

# Designing a Conceptual Model of Explainable Artificial Intelligence (XAI) Based on the clarification jihad Approach

Mohammad Elyasi<sup>1✉</sup> 

1. Former PhD Student in Tourism, Master's Degree in Tourism Marketing, Email: [M.Elyasi@ut.ac.ir](mailto:M.Elyasi@ut.ac.ir)

## Article Info

**Article type:**  
Research Article

**Article history:**

**Received:** 13 March 2025

**Received in revised form:**

22 May 2025

**Accepted:** 1 July 2025

**Available online:**

23 August 2025

**Keywords:**

Explainable Artificial Intelligence (XAI), clarification jihad, Supreme Leader's Statements, Conceptual Model, Cognitive Defense.

## ABSTRACT

This study seeks to integrate the technical capabilities of Explainable Artificial Intelligence (XAI) with the strategic principles of clarification jihad, aiming to provide a transparent tool for clarifying decisions of intelligent systems, neutralizing misinformation, and enhancing the cognitive resilience of society. Adopting a problem-oriented approach, the study addresses two contemporary challenges: first, the "black-box" nature of AI systems; and second, the threat of cognitive warfare through the dissemination of fake news and information distortion. The research employs a qualitative-integrative and applied-developmental methodology. Data were collected through a systematic literature review, thematic analysis of the Supreme Leader's statements from 2016 to 2025, and 10 semi-structured interviews with experts. Data analysis utilized reflexive thematic analysis and the design of a dynamic three-layer model.

The findings resulted in a three-layer conceptual model:

1. Foundational Layer – clarification jihad: encompassing the mission of clarification, effective methods (explicitness, reasoning, audience analysis), and authentic religious content;
2. Intermediate Layer – Explainable AI: including technical-ethical capabilities (transparency, explainability, bias control) and technical requirements (indigenous knowledge bases, anti-bias systems);
3. Operational Outer Layer: featuring a human-centered approach, local infrastructure, and a management and evaluation system based on indicators and feedback cycles.

The value-to-technology engine, aiming to translate clarification jihad principles into XAI requirements, alongside a combined jurisprudential-technical governance and feedback-driven lifecycle, transforms the clarification jihad concept from an abstract idea into an operational, value-oriented, dynamic, indigenous, and defensible system. This model provides an effective framework for countering cognitive warfare and advancing responsible AI within an Islamic discourse

**Cite this article:** Elyasi, M. (2025). Designing a Conceptual Model of Explainable Artificial Intelligence (XAI) Based on the clarification jihad Approach. *Digital Islamic Studies and Humanities*, 1 (1), 311-354. <https://doi.org/10.22034/disah.2025.2076072.1054>



© The Author(s). **Publisher:** Research Center for Digital Islamic Studies and Humanities (RCDISAH).

**DOI:** <https://doi.org/10.22034/disah.2025.2076072.1054>

## طراحی مدل مفهومی هوش مصنوعی XAI با رویکرد جهاد تبیین

محمد الیاسی<sup>۱</sup> ✉

۱. دانشجوی سابق دکتری گردشگری، کارشناس ارشد بازاریابی گردشگری، رایانامه: M.Elyasi@ut.ac.ir

### اطلاعات مقاله

### چکیده

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت: ۱۴۰۳/۱۲/۲۳

تاریخ بازنگری: ۱۴۰۴/۰۳/۰۱

تاریخ پذیرش: ۱۴۰۴/۰۴/۱۰

تاریخ انتشار: ۱۴۰۴/۰۶/۰۱

کلیدواژه‌ها:

هوش مصنوعی توضیح‌پذیر (XAI)،

جهاد تبیین،

بیانات رهبری،

مدل مفهومی،

پدافند شناختی.

