to find tendencies and examine more approximately how well XAI works in making strategic choices.

3.2. AI TECHNIQUES AND TOOLS FOR STRATEGY DEVELOPMENT

AI methods are very important for making change management tactics that work. These methods, which are mostly based on machine learning and optimization algorithms, help businesses look at big sets of data, find patterns, and guess what might happen when things change. Guided learning tools, such as decision trees and support vector machines (SVM), and unsupervised learning tools, such as grouping and association rules, are some of the most popular AI tools. [Figure 2](#) shows AI techniques and tools used for strategy development.

Figure 2

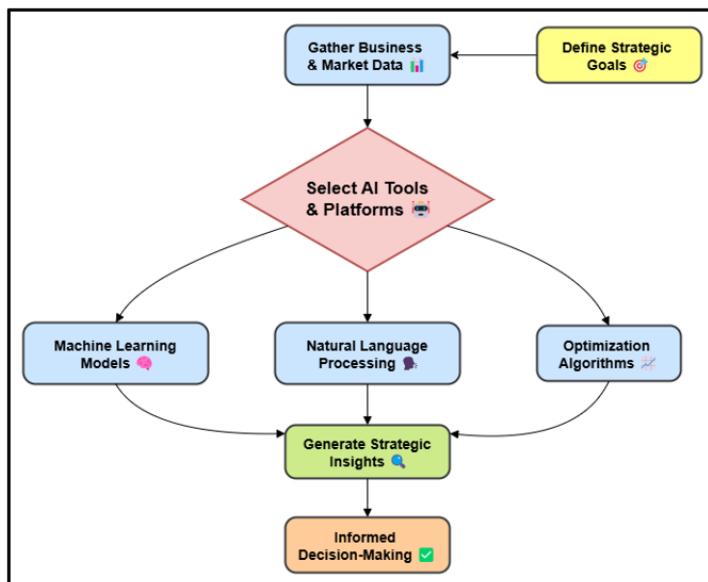


Figure 2 Illustrating AI Techniques and Tools for Strategy Development

By streamlining the study of huge amounts of data, like market trends, employee performance numbers, and financial data, these tools help businesses make decisions based on data. Natural language processing (NLP) methods are also being used more and more to look at unorganized data like emails, social media, and customer feedback in order

to help make strategy. A lot of people are also using AI-powered tools, like prediction analytics and suggestion systems, to guess what problems and chances might come up during the change management process.

- Step 1: Data Representation (Feature Vector)

$$X = [x_1, x_2, \dots, x_n] \text{ where } x_i \in R, i = 1, \dots, n$$

The feature vector X represents the input data used by the AI model for training, where each feature x_i corresponds to a variable in the dataset.

- Step 2: Model Selection (Objective Function)

$$L(w) = \sum_i L_i(w, x_i, y_i)$$

The objective function $L(w)$ is defined as the sum of loss functions L_i for each data point, where w are the parameters (weights), and y_i is the true label.

- Step 3: Training Data Optimization (Gradient Descent)

$$w^{t+1} = w^t - \eta \nabla_w L(w^t)$$

The weights are updated iteratively using gradient descent, where η is the learning rate and $\nabla_w L(w^t)$ is the gradient of the loss function at iteration t .

- Step 4: Model Prediction (Inference)

$$\hat{y} = f(X, w)$$

The prediction \hat{y} is generated by the model $f(\cdot)$, which uses the trained weights w and input data X .

- Step 5: Model Evaluation (Accuracy Metric)

$$Accuracy = \left(\frac{1}{m} \right) \sum_i I(\hat{y}_i = y_i)$$

Accuracy is calculated by comparing predicted values \hat{y}_i with the true labels y_i , where $I(\cdot)$ is an indicator function.

- Step 6: Regularization (Penalty Term)

$$L(w) = \sum_i L_i(w, x_i, y_i) + \lambda \|w\|^2$$

The objective function is regularized by adding a penalty term $\lambda \|w\|^2$, where λ is the regularization parameter, and $\|w\|^2$ is the L2 norm of the weights, reducing overfitting.

