TRENDS REPORT

The State Of Explainable AI, 2024

February 6, 2024

By Brandon Purcell with Aaron Katz, Evan Megan, Jen Barton

FORRESTER®

Summary

For enterprises to get value from their investments in AI, they must build trust with users and provide a view into how these systems reach their decisions. Enter explainable AI. Enterprises use explainability techniques and solutions across use cases and industries. These use cases continue to expand as enterprises apply AI to more aspects of their business. This report explores the current state of explainable AI, how companies access and use it, and what they should consider when implementing it.

Explainability Bridges The AI Trust Gap

Artificial intelligence suffers from a trust problem that spans stakeholder groups. Enterprises need to work toward minimizing this trust gap to realize the full potential value of their AI investments. Fortunately, enterprises have seven levers available to them to build trust in AI systems. One of these levers is transparency, the perception that an AI system is leading to decisions in an open and traceable way. Explainable AI (XAI) is a critical transparency mechanism that Forrester defines as:

Techniques and software capabilities for ensuring that people understand how AI systems arrive at their outputs.

In Forrester's Priorities Survey, 2023, 55% of business and technology professionals with knowledge of XAI say that their organization has implemented XAI; a further 13% report plans to adopt it in the next 12 months. These forward-thinking enterprises are adopting XAI to bridge the trust gap with:

- **Employees who use AI.** Most employees who will leverage AI systems to perform their daily functions are not data scientists. Clients tell us that adoption of AI systems suffers when employees don't have at least a high-level understanding of how the system produces results. In Forrester's Data And Analytics Survey, 2023, 25% of data and analytics decision-makers say that lack of trust in AI systems is a major concern in using AI; 21% cite a lack of transparency with AI/ML systems and models. By delivering an understanding of how AI systems make critical decisions, XAI generates trust among employees, which in turn increases adoption of AI systems.

- **Customers who experience AI.** From fraud detection to product recommendations to medical diagnoses, AI is increasingly playing a role in customer experiences. But as AI proliferates, so do the horror stories of AI gone awry. As of January 2024, the AI Incident Database, which tracks AI harms, contained over 3,100 reports of incidents spanning such categories as misinformation, wrongful arrests, and autonomous vehicular fatalities. As customers hear these stories, their willingness to engage with your AI experiences erodes. In Forrester's Technology And Telecom Topic Insights Survey, 2023, only 26% of online adults agreed that they would trust information provided by generative AI tools. For this reason, when Google and Microsoft launched generative AI capabilities in search, they included source citations to inspire trust among users.

- **Regulators who oversee AI.** Companies that don't invest in explainable AI to build trust with employees and customers may soon have to do so to appease regulators. Between the White House's recent executive order on AI and the European Commission's impending AI Act, regulatory action is imminent. Article 13 of the EU's proposed AI Act states: "High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately." And Section 8 of President Biden's executive order encourages agencies to consider rulemaking "related to the transparency of AI models and regulated entities' ability to explain their use of AI models."

Explainable AI Technology Caters To Different Enterprise Users

Explainable AI emerged to overcome the opacity of neural networks and other modern machine learning techniques. Researchers made great strides in developing interpretability techniques like SHAP and LIME, which are open source and have been popular among data scientists. But data scientists aren't the only ones who desire explainability. Enterprises can choose persona-specific solutions via XAI embedded in:

- **Machine learning platforms for data scientists.** Many larger machine learning platform vendors offer explainable AI capabilities on top of their existing model development functionality. For the most part, these vendors serve a wide array of responsible AI needs — from model interpretability to bias detection to model lineage, documentation, and of course, explainability. Some vendors, like DataRobot, offer SHAP natively in their platforms while others, like IBM, offer proprietary explainability methods as well.