پژوهش حاضر در پی تلفیق قابلیت‌های فنی هوش مصنوعی توضیح‌پذیر (XAI) با مبانی راهبردی «جهاد تبیین» است تا ابزاری شفاف برای روشن‌سازی تصمیم‌های سامانه‌های هوشمند، خنثی‌سازی شبهات و تقویت تاب‌آوری شناختی جامعه فراهم آورد. این مطالعه با رویکرد مسئله‌محور، دو چالش اساسی معاصر را مدنظر قرار داده است: نخست، ماهیت «جعبه‌سیاه» سامانه‌های هوش مصنوعی؛ و دوم، تهدید جنگ شناختی از طریق انتشار اخبار جعلی و تحریف اطلاعات. روش پژوهش، کیفی-تلفیقی و کاربردی-توسعه‌ای است. داده‌ها از طریق مرور نظام‌مند ادبیات، تحلیل تماتیک بیانات مقام رهبری در بازه زمانی ۱۳۹۵ تا ۱۴۰۴، و انجام ۱۰ مصاحبه نیمه‌ساختاریافته با خبرگان گردآوری شده‌اند. تحلیل داده‌ها با بهره‌گیری از روش تحلیل تماتیک بازتابی و طراحی مدل سه‌لایه پویا صورت گرفته است. نتایج پژوهش منجر به ارائه یک مدل مفهومی سه‌لایه شده است: ۱. لایه بنیادین «جهاد تبیین» شامل رسالت تبیین، شیوه‌های مؤثر (صراحت، استدلال، مخاطب‌شناسی) و محتوای اصیل دینی؛ ۲. لایه میانی «هوش مصنوعی توضیح‌پذیر» شامل قابلیت‌های فنی-اخلاقی (شفافیت، قابلیت توضیح، کنترل سوگیری) و الزامات فنی (پایگاه دانش بومی، سامانه‌های ضدسوگیری)؛ ۳. لایه بیرونی عملیاتی شامل رویکرد انسان‌محور، زیرساخت بومی، و نظام مدیریت و ارزیابی مبتنی بر شاخص‌ها و چرخه بازخورد. موتور تبدیل ارزش به فناوری، با هدف ترجمه مفاهیم جهاد تبیین به الزامات هوش مصنوعی توضیح‌پذیر، در کنار حاکمیت ترکیبی فقهی-فنی و چرخه حیات بازخوردمحور، مدل جهاد تبیین را از یک مفهوم انتزاعی به سامانه‌ای عملیاتی، ارزش‌محور، پویا، بومی و قابل دفاع تبدیل می‌کند. این مدل، چارچوبی مؤثر برای مقابله با جنگ شناختی و توسعه هوش مصنوعی مسئولیت‌پذیر در گفتمان اسلامی ارائه می‌دهد.

استناد: الیاسی، محمد (۱۴۰۴). طراحی مدل مفهومی هوش مصنوعی XAI با رویکرد جهاد تبیین. علوم انسانی و اسلامی

دیجیتال، ۱ (۱)، ۳۱۱-۳۵۴. <https://doi.org/10.22034/disah.2025.2076072.1054>



ناشر: پژوهشگاه علوم اسلامی و انسانی دیجیتال (مرکز تحقیقات کامپیوتری علوم اسلامی نور). © نویسندگان.

## مقدمه

در دهه‌های اخیر، هوش مصنوعی (AI) به ستون فقرات تصمیم‌گیری‌های حیاتی در حوزه‌هایی چون پزشکی، مالی، دفاع و امنیت تبدیل شده است. با این حال، ماهیت «جعبه سیاه» مدل‌های پیشرفته - به ویژه شبکه‌های عصبی عمیق - درک منطق تصمیم‌گیری‌ها را دشوار ساخته و چالش‌های بنیادینی در زمینه‌های اعتمادپذیری، شفافیت، عدالت و پاسخ‌گویی ایجاد کرده است (Gunning & Aha, 2019). در پاسخ به این چالش‌ها، حوزه پژوهشی «هوش مصنوعی توضیح‌پذیر» (XAI) با هدف توسعه روش‌هایی قابل فهم و کنترل‌پذیر برای فرآیندهای تصمیم‌گیری سامانه‌های هوشمند شکل گرفته است (Dwivedi et al., 2023).

هم‌زمان با این تحولات، در گفتمان جمهوری اسلامی ایران، راهبرد «جهاد تبیین» بر اساس بیانات مقام معظم رهبری به‌عنوان فریضه‌ای اساسی برای مقابله با جنگ نرم، تحریف‌های رسانه‌ای و شبهه‌افکنی‌های سازمان‌یافته معرفی شده است (خامنه‌ای، ۱۳۹۸). این راهبرد بر روشنگری فعال، مستدل و اقناع‌گرایانه تأکید دارد تا ادراک عمومی را در برابر هجوم‌های شناختی حفظ کند. پژوهش حاضر استدلال می‌کند که تلفیق هم‌افزایانه قابلیت‌های فنی XAI (شفاف‌سازی فناورانه) با مبانی راهبردی جهاد تبیین (روشنگری فرهنگی - رسانه‌ای) می‌تواند چالش‌های اخلاقی و بحران اعتماد در حوزه هوش مصنوعی را مرتفع سازد، ابزاری مؤثر برای پدافند غیرعامل شناختی فراهم آورد و با ایجاد شفافیت در سامانه‌های اطلاعاتی، بستری برای تبیین دقیق‌تر حقایق فراهم کند.

مسئله اصلی این تحقیق، چالش دوگانه فنی - اخلاقی و راهبردی - امنیتی است. ماهیت غیرشفاف مدل‌های پیشرفته هوش مصنوعی موسوم به «جعبه سیاه»، اعتماد ذی‌نفعان را تضعیف کرده و ارزیابی معیارهایی چون عدالت، تشخیص سوگیری‌ها و ریشه‌یابی خطاها را دشوار می‌سازد. این مسئله در حوزه‌های حساس مانند قضاوت قضایی، پزشکی، امنیت ملی و دین و مذهب، پیامدهایی جبران‌ناپذیر به همراه دارد. از سوی دیگر، فضای مجازی در عصر دیجیتال به عرصه اصلی جنگ شناختی تبدیل شده است؛ جایی که بازیگران متخاصم از طریق اخبار جعلی، اطلاعات گمراه‌کننده و تحریف واقعیت‌ها، ادراک عمومی را هدف قرار می‌دهند. بر اساس بیانات مقام معظم رهبری، راهبرد جهاد تبیین برای مقابله با این هجوم‌ها بر تبیین فعال، مستدل و تهاجمی تأکید دارد.

ضرورت انجام این پژوهش از سه جنبه قابل بررسی است:

- جنبه علمی و نظری: ایجاد تلفیقی میان رشته‌ای و نوآورانه میان مبانی نظری هوش مصنوعی توضیح‌پذیر و اصول علوم انسانی-اسلامی راهبرد جهاد تبیین، به تولید دانش نوین کمک می‌کند.
- جنبه کاربردی و فناورانه: پژوهش حاضر، مبنایی علمی برای توسعه ابزارها و سامانه‌های هوشمند بومی فراهم می‌آورد؛ از جمله چت‌بات‌های تبیین‌گر هوشمند برای پاسخ‌گویی مستدل به شبهات، سامانه‌های رصد رسانه‌ای برای شناسایی و تحلیل اخبار جعلی و جریان‌های تحریفی، و پلتفرم‌های تصمیم‌یار شفاف برای نهادهای حاکمیتی.
- جنبه راهبردی و ملی: این تحقیق با هم‌سویی کامل با تأکیدات مقام معظم رهبری بر فریضه جهاد تبیین و ضرورت پیشرفت در حوزه هوش مصنوعی، می‌تواند به‌عنوان سندی پشتیبان برای سیاست‌گذاری‌های کلان در زمینه «هوش مصنوعی مسئولیت‌پذیر، اخلاق‌مدار و بومی» مورد استفاده قرار گیرد. نتایج آن در تقویت پدافند شناختی کشور در برابر تهدیدات سایبری و جنگ نرم نقش‌آفرین خواهد بود.

هدف اصلی پژوهش، طراحی مدل مفهومی برای به‌کارگیری XAI در چارچوب جهاد تبیین به‌منظور افزایش شفافیت سامانه‌های هوش مصنوعی و عملیاتی‌سازی مقابله با جنگ شناختی است. اهداف فرعی شامل موارد زیر هستند:

- شناسایی مؤلفه‌ها و ابعاد کلیدی XAI از طریق مرور نظام‌مند ادبیات؛
  - استخراج مؤلفه‌ها و شاخص‌های جهاد تبیین از متون معتