

# How 'Explainability' is Driving the Future of Artificial Intelligence

A Kyndi White Paper



The term “black box” has long been used in science and engineering to denote technology systems and devices that function without divulging their inner workings. The inputs and outputs of the “black box” system may be visible, but the actual implementation of the technology is opaque, hidden from understanding or justifiability.

The “black box” concept has been exploited by the likes of Silicon Valley start-ups to Wall Street investment firms, usually in their efforts to protect intellectual property and maintain competitiveness. “We’ve developed this powerful new algorithm to generate awesome results and returns for you, but don’t ask us how it works or why. Just trust us.”

But, “just trust us” is not cutting it anymore as new technologies such as artificial intelligence (AI) are seeping into virtually every facet of life. As AI becomes an increasingly essential part of how organizations of all types and sizes operate, there is a growing recognition that the old “black box” approach used by technology companies (including AI providers) is not sufficient or appropriate. The fact is, many companies doing business in highly regulated sectors as well as governmental entities that operate under constant

oversight scrutiny, need to be able to explain the “how’s” and “why’s” of AI-generated results. In many cases, the law mandates this level of openness and accountability.

**A November 2017 [commentary](#) in the Wall Street Journal outlined the growing concerns about the AI “black box”:**

*Everyone wants to know: Will artificial intelligence doom mankind—or save the world? But this is the wrong question. In the near future, the biggest challenge to human control and acceptance of artificial intelligence is the technology’s complexity and opacity, not its potential to turn against us like HAL in “2001: A Space Odyssey.” This “black box” problem arises from the trait that makes artificial intelligence so powerful: its ability to learn and improve from experience without explicit instructions.*

The MIT Technology Review recently published an [article](#) on this same topic, highlighting the growing demand for AI solutions whose results are “explainable” and auditable. The article quotes an executive from a leading financial company, who requires explainability in his AI solutions as a matter of regulatory compliance:

*Adam Wenchel, vice president of machine learning and data innovation at Capital One, says the company would like to use deep learning for all sorts of functions, including deciding who is granted a credit card. But it cannot do that because the law requires companies to explain the reason for any such decision to a prospective customer. Late last year Capital One created a research team, led by Wenchel, dedicated to finding ways of making these computer techniques more explainable.*

Ryan Welsh, Founder and CEO of Kyndi, a Silicon Valley-based AI solutions company, believes that the technology industry must step up its efforts to embrace “explainable AI” and make its results more explainable and auditable. Kyndi is building the first Explainable AI platform for government, financial services, and healthcare.

“Our mission is to build Explainable AI products and solutions that help to optimize human cognitive performance. A cornerstone of that mission is never to operate as a ‘black box,’” said Welsh. “Explainable AI means that the system can justify its reasoning. Kyndi’s product exists because Deep Learning is a ‘black box’ and cannot be used in regulated industries where organizations are required to explain the reasons for any advice on any decision.

By creating explainable AI solutions, Kyndi is also helping to mitigate the human bias that can arise in the process of extracting knowledge and answers from data.



**The Wall Street Journal’s commentary weighed in on the value of creating AI that is both accountable and explainable:**

*A better solution is to make artificial intelligence accountable. The concepts of accountability and transparency are sometimes conflated, but the former does not involve disclosure of a system’s inner workings. Instead, accountability should include explainability, confidence measures, procedural regularity, and responsibility.*

*Explainability ensures that nontechnical reasons can be given for why an artificial-intelligence model reached a particular decision. Confidence measures communicate the certainty that a given decision is accurate. Procedural regularity means the artificial-intelligence system’s*



*decision-making process is applied in the same manner every time. And responsibility ensures individuals have easily accessible avenues for disputing decisions that adversely affect them.*



## **US Government Advancing ‘Explainable AI’ Through Major DARPA Project**

The US Defense Department of Defense (DOD) is pushing Explainable AI because it cannot invest in technology “black boxes” based solely on the promise of “trust us.” The DOD’s Defense Advanced Research Projects Agency (DARPA) has responded to the growing need for greater explainability in AI by launching a major [Explainable AI research project](#). Here is how DARPA describes the rationale for its groundbreaking Explainable AI program:

*Dramatic success in machine learning has led to a torrent of Artificial Intelligence (AI) applications. Continued advances promise*

*to produce autonomous systems that will perceive, learn, decide, and act on their own. However, the effectiveness of these systems is limited by the machine’s current inability to explain their decisions and actions to human users. The Department of Defense is facing challenges that demand more intelligent, autonomous, and symbiotic systems. Explainable AI—especially explainable machine learning—will be essential if future warfighters are to understand, appropriately trust, and effectively manage an emerging generation of artificially intelligent machine partners.*

*The Explainable AI (XAI) program aims to create a suite of machine learning techniques that:*

- *Produce more explainable models, while maintaining a high level of learning performance (prediction accuracy); and*
- *Enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners.*

*New machine-learning systems will have the ability to explain their rationale, characterize their strengths and weaknesses, and convey an understanding*

*of how they will behave in the future. The strategy for achieving that goal is to develop new or modified machine-learning techniques that will produce more explainable models.*