- **Model intelligence platforms for AI governance teams.** As AI proliferates, so do AI governance efforts. AI governance teams need a "single pane of glass" into their enterprise's use of AI. Enter model intelligence platforms, which sit on top of previously developed and deployed models and provide transparency, interpretability, bias detection, and other explainability-related model assessments. These platforms are agnostic to model types and origin and provide a business intelligence interface or metrics cockpit so governance, risk, and compliance pros, data scientists, developers, and even business users can supervise model performance and intervene if an issue arises. Some model intelligence platforms like TruEra have proprietary interpretability techniques while others package open-source approaches.

- **Domain solutions with embedded explainability for business users.** Several smaller vendors offer machine learning engines with explainable AI techniques for

specific use cases. Many of them focus on high-risk or highly regulated use cases such as credit determination and hiring. While these vendors primarily focus on explainability, they can also help with bias detection — once the models are transparent, they are also auditable. For example, fraud detection provider Feedzai offers multiple explainability approaches as well as bias detection and mitigation.

Emerging Technology Use Cases And Heatmap

Explainable AI has the potential to be useful for horizontal use cases as well as vertical-specific ones. In some cases, XAI is a mandatory part of regulatory compliance; in others, it can unlock the business value of AI algorithms. The eight most important use cases for explainable AI today are customer insights, fraud detection, human resources, medical diagnosis, operational insights, personalized pricing, and public safety (see Figure 1).

Figure 1

Use Cases For Explainable AI

Use case	Timing	Users	Benefits and risks
Credit determination	Now	Data scientists, business analysts, originations teams	+ Regulatory compliance + Insights can inform future credit risk models - Perception of loss of accuracy
Customer insights	Now	Data scientists, business analysts, marketers, CX pros	+ Insights into drivers of customer behavior + Increased employee trust and adoption - Perception of loss of accuracy and reduced speed to insights
Fraud detection	Now	Data scientists, business analysts, fraud teams	+ Improved fraud detection with neural networks + Additional fraud insights - Perception of reduced speed to insights
Human resources	1 to 2 years	Data scientists, business analysts, human resources	+ Regulatory compliance (i.e., Equal Employment Opportunity Act) + Increased employee trust and adoption - Expose biases in existing practices
Medical diagnosis	2 to 4 years	Data scientists, business analysts, healthcare providers	+ Improved trust among healthcare providers + Additional medical insights
Operational insights	Now	Data scientists, business analysts, ops teams	+ Insights into drivers of asset failure, supply chain interruptions, etc. + Increased employee trust and adoption - Correlative insights do not prove causality
Personalized pricing	1 to 2 years	Data scientists, business analysts, pricing teams	+ Insights can inform future pricing and risk models + Increased employee trust and adoption + Increased customer trust - Regulatory exposure
Public safety	2 to 4 years	Data scientists, business analysts, public agencies	+ Decreased bias in policing and criminal justice + Increased public trust in agencies

Emerging Technology Heatmap

Forrester analyzed the importance of these use cases in six major industry groups and found that the adoption maturity of explainable AI varies (see Figure 2). Major areas of differences across industries can be linked to:

- **Overall AI maturity.** Explainability is a "gold-plated problem" — only businesses with a high level of AI maturity are confronted with explainability challenges. For example, many industries (e.g., manufacturing) have historically used predictive ML techniques where the models are inherently explainable. Explainability becomes a challenge as enterprises begin to leverage opaque methods like neural networks for additional predictive power. Those methods tend to be concentrated in industries where AI maturity is more advanced (e.g., financial services).

- **Regulatory exposure.** Some industries (e.g., financial services, insurance, healthcare) are more likely to have adopted XAI out of necessity for regulatory compliance. In financial services, explainable AI is widely applied to use cases such as credit determination and fraud detection, where regulators want to see how models arrive at outcomes. As more healthcare providers adopt AI for medical diagnoses and more payers adopt the technology for fraud detection, they are also gravitating toward explainability to satisfy regulatory requirements.