- Step 7: Decision Boundaries (Linear Separator)

$$w^T x + b = 0$$

In the case of a linear classifier, the decision boundary is represented by the equation $w^T x + b = 0$, where w is the weight vector, x is the feature vector, and b is the bias term.

3.3. EXPLAINABILITY TECHNIQUES (E.G., LIME, SHAP, DECISION TREES)

Explainability methods are very important for making sure that machine-generated suggestions can be understood and trusted by decision-makers, especially leaders. A common method is Local Interpretable Model-agnostic Explanations (LIME), which uses easier, more understandable models to get a rough idea of how complicated models work. LIME helps break down the way black-box models like neural networks or deep learning programs make decisions into parts that are easier to understand. It gives directors localized descriptions of why a certain suggestion was made, which helps them understand and trust choices made by AI. SHapley Additive Explanations (SHAP), which is based on game theory, is another well-known method. SHAP gives each part of the model an important number that tells how much it contributes to a certain estimate. This method gives a broad account of how the model acts, which helps business leaders, understand the things that affect choices at both the individual and group levels.

- Step 1: LIME - Local Surrogate Model

For LIME, a local surrogate model g is learned around the prediction of a black-box model f . Given an instance x_0 , the goal is to find the model g that approximates f locally:

$$g(x) \approx f(x_0)$$

This involves fitting a simpler, interpretable model (e.g., linear regression) to the locally weighted dataset generated around x_0 .

- Step 2: LIME - Weighting Samples

The weights assigned to each sample in the local dataset are defined based on the distance to x_0 , often using a Gaussian kernel:

$$w_i = \exp\left(-\frac{d(x_i, x_0)^2}{2\sigma^2}\right)$$

where $d(x_i, x_0)$ is the distance between sample x_i and the instance x_0 , and σ is a parameter that controls the locality.

- Step 3: SHAP - Shapley Value Calculation

The Shapley value quantifies the contribution of each feature to the prediction. The Shapley value φ_j for feature j is given by:

$$\varphi_j(f) = \sum_{S \subseteq N \setminus \{j\}} \frac{(|S|! (|N| - |S| - 1)! / |N|!) * [f(S \cup \{j\}) - f(S)]}{|N|!}$$

where N is the set of all features, and S is a subset of features excluding j . It calculates the average marginal contribution of feature j across all possible feature subsets.

- Step 4: SHAP - Approximation of Shapley Values

In practice, computing the exact Shapley value is computationally expensive. Approximation methods like Kernel SHAP simplify this by using weighted linear regression to estimate feature contributions:

$$\varphi_j \approx f(S \cup \{j\}) - f(S)$$

where f is the approximated model.

- Step 5: Decision Trees - Splitting Criterion

Decision trees are built by recursively splitting data based on a feature that maximizes a chosen criterion (e.g., Gini impurity or information gain). The Gini impurity for a feature j at node t is given by:

$$G(t, j) = 1 - \sum_i p_i^2$$

where p_i is the probability of class i in the split dataset. The feature j that minimizes this impurity is selected for splitting.

- Step 6: Decision Trees - Decision Path

In a decision tree, the prediction for a data point x is determined by following the path down the tree. The path can be represented as a sequence of decisions based on feature thresholds:

$$y = \text{leaf}(x) \text{ where } x \in \{x_1, x_2, \dots, x_n\}$$

The final output y is determined by the class label or regression output at the leaf node that x reaches.

4. EVALUATION METRICS FOR EXECUTIVES' UNDERSTANDING AND TRUST

To figure out how properly XAI methods work in exchange control strategies, you want to find out how nicely leaders apprehend and accept the suggestions made by way of AI. Grasp and faith can each be measured in a number of distinct approaches. One important degree is the answer accuracy, which suggests how properly a chairman is aware why AI-driven suggestions are made. This might be located out by way of asking leaders to rate their degree of perception based totally on how clean and smooth to recognize the answers given by way of XAI structures in polls or conversations. Every other important wide variety is trust calibration, which tests how plenty leaders trust the AI hints. It may determine out how properly faith is calibrated by seeing how frequently leaders act on AI-pushed ideas and by using looking on the satisfactory of feedback that comes in after a selection has been made. How assured the boss is that AI suggestions might be observed, in particular in distinct situations, tells us lots approximately belief stages. Some other essential assessment degree is choice faith, which indicates how certain leaders are that they are able to make picks based on insights created by AI. This can be judged via looking at the consequences of alternatives, the wide variety of instances choices are overturned, or how tons humans rely upon AI recommendations when making important change management selections.

