PEGA®

Building transparency and trust into AI-powered decisioning

Artificial intelligence is reshaping industries, redefining customer experiences, and reimagining how organizations are conducting daily business operations. It's enabling organizations to be more efficient and to operationalize data in a way that may finally help companies break down the data silos that plague our tech stacks and contribute to fragmented customer experiences.

Such powerful tools also carry inherent risks. And while we haven't yet seen robots take over the world as foreshadowed by sci-fi Hollywood, AI unpredictability can still do real damage to a brand's reputation. Look no further than when Goldman Sachs was accused of gender bias discrimination in its Apple Card application algorithm. Borrowers with similar assets and credit scores were being offered different levels of credit – and it appeared that gender was a determining factor. Even though the organization has since been cleared of wrongdoing, because its algorithm was opaque and the decision-making process unexplainable, it has lost upwards of $3 billion and a priceless amount of consumer trust.

Being able to explain AI-powered decisioning, also referred to as "transparency," is one of the four pillars that defines responsible AI systems. At Pega, we define responsible AI systems as being transparent, robust, fair, and empathetic. But what exactly does transparency in AI mean, and why is it critical for businesses to incorporate it into their AI strategy?

FAIR
unbiased and balanced
for all groups

TRANSPARENT
explainable to a human
audience

EMPATHETIC
informed by customer need

ROBUST
hardened to real
world exposure

ACCOUNTABLE
fully traceable through
its lifecycle

Understanding AI transparency

At its core, transparency in AI refers to the ability to understand and trace how AI systems make decisions. It's about making the inner workings of AI algorithms clear to humans, particularly those who use, regulate, or are affected by these systems. In the above example, it would mean being able to explain why borrower A received a credit limit of $10,000 while borrower B received a credit limit of $100,000.

These systems learn from vast amounts of data, often making decisions in ways that are not inherently clear, even to their creators. When an AI algorithm operates as a "black box" we call this opaque AI. AI systems can inadvertently perpetuate and amplify biases present in their training data. And if left unchecked, these systems may reinforce those biases. When AI-powered decisions are trained with data points like age, race, gender, sexual orientation, or geography, they may inadvertently favor certain demographics over others, highlight preference for a certain gender or race in historical data, or introduce bias from humans where data needs to be labeled.

For example, the 2019 "Women in Stem" study commissioned by the Equal Employment Opportunity Commission highlights the disparity between the number of women versus men in STEM fields. If an AI-driven recommendation system that suggests jobs to candidates was based on just the underlying data of this study, the algorithm may start to pick up gender-stereotypical patterns. This, in turn, might lead to fewer women being recommended STEM-related jobs, thereby reinforcing the gender disparity.
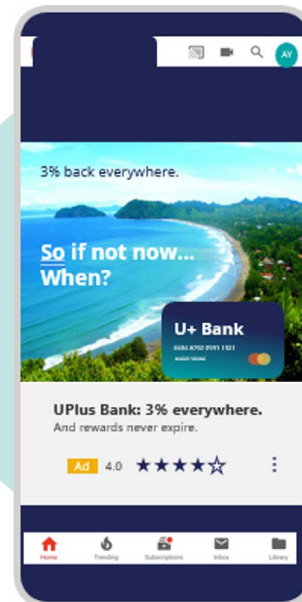
Transparency allows for the examination and understanding of how these biases may occur, leading to more ethical and fair AI systems.

CAN YOU EXPLAIN THIS AI-POWERED DECISION?

Jenna
Employed
Female
Credit
Score 702

Offer

3% back everywhere.

So if not now... When?

U+ Bank

UPlus Bank: 3% everywhere.
And rewards never expire.

Ad   4.0   ★★★★☆

Jason
Employed
Male
Credit
Score 702

No offer

Why is transparency important?

