

Conclusion

*“With the right programming, a computer can become a theater,
a musical instrument, a reference book, a chess opponent.
No other entity in the world except a human being has such an adaptable,
universal nature”.*

Daniel Hillis⁷⁴



This study has presented Explainable Artificial Intelligence (XAI), its fundamentals, context and techniques for improving model interpretability. The main challenges facing artificial intelligence models in terms of interpretability and how technology can help to address them have been discussed, and a use case developed with ModelCraft™ has been shared to demonstrate how these techniques can be employed to understand and explain AI models.

The AI discipline, and within it XAI, has grown in importance worldwide in recent years as developing high-performance AI technologies has become a priority for many sectors, from health to security, financial services to energy and many others. Interpretability arises as the need to understand and improve AI models, which is particularly complex in the case of certain techniques.

As seen, it can be difficult for AI models to explain their outcome or the logic behind their decisions. This is because these models use deep learning techniques and complex algorithms to learn from data, which are often difficult to interpret, and this poses challenges in evaluating AI models and the reliability of their output.

As a result, the AI regulatory framework is evolving rapidly, and organizations are expected to adapt to new requirements for transparency, explainability and fairness in the use of AI models. This implies the need for a comprehensive approach to integrate interpretability and explainability into each organization and its processes, encompassing interpretability techniques, model risk management, interdisciplinary collaboration and XAI training for professionals involved in AI development and implementation, among other areas.

In conclusion, the interpretability of artificial intelligence models is an emerging area of research, and it is expected to continue to develop and grow in importance as AI models become more complex, regulation continues to proliferate, and their use extends to more highly sensitive domains.

⁷⁴Daniel Hillis (b. 1956), American inventor, entrepreneur and scientist, pioneer of parallel computing and its use in the field of artificial intelligence, with more than 300 published patents.