## **Explainable AI Initiatives On the Rise Worldwide**

A recent [Wired article](#) examined how government entities across the US and around the world have come to the same conclusion as DARPA. They have realized that the old AI “black box” is neither appropriate nor, in many cases, legal, and that AI results need to be explainable and justifiable. The Wired story, “AI Experts Want to End 'Black Box' Algorithms in Government,” reported on the broad range of Explainable AI initiatives that are now cropping up around the world:

*On Sunday the UK government released a review that examined how to grow the country’s AI industry. It includes a recommendation that the UK’s data regulator develop a framework for explaining decisions made by AI systems.*

*On Monday New York’s City Council debated a bill that would require city agencies to publish the source code of algorithms used*

*to target individuals with services, penalties, or police resources.*

*On Tuesday a European Commission working group on data protection released draft guidelines on automated decision making, including that people should have the right to challenge such decisions. The group’s report cautioned that “automated decision-making can pose significant risks for individuals’ rights and freedoms which require appropriate safeguards.” Its guidance will feed into a sweeping new data protection law due to come into force in 2018, known as the GDPR.*



## **Trust and Regulatory Compliance Driving Growing Demand for Explainable AI**

To realize AI’s full potential, trust is crucial. Trust comes from understanding — and being able to justify — the reasoning behind

an AI system's conclusions and results. Kyndi believes that Explainable AI achieves the level of trust that is so important for accelerated growth and acceptance of AI. Crucially, it does so without the all too familiar "black box" approach.

For Kyndi, Explainable AI means that its software's reasoning is apparent to the user. This visibility allows them to have confidence in the system's outputs, be aware of any uncertainties, anticipate how the software will work in the future, and know how to improve the system. Such knowledge is essential to confident analysis and decision making.

Explainable AI is also necessary to provide a natural feedback mechanism so that users can tailor the results to their needs.

Because users know why the system produced specific outputs, they will also know how to make the software smarter. Using a process called "calibration," Kyndi's customers can teach the software to produce better results in the future.

Explainable AI thus becomes the foundation for ongoing iteration and improvement between human and computer.

Kyndi's novel approach to AI, which unifies probabilistic and logical methods, was built with explainability as a fundamental

requirement. A critical function of the software is to answer questions, recognize similarities, and find analogies rapidly. These features enable Kyndi to build models that are made up of a series of questions, for which the software attempts to generate answers from the data provided by customers.

Kyndi's solutions justify their reasoning by pointing to specific instances in user data and highlighting the relevant words and phrases. By providing auditability, government and enterprise users can confidently assess the results when applying them to further analysis or to make immediate decisions. All this information is readily available through Kyndi's user-friendly interface.

Kyndi's Explainable AI software is especially relevant to regulated sectors – government, financial services, and healthcare – where organizations are required to explain the reason for any decision. Because Kyndi's software logs every step of its reasoning process, users can transform regulated business functions with AI. And they will always do so with the knowledge that Kyndi's AI system allows them to justify their decisions when necessary.

## Underscoring its Explainable AI Leadership, Kyndi Named to 'AI 100' for 2018

In recognition for its leadership and innovation in Explainable AI, Kyndi was recently named to the prestigious [AI 100 for 2018](#). Sponsored by CB Insights, the Second Annual [AI 100](#) honors a select group of “promising private companies working on groundbreaking artificial intelligence technology.” Kyndi and the other AI companies selected for this year’s AI 100 were culled from a group of more than 1,000 technology firms.

Here is how CB Insights summed up Kyndi’s achievements in its recent AI 100 news release:

*Founded in 2014, Kyndi transforms business processes by offering auditable AI products. Its novel approach to AI, which unifies probabilistic and logical methods, enables organizations to analyze massive amounts of data to create actionable knowledge significantly faster and without having to sacrifice explainability. Kyndi’s Explainable AI Platform supports the following solutions: Intelligence, Defense, Compliance (i.e., for financial services and healthcare), and Research.*

Kyndi Founder and CEO Ryan Welsh commented on Kyndi’s naming to the 2018 AI 100:

“Being named to CB Insights’ AI 100 is an incredible honor. It is a major industry recognition, and I think it underscores the importance of moving past ‘black box’ machine learning towards Explainable AI products that have auditable reasoning capabilities. Explainability is especially crucial for critical organizations that are required to explain the reason for any decision.”



## Explainability is the Future of AI – Right Now

Explainability is at the core of Kyndi’s breakthrough AI products and solutions. Explainability allows users to have confidence in the AI system’s outputs, be aware of any uncertainties, anticipate how

the software will work in the future, and know how to improve the system. Such knowledge is essential to confident analysis and decision making. It's what gives Kyndi's customers a strong competitive edge.

For more information on Kyndi's Explainable AI products and solutions, visit [www.Kyndi.com](http://www.Kyndi.com) or call (650) 437-7440.



## About Kyndi

Kyndi is an artificial intelligence company that's building the first Explainable AI platform for government, financial services, and healthcare. Kyndi transforms business processes by offering auditable AI systems. Our product exists because critical organizations cannot use "black box" machine learning when they are required to explain the reason for any decision. Based in Silicon Valley, Kyndi is backed by leading venture investors.