Transparency builds trust with consumers, employees, and stakeholders. When users understand how and why an AI system makes decisions, they are more likely to trust and accept it. But the level of AI opacity may vary depending on the industry. For example, in highly regulated industries, transparency is paramount for compliance with the law. However, the regulatory environment is often much slower than the speed of innovation and various regions promote different governing strategies. For example, in the U.S. there's the possibility that 50+ different privacy laws could govern AI depending on the legislative appetite in each state. Whereas in Europe, there is a consensus approach between EU member states.

If a regulatory standard doesn't compel businesses to be transparent, what will? The answer is ethics – and that's where empathy comes in. If transparency is part of an organization's core values, and if it's incorporated into AI strategies, the organization is demonstrating empathy for customers and stakeholders because the business prioritizes fairness, respect, and privacy, which is in the best interest of us all.

Ethics also play a crucial role in a company's revenue. Customers are essential for any business, and their trust in a brand is fundamental. This trust can be significantly impacted when it comes to the use of AI by businesses. If consumers believe that AI systems are biased or make unfair decisions,

they may be hesitant to engage with companies that rely on this technology. Therefore, maintaining ethical standards in AI applications is not just about doing the right thing; it's also about sustaining customer trust and, by extension, the business' financial health.

Pega partnered with research firm Savanta to study consumers' perceptions of AI. More than 5,000 people worldwide were surveyed on their views on AI, its continued evolution, and the ways in which people interact with the technology. Resulting data confirmed consumer suspicions of both brands and AI:

- Consumers do not believe that brands have their best interests at heart.

- They do not trust AI to make decisions that impact their life, like lending, as much as they would a human.

- And lastly, that businesses have an ethical responsibility to behave morally beyond what is legal.

These findings demonstrate the need for businesses to demonstrate transparency and fairness in both their technology application and commitment to consumer protection overall.

**Do you believe that businesses you interact with truly have the customer's best interests at heart?**

40%

60%
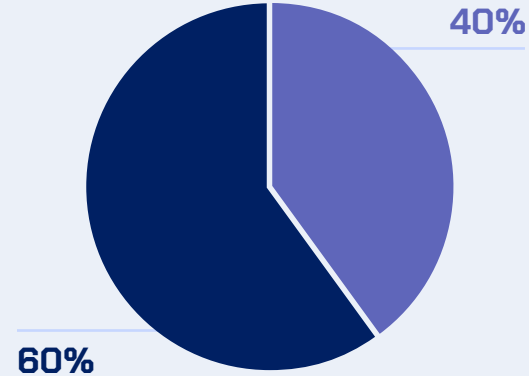
- Yes
- No

Does an organization have an obligation to do what is morally right for the customer, beyond what is legally required?

26%

74%
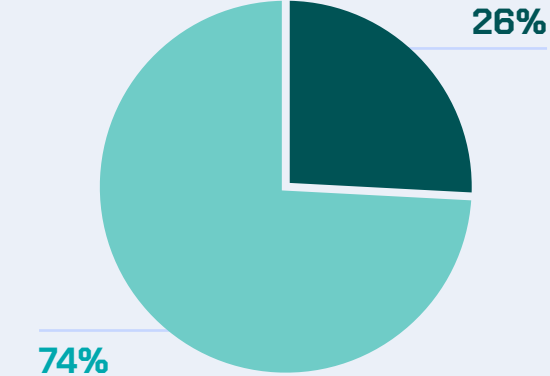
- Yes - they have a responsibility to do what is morally right
- No - their responsibility stops at what is legally required

Who would you trust more to make an objective, unbiased decision about whether or not to give you a bank loan?

32%

68%
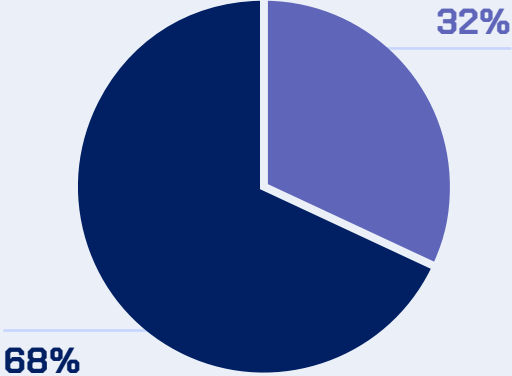
- A human bank employee
- An AI solution for banking

Challenges to achieving transparency in AI

Developing more explainable AI models is the core tactic for achieving transparency, but that's often easier said than done. And many view AI models and algorithms as a secret sauce that, if exposed, would be just as bad as giving up your competitive advantage. Algorithms can be classified by some as intellectual property.

There's also a relationship between opacity and predictive power. More opaque models are often more powerful. Marketers can look at this as a similar comparison to the relationship between audience reach and accuracy in data-driven campaigns: The wider an audience is the less relevant messaging might be. Whereas if an audience is more granular, you may reach fewer people, but the messaging may resonate more. Think podcast advertisement versus SMS text message. It's a tradeoff we must analyze against our goals and budgets.

When it comes to statistical and machine learning models, they range from simple and transparent to complex and opaque. Some AI models are incredibly complex and difficult to interpret. Two examples of more complex models are Deep Neural Networks (DNN) and Bayesian networks. Some examples of technology that uses DNN are voice assistants like Siri and Alexa, recommendation algorithms like those used by Netflix and YouTube, language translation services, and self-driving cars.

Examples of more simplistic models are linear regression and decision trees. A decision tree can be made on a simple piece of paper by someone who isn't a data scientist. I could make one right now to figure out where I'm going to go on vacation this year. It's extremely easy to see the decision path and how you arrive at an outcome. Decision trees can be used for medical diagnosis and loan approval processes, and linear regression is used in credit scoring and real estate pricing.

You'll notice the tradeoff between accuracy and opacity. My Netflix recommendations are going to be a lot more accurate, more of the time, than a human using a decision tree to figure out if I have a virus. And though there is an algorithm that's widely used for real estate appraisals, the process varies based on factors outside of the model, including who's performing the evaluation.
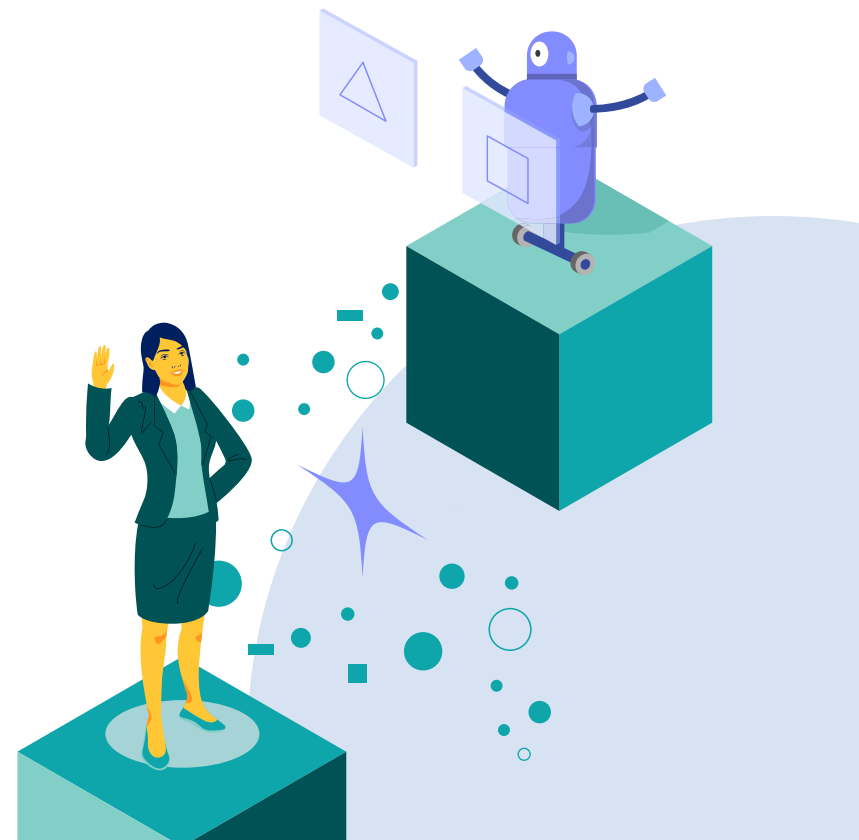
Strategies for enhancing transparency and explainability

Integrating explainability considerations into your AI systems from the beginning of the development process is one of several ways to build transparency into your organization. Below is an overview of a holistic set of steps to enhance transparency across the enterprise.

AI TRANSPARENCY CONSIDERATIONS

Opaque AI: AI that is powerful but often difficult to explain
Example: Pega Customer Decision Hub Adaptive Gradient Boosting

Transparent AI: AI that is less powerful but easier to explain
Example: Pega Customer Decision Hub Naïve Bayes Adaptive Learning

AI Guardrails: Controls that mandate the best tool for the job
Example: Pega Customer Decision Hub Transparency T-Switch

AI Explanations: Tools to quickly explain high-risk decisions
Example: Pega Customer Decision Hub Customer Profile Viewer

CREATE A TRANSPARENT CULTURE

Algorithmic transparency goes hand in hand with creating an organizational culture that strives for transparency. Accountability should be shared – not just taken on by technologists, but from all functional areas like marketing, operations, sales, customer service, and beyond. Making this value part of the culture reinforces its importance to business. Educate employees about policies and processes, like data collection, to increase their knowledge around best practices and convert them into transparency stewards.

AUDIT AND OVERSEE FOR ACCOUNTABILITY

Continuous monitoring by a human in the loop to oversee AI-powered decisions and performance is essential in maintaining transparency. If there is a problem or bias emerges, a human auditor can catch it early before it's reinforced over and over. Data freshness and recency should also be a consideration; use the most recent data to make the most accurate and relevant decisions for customers. Businesses should also clearly state and publicize how data is collected, used, processed, cleaned, and handled – because our AI systems are only as fair and accurate as the data we feed into them. Not only does this enhance transparency, but it also enhances consumer trust. Most organizations who handle consumer data post their privacy policies online, and if we do the same with AI governance policies, this may further build trust and foster adoption.

## UTILIZE TECHNOLOGY THAT HAS MONITORING FEATURES AND TOOLS.

As AI becomes more integrated into the enterprise, innovation has produced tools to help make sure we can explain how our systems work and catch issues before they become business-impacting problems. At Pega, we have transparency-enhancing capabilities within our AI-powered Pega Customer Decision Hub™. Brands like Wells Fargo, Verizon, T-Mobile, Etisalat, Citi, and beyond use Customer Decision Hub as a central brain to power relevant, perfectly timed customer interactions in real time. With both adaptive and predictive modeling capabilities, the AI is capable of sensing customer data across brand channels, learning from it instantly and then delivering the customer the next best action in that customer's journey according to their needs in the moment. Customer Decision Hub helps brands inject transparency into their strategy in two ways:

The Transparency Switch (T-Switch) is a guardrail that enables brands to deploy AI algorithms based on transparency thresholds required for their business needs across products or functional areas. With the T-Switch, brands can reduce opacity to mitigate risks, maintain regulatory compliance, and provide differentiated experiences safely to their customers.
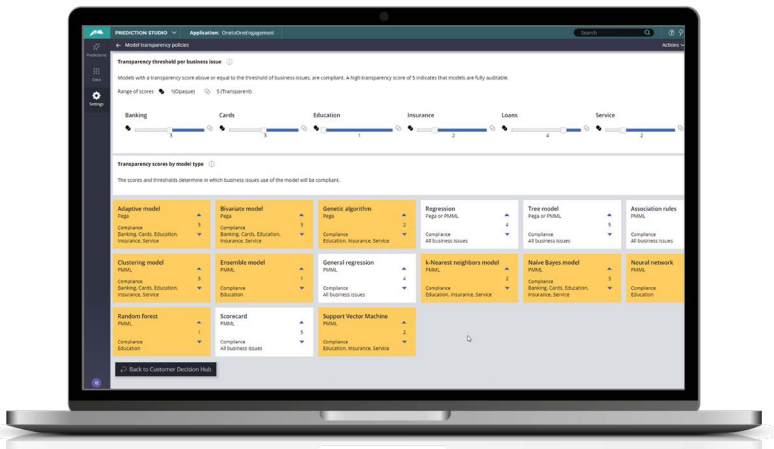
Customer Profile Viewer provides a comprehensive view of customer behavior and the detailed insights that explain how every AI-powered decision is made. Through suitability, eligibility, propensity to accept, and business goals, every decision calculation is visible to users, showcasing why each customer received each action.
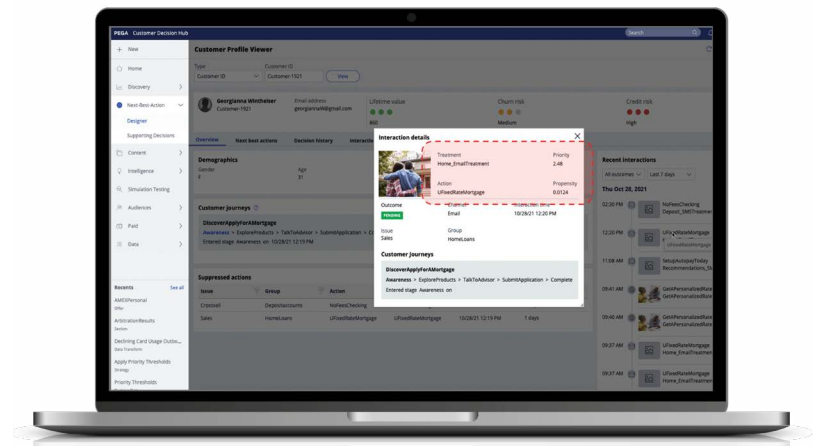
Beyond the responsibility of each organization to achieve transparency, setting industry standards is also important and achievable. This requires organizations to come together and develop a framework for responsible AI best practices by establishing agnostic organizations, like the Responsible AI Institute, which develop and maintain standards, offer benchmarking, and conduct research to measure adherence to such frameworks.

As AI continues to integrate into more aspects of enterprise operations and the everyday lives of consumers, transparency will be critical to unlocking its full potential. It's central to building consumer trust, ensuring fairness for marginalized groups, and meeting regulatory standards across many industries. While technologists are still solving challenges that contribute to the opacity of AI algorithms, we can simultaneously come together to create accountable cultures, best practices, and agreed-upon frameworks in pursuit of more widespread explainability.

SPOTLIGHT ON ORGANIZATIONAL TRANSPARENCY

T-Mobile chose the Pega Platform™ and Pega Customer Decision Hub™ to harness the power of AI and address adoption challenges. The company implemented next-best-action (NBA) capabilities within Customer Decision Hub to provide its representatives with the right action to strengthen customer relationships through proactive service.

T-Mobile's approach focused on building trust with frontline users by emphasizing that AI is meant to enhance their capabilities, not replace them. The framework provided suggested wording to assist with proactive conversations and showed how AI could optimize the investment of time during visual audits. Developed program reporting that went beyond revenue, to include cost and customer experience metrics, allows for trust-building conversations with operations leaders about where investments would be made and where benefits would be realized.

> **The journey we've been on is getting those stakeholders in and trusting the technology, helping them see what it does, making sure that they're part of it.**
>
> **This isn't a black box. We welcome all teams to work with us directly. We share our information, everything's transparent. This is a journey that we are on as a company. This is everybody at T-Mobile trying to put the customer at the center of what we do.**

BRIAN SPRIGINGS
DIRECTOR – AI STRATEGY, T-MOBILE

To see their results visit: https://www.pega.com/customers/t-mobile-drives-ai-customer-decision-hub