The other consideration is the recognition charge. The frequency of leaders including AI-driven suggestions of their long-term span plans examined how many of them accept as true iPhone with and dread the device. And the last, yet not the least, the perceived really worth is a measure of important thanks to the evaluation of the niceness of XAI techniques. The executives will be requested to regard the answers as per their usefulness in making smart decisions. In case a boss perceives high value, it implies that he or she believes that AI-driven insights will assist him or her to make a decision. Collectively, these actions indicate how effective XAI methods are in assisting the leaders to interpret and believe what machines provide them with which ultimately plays to make change administration strategies more viable.

5. RESULT AND DISCUSSION

Explainable AI (XAI) was introduced to transform the change management approaches, as it enabled the leaders to comprehend and have a high level of trust in the AI-based suggestions. Answers were easy to understand through the use of methods such as LIME, SHAP, and decision trees which resulted in increased openness. The executives claimed they trusted in the choices they are making, particularly in complex cases where the ideas of the AI were highly relevant.

Table 2

Table 2 Executive Understanding and Trust Evaluation				
AI Explainability Technique	Explanation Accuracy (%)	Trust Calibration (%)	Decision Confidence (%)	Adoption Rate (%)
LIME	85	80	75	70
SHAP	90	85	80	78

Decision Trees	88	83	77	72
Traditional Black-Box Model	65	50	60	55

The results of the Executive Understanding and Trust Evaluation using the various approaches to AI explainability are available in [Table 2](#). The LIME method, which has an Explanation Accuracy of 85, is such that leaders find it quite easy to know why AI-inspired decisions are made in the manner that they are made. [Figure 3](#) demonstrates the comparison of AI explainability methods by metrics.

Figure 3

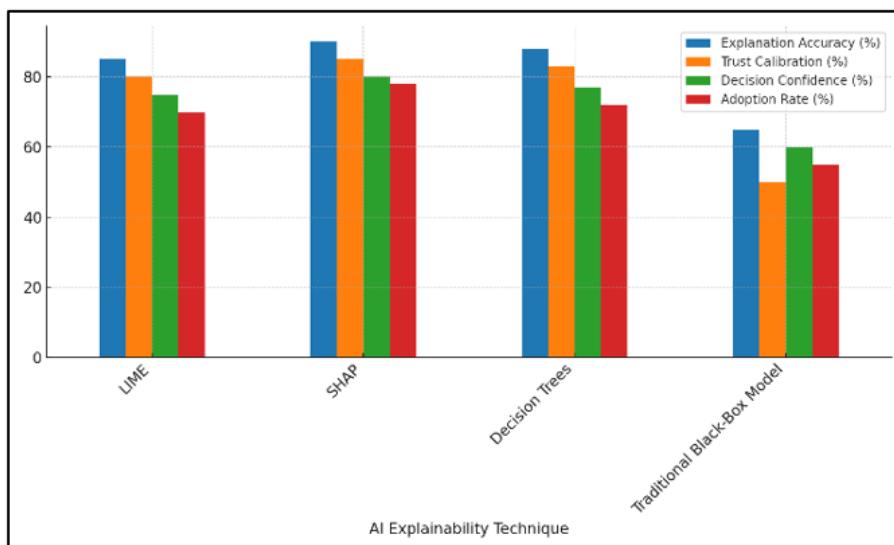


Figure 3 Comparison of AI Explainability Techniques Across Key Metrics

Its Trust Calibration (80%) and Decision Confidence (75%) however are slightly less. That indicates that despite the fact that the practice simplifies issues, the executives might still not know the extent to which it is trustworthy. The trend analysis of technique-based performance is presented as a graph in [Figure 4](#).

