Chapter 3
Explainable Artificial Intelligence as a Cybersecurity Aid

Ruchi Doshi
https://orcid.org/0000-0002-7259-8481
Universidad Azteca, Mexico

Kamal Kant Hiran
Symbiosis University of Applied Sciences, India

ABSTRACT

Within the span of just a few short years, artificial intelligence (AI) methods have spread across every facet of modern society. Even though AI models produce results, those results are often not easily explicable. XAI, or "explainable artificial intelligence," is a rapidly expanding area of study that aims to maximise the clarity of data extraction and visualisation processes. At the heart of the current investigation is an examination of the connections between cybersecurity and the application of XAI. The increasingly sophisticated and automated nature of attacks necessitates similarly mechanised approaches to defence. Due to its unique characteristics, XAI is suitable for this purpose. Cybersecurity is the practise of safeguarding computer systems, networks, and software from intrusion. An enormous amount of hope rests on the shoulders of XAI for foreseeing such assaults.

1. INTRODUCTION

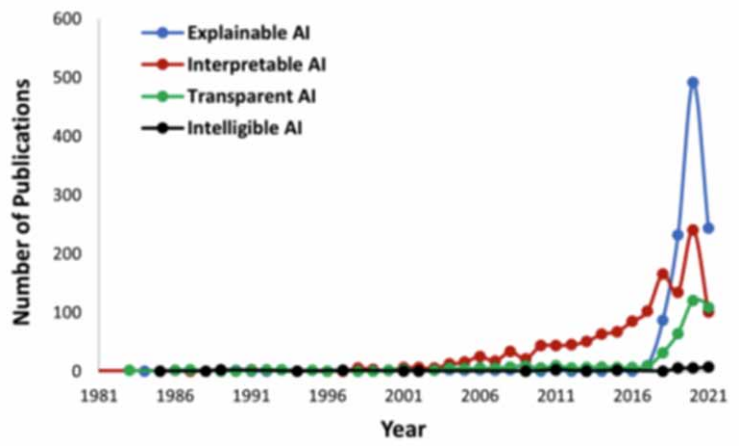
Human beings possess a cognitive ability, such as learning and problem solving, that has been investigated for several years through different studies and experiments in order to understand how it works, i.e., how we comprehend, predict or manipulate what's around us. Artificial Intelligence (AI), however, is not just limited to understanding but also to building intelligent entities.

DOI: 10.4018/978-1-6684-6361-1.ch003

Certain security properties are guaranteed by Cybersecurity in order to avoid possible risks in cyberspace, such as integrity, availability, and confidentiality. However, Cybersecurity is not just focused on protecting cyberspace but also on protecting whatever operates within cyberspace and any of its assets that may have a direct or indirect relationship with cyberspace. AI is gradually being integrated into business, education, medicine, and other fields. It is widely employed in a variety of application scenarios. However, not all sectors have reached the same level of development. In terms of implementing artificial intelligence, the information technology and telecommunications sectors are the most advanced (Easwaran et al., 2022).

According to a worldwide survey [survey] of over 5,000 technology stakeholders from various industries, 46 percent of large companies and 38 percent of SMEs have already incorporated artificial intelligence in their organizations (Colaner, 2022). It is also noticeable the increasing amount of research regarding accountability of artificial intelligence to assure trustworthy decision in the latest decade. Annual publications represent the rapid and widespread rise of XAI, Interpretable, Intelligible, and Transparent AI, with XAI first appearing in 2017 alongside the US DoD DARPA XAI initiative, as observable in Figure 1.

Figure 1. Relationship with grap XAI evolution



1.1 Cybersecurity

Any business should be concerned about cybersecurity. The weakest link in a security chain is frequently the human element, not computers or other technology. Users can make mistakes because of risky behaviours, ignorance of the online world, or disagreement with security protocols (Hiran et al., 2014). Users may intentionally or unintentionally expose sensitive data to criminal third parties, which allows those parties to steal organisational data that was not initially exposed. Hackers target technology users who have access to information systems in order to breach the networks due to an increase in intense social engineering intrusions. Therefore, educating and training computer users to consider security-conscious performance as an integral component of their work will have a favourable impact on the operation's overall information security (Hariharan et al., 2021; Nekhai et al., 2022).