- **Importance of unstructured data.** Currently, opaque neural networks are the only game in town for analyzing text, images, video, and audio at scale. Industries with more use cases involving unstructured data will therefore be more inclined to invest in explainability to shed light into these models. For example, healthcare providers using computer vision to diagnose illness and insurers using NLP to detect fraud will need explainability to convey how neural networks make these critical decisions. Eventually, with the explosion of generative AI-enabled natural language interactions, all companies will need to invest in explainability.

Figure 2
Explainable AI Heatmap

Use case	Financial services	Insurance	Healthcare	Public sector	Retail	Smart manufacturing
Customer insights	Warm	Warm	Warm	Cold	Warm	Cold
Operational insights	Warm	Warm	Warm	Cold	Cold	Warm
Personalized pricing	Hot	Hot	Cold	Cold	Warm	Cold
Human resources	Warm	Warm	Warm	Cold	Warm	Warm
Credit determination	Hot	Cold	Cold	Warm	Warm	Cold
Fraud detection	Hot	Hot	Warm	Warm	Warm	Cold
Medical diagnosis	Cold	Cold	Hot	Cold	Cold	Cold
Public safety	Cold	Cold	Cold	Hot	Cold	Cold

XAI Remains Critical In The Age Of Generative AI

The explosion of genAI is fueling an unfortunate myth that we have moved beyond explainability. The large language models (LLMs) and latent diffusion models, which undergird most of today's genAI, are inherently not explainable using existing interpretability techniques. However, predictive (not generative) AI accounts for the vast majority of models that are currently tasked with decision-making in enterprises. Predictive models are ripe for explainability. Do not dismiss the importance of XAI — it is alive, well, and remains critical. Enterprises seeking explainability need to:

- **Inventory current AI implementations and classify risk.** Risk varies widely across AI use cases. To categorize risk, many companies are borrowing from the EU's AI Act which classifies AI systems into four categories — unacceptable risk, high risk, medium risk, and low risk. Enterprises can then define explainability requirements for each tier. For example, high-risk use cases may require complete transparency while interpretability may suffice for moderate risk use cases.

- **Demand explainability from vendors.** Explainability isn't only applicable to AI you build, it's also applicable to AI you buy. Beware the black box — you will be held accountable for any vulnerabilities or flaws lurking within. In partnering with AI vendors, third-party due diligence is critical and will require effort from both your procurement team and your data science team. If the vendor can't produce a detailed explanation of how their system works, run!

- **Take a holistic view of explainability.** Most of the emphasis in the world of XAI today is on individual models, but multiple models chained together into a decision system is increasingly common. Ensure that explainability goes beyond individual models and covers how the entire system works — the interoperability of all the pieces — and measure business outcomes and customer satisfaction as well as model performance to ensure the system is delivering as expected.

FORRESTER®

We help business and technology leaders use customer obsession to accelerate growth.

FORRESTER.COM

Obsessed With Customer Obsession

At Forrester, customer obsession is at the core of everything we do. We're on your side and by your side to help you become more customer obsessed.

Research

Accelerate your impact on the market with a proven path to growth.

- Customer and market dynamics
- Curated tools and frameworks
- Objective advice
- Hands-on guidance

Learn more.

Consulting

Implement modern strategies that align and empower teams.

- In-depth strategic projects
- Webinars, speeches, and workshops
- Custom content

Learn more.

Events

Develop fresh perspectives, draw inspiration from leaders, and network with peers.

- Thought leadership, frameworks, and models
- One-on-ones with peers and analysts
- In-person and virtual experiences

Learn more.

FOLLOW FORRESTER

Contact Us

Contact Forrester at www.forrester.com/contactus. For information on hard-copy or electronic reprints, please contact your Account Team or reprints@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.

Forrester Research, Inc., 60 Acorn Park Drive, Cambridge, MA 02140 USA
Tel: +1 617-613-6000 | Fax: +1 617-613-5000 | forrester.com