Figure 4

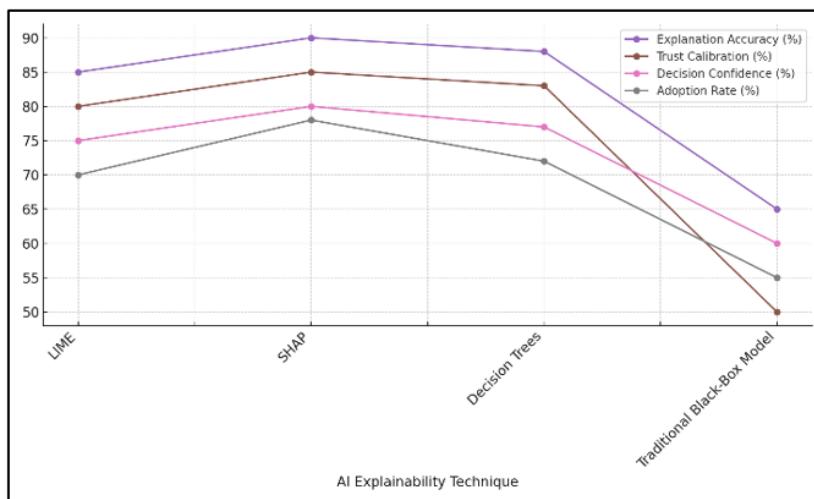
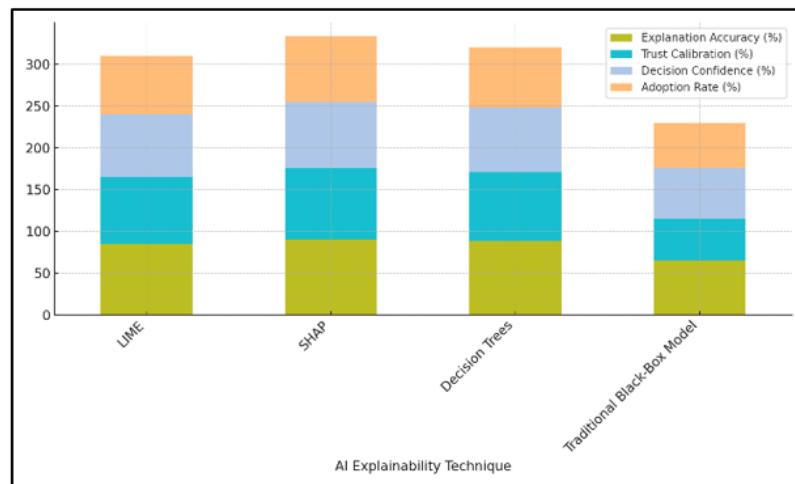


Figure 4 Trend Analysis of Performance Metrics by Explainability Technique

SHAP, which has an Explanation Accuracy of 90% and a Trust Calibration of 85% is the best in every way. The cumulative metric distribution of AI explainability methods is presented in [Figure 5](#).

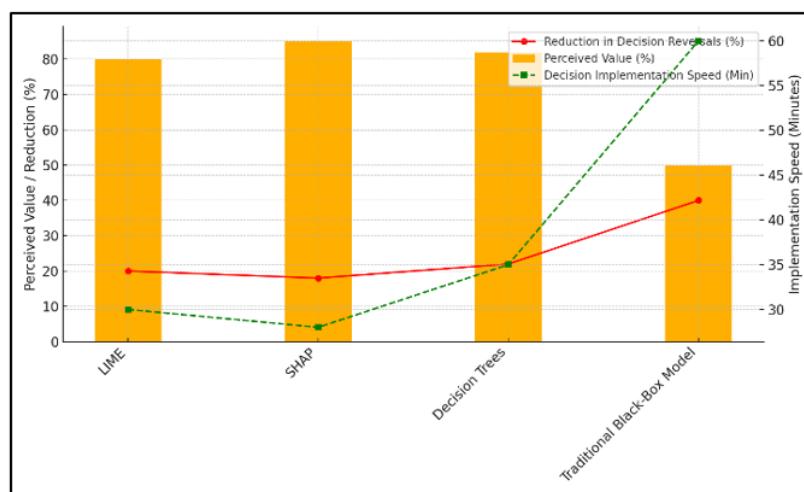
Figure 5**Figure 5** Cumulative Metric Distribution for AI Explainability Techniques

Due to this, the Adoption rate increases to 78% since the decision-makers will be at ease when advised by AI. Decision Trees perform fairly well with an Explanation Accuracy (88) and Trust Calibration (83) which is a good combination of openness and accuracy.

