Tutorial: XAI or Explainable Artificial Intelligence and how it can be applied to a software coding project

Introduction:

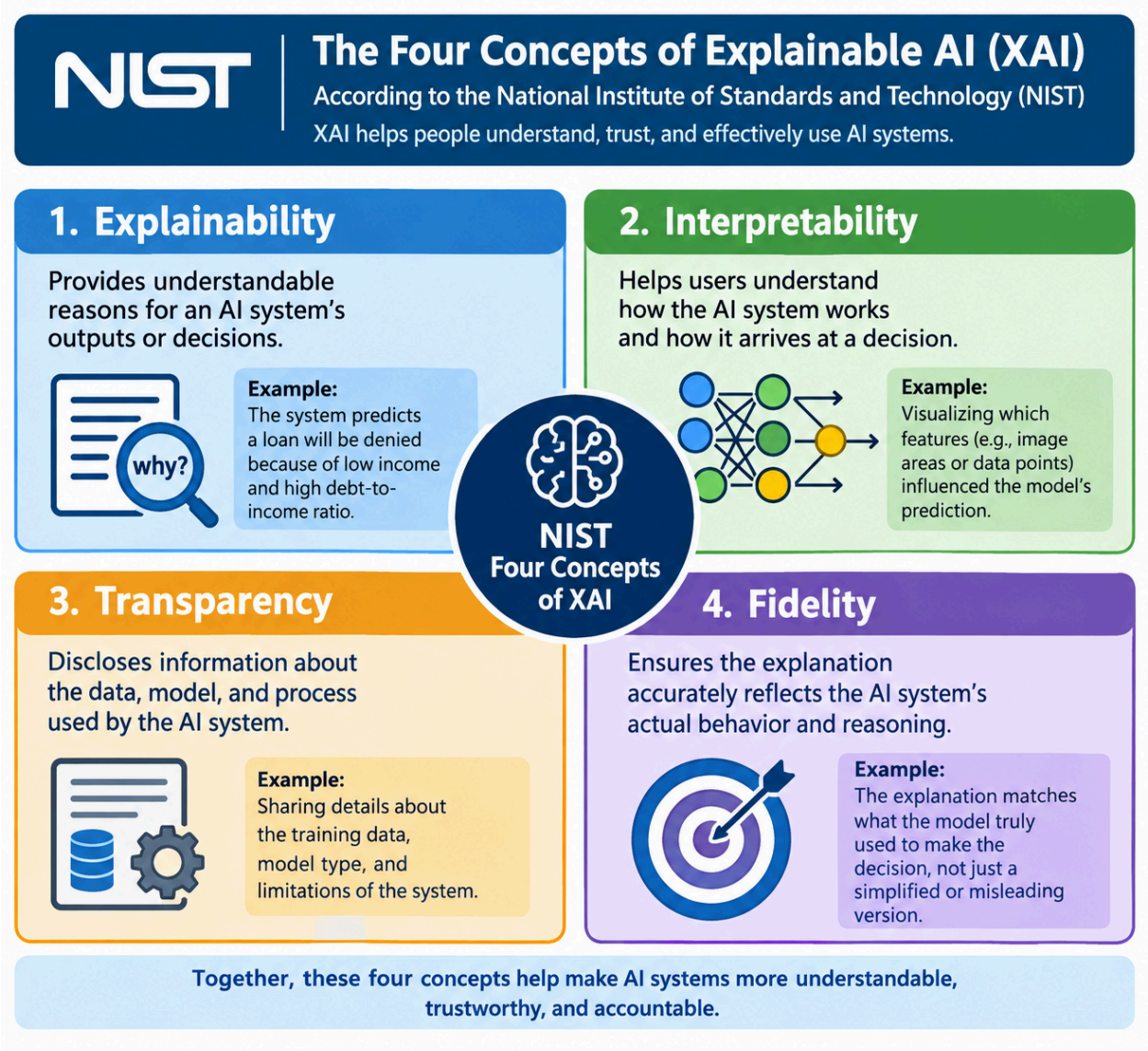
Artificial Intelligence can make estimations, predictions, develop software code, find patterns in code and network functions, and automate difficult tasks or digital chores.

How did Artificial Intelligence have the ability to answer your questions?

XAI commonly called Explainable AI or Explainable Artificial Intelligence can answer this inquiry. Scientists developed Explainable AI to assist computer professionals in understanding the logic, evidence, and concepts behind an AI system's actions and results.

The National Institute of Standards and Technology has four concepts of explainable Artificial Intelligence which are:

- Explanation
- Meaningfulness
- Explanation accuracy
- Knowledge limits



Software developer can implement XAI in two basic ways:

1. Explaining decisions and logic created by AI applications.
2. Explaining and assessing code developed by AI programming tools.

An in-depth definition of Explainable AI

Legacy software usually uses instructions created by a human software programmer.

For example, let's analyze how a person would write code to check a condition:

```
Age = 25
  If   Age >= 18;
          Print (“Adult”)
 else:
          Print (“Minor”)
```

A human software engineer can simply describe why the program produces this outcome.

The reason is simple:

```
IF          Age >=18

             |
             V
           Adult
```

Machine-learning or Artificial Intelligence systems can be more complex than human thought.

An AI machine learning agent might receive:

- Age
- Income
- Credit History
- Debt
- Employment History
- Payment History

and generate the following logic:

Prediction: This person has a low credit risk.

The question become hard to answer when you ask yourself:

Why did the AI agent model say that this individual has a low credit risk?

Explainable Artificial Intelligence or XAI was developed by engineers and scientists to provide a comprehensive explanation.

Why is AI having the ability to explain processes and code is important in software development?

Debugging

When an AI agent model develops an unexpected outcome, software engineers need to understand why those outcomes happened.

Trust

People and system users are more trusting of an AI agent model when it has the ability to explain its logic and decisions.

Security

When Artificial Intelligence explains itself software developers and engineers can identify unusual behavior or vulnerabilities in software code or decisions on an information system.

Compliance

Some software applications will make it mandatory for organizations to note how AI systems make informed decisions.

Quality Control

Software engineers and developers can use Explainable Artificial Intelligence or XAI to decide whether an AI system is using the necessary data to make this kind of decision for an information system.

Who Invented Explainable AI?

XAI was not exactly invented by one person. This technological innovation was developed over several decades of research and scientific development in order to make computer and machine learning logic or choices comprehensive to people.

XAI developmental Milestones:

- **1970's-1980's:** Researchers in professional information systems and machine learning created technological infrastructures that could describe their decisions. Example. The MYCIN project at Stanford are early instances of AI that can assist with explanations for its implications.
- **1990's:** Research scientists implemented an accelerated effort on studies on interpretable and explainable machine learning systems.
- **2010's:** The scientific researchers at this time coined the term "Explainable AI" (XAI) became a prominent term in the Computer Science field. DARPA implemented an Explainable Artificial Intelligence (XAI) program in the year 201 and assisted in building XAI asa major area of scientific research.

- Marco Tulio Ribiero, Saneer Singh, and Carlos Gusetrin created LIME in 2016 which is a highly used procedure for explaining industrial artificial intelligence predictions and estimations.
- Scott Lundbeg and Su-InLee created SHAP in 2017, which is a widely used method for designing machine learning decisions.

Conclusion

Explainable Artificial Intelligence or XAI is a crucial step in the creation of machine learning systems that human beings can understand, assess, and respect. Instead of just writing random answers, XAI assists people in comprehending why a machine learning system reached a particular outcome or answer to avoid the feeling that the responses are not just random answers without justification.

Milestones of Developing Explainable AI (XAI)

From early expert systems to modern AI explanation techniques

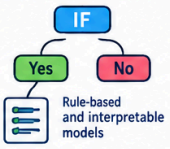
1970s – 1980s
Early Expert Systems

Researchers in expert systems and artificial intelligence developed systems that could explain their reasoning.

The **MYCIN** project at Stanford is a well-known early example.

First steps toward AI explanations

1990s
Interpretable Machine Learning



Researchers increasingly studied interpretable and explainable machine-learning systems, including decision trees, rule-based models, and simpler models that were easier to understand.

Focus on transparency and interpretability

2016
DARPA's XAI Program

DARPA launched the Explainable Artificial Intelligence (XAI) program to advance research in understanding and explaining AI decisions.

Helped establish XAI as a major research area

2016
LIME

Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin developed LIME (Local Interpretable Model-agnostic Explanations) for explaining individual machine-learning predictions.

Explained individual predictions

2017
SHAP

Scott Lundberg and Su-In Lee introduced SHAP (SHapley Additive exPlanations), a powerful method for explaining machine-learning predictions.

Provided consistent and accurate explanations

2020s – Present
Wider Adoption

XAI is now widely used in education, healthcare, cybersecurity, business, transportation, and many other fields.

Research continues to improve fairness, accuracy, and usability of AI explanations.

More trustworthy, responsible, and human-centered AI

The Journey Continues

XAI did not come from a single inventor, but from decades of research and collaboration. From early expert systems to modern techniques like LIME and SHAP, each milestone has helped make AI more transparent, understandable, and trustworthy.

Understand the why behind the AI.

References

Chamola, V., Hassija, V., Sulthana, A. R., Ghosh, D., Dhingra, D., & Sikdar, B. (2023). A review of trustworthy and explainable artificial intelligence (XAI). *IEEe Access*, *11*, 78994-79015.

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., ... & Ranjan, R. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM computing surveys*, *55*(9), 1-33.

Phillips, P. J., Phillips, P. J., Hahn, C. A., Fontana, P. C., Yates, A. N., Greene, K., ... & Przybocki, M. A. (2021). Four principles of explainable artificial intelligence.

Zodage, P., Harianawala, H., Shaikh, H., & Kharodia, A. (2024). Explainable AI (XAI): History, basic ideas and methods. *International Journal of Advanced Research in Science, Communication and Technology*, *3*.