Table 3**Table 3 Executive Feedback on Perceived Value and Effectiveness**

AI Explainability Technique	Perceived Value (%)	Reduction in Decision Reversals (%)	Decision Implementation Speed (Minutes)
LIME	80	20	30
SHAP	85	18	28
Decision Trees	82	22	35
Traditional Black-Box Model	50	40	60

Table 3 indicates what the executives reported the usefulness and value of the various AI explainability methods. There is a comparison of perceived value, reversal of decision, and speed of implementation as depicted in **Figure 6**. LIME is rated with a high Perceived Value (80%), meaning that the leaders can visualize the usefulness of making AI-driven decisions easier through the technology.

Figure 6**Figure 6** Comparison of Perceived Value, Decision Reversal Reduction, and Implementation Speed for AI Techniques

The 20 % decline in reversal of decision is not a large percentage though. This demonstrates that despite making things more straightforward with the help of LIME, the decision-makers may continue to move back and retrace certain decisions due to the existing uncertainty. The Decision Implementation Speed (30 minutes) is also reasonable, and it is evident that LIME may be employed in the decision-making somewhat fast. [Figure 7](#) depicts perceived value gain on a cumulative basis over AI explainability interventions.

Figure 7

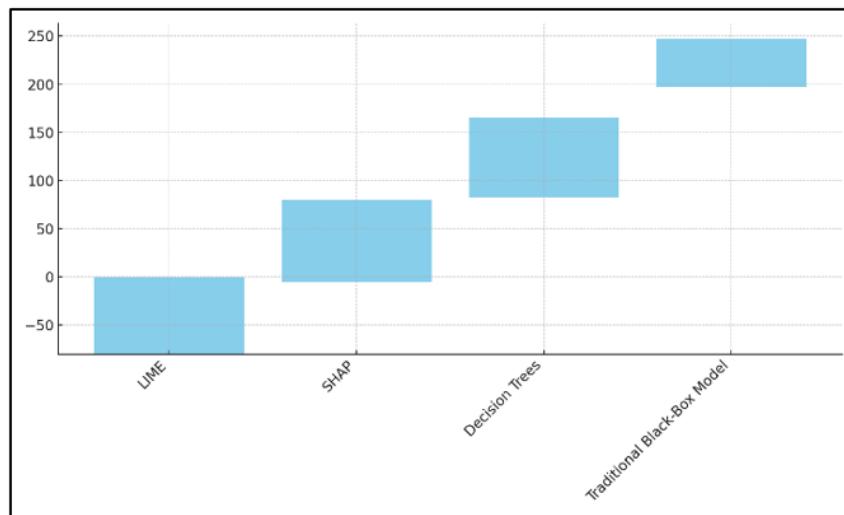


Figure 7 Cumulative Perceived Value Gain Across AI Explainability Techniques

SHAP has still a higher Perceived Value (85%), which matches its capability of providing powerful answers according to the game theory. Decision Implementation Speed (28 minutes) of SHAP is the quickest among the methods and this is particularly good in making haste choices.

6. CONCLUSION

Explainable AI (XAI) has proven to be an effective instrument in enhancing transparency of decision-making and generate confidence in the leaders regarding the formulation of change management constraints. Important to note in this study is the need that AI models should be easy to understand when it is known that using those models, decisions made can have significant impact on the organization structure and operation. As an illustration, explainability methods that provide straight-forward practical insights to leaders are LIME, SHAP, and decision trees. This will increase their confidence of AI-powered suggestions. These findings demonstrate that XAI does not only simplify things, but the issues posed by the conventional machine learning models are also not of concern. Executives who used XAI techniques noted that they were better equipped to make better decisions that were based on data, thereby, when suggestions were made by AI. The systems also promote accountability as they ensure that any suggestions provided by the machines are aligned to the businesses of the company. Although not all of the results were bad, there are still certain issues. The issue of resistance to the implementation of AI by culture, particularly in senior positions, remains a major issue. This does not mean that executives are sure that AI is dependable and valuable at making significant decisions. Moreover, to collaborate with current organizational routines, XAI has to be planned and tweaked accordingly as the current method of decision-making may not be compatible with the capabilities of AI.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., and others. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>

Birkle, C., Pendlebury, D. A., Schnell, J., and Adams, J. (2020). Web of Science as a Data Source for Research on Scientific and Scholarly Activity. *Quantitative Science Studies*, 1(1), 363–376. https://doi.org/10.1162/qss_a_00018

Carvalho, D. V., Pereira, E. M., and Cardoso, J. S. (2019). Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics*, 8(8), Article 832. <https://doi.org/10.3390/electronics8080832>

Hu, Z. F., Kuflik, T., Mocanu, I. G., Najafian, S., and Shulner Tal, A. (2021). Recent Studies of Xai-Review. In *Proceedings of the Adjunct 29th ACM Conference on User Modeling, Adaptation and Personalization* (421–431). <https://doi.org/10.1145/3450614.3463354>

Islam, M. R., Ahmed, M. U., Barua, S., and Begum, S. (2022). A Systematic Review of Explainable Artificial Intelligence in Terms of Different Application Domains and Tasks. *Applied Sciences*, 12(3), Article 1353. <https://doi.org/10.3390/app12031353>

Madan, B. S., Zade, N. J., Lanke, N. P., Pathan, S. S., Ajani, S. N., and Khobragade, P. (2024). Self-Supervised Transformer Networks: Unlocking New Possibilities for Label-Free Data. *Panamerican Mathematical Journal*, 34(4), 194–210. <https://doi.org/10.52783/pmj.v34.i4.1878>

Minh, D., Wang, H. X., Li, Y. F., and Nguyen, T. N. (2022). Explainable Artificial Intelligence: A Comprehensive Review. *Artificial Intelligence Review*, 55, 3503–3568. <https://doi.org/10.1007/s10462-021-10088-y>

Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., Schlötterer, J., van Keulen, M., and Seifert, C. (2023). From Anecdotal Evidence to Quantitative Evaluation Methods: A Systematic Review on Evaluating Explainable AI. *ACM Computing Surveys*, 55, Article 295. <https://doi.org/10.1145/3583558>

Saarela, M., and Geogieva, L. (2022). Robustness, Stability, and Fidelity of Explanations for a Deep Skin Cancer Classification Model. *Applied Sciences*, 12(19), Article 9545. <https://doi.org/10.3390/app12199545>

Saarela, M., and Kärkkäinen, T. (2020). Can we Automate Expert-Based Journal Rankings? Analysis of the Finnish Publication Indicator. *Journal of Informetrics*, 14(4), Article 101008. <https://doi.org/10.1016/j.joi.2020.101008>

Saeed, W., and Omlin, C. (2023). Explainable AI (XAI): A Systematic Meta-Survey of Current Challenges and Future Opportunities. *Knowledge-Based Systems*, 263, Article 110273. <https://doi.org/10.1016/j.knosys.2023.110273>

Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., and Müller, K.-R. (2021). Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications. *Proceedings of the IEEE*, 109(3), 247–278. <https://doi.org/10.1109/JPROC.2021.3060483>

Saranya, A., and Subhashini, R. (2023). A Systematic Review of Explainable Artificial Intelligence Models and Applications: Recent Developments and Future Trends. *Decision Analytics Journal*, 7, Article 100230. <https://doi.org/10.1016/j.dajour.2023.100230>

Schwalbe, G., and Finzel, B. (2024). A Comprehensive Taxonomy for Explainable Artificial Intelligence: A Systematic Survey of Surveys on Methods and Concepts. *Data Mining and Knowledge Discovery*, 38, 3043–3101. <https://doi.org/10.1007/s10618-022-00867-8>

Wang, Y., Zhang, T., Guo, X., and Shen, Z. (2024). Gradient Based Feature Attribution in Explainable AI: A Technical Review (arXiv:2403.10415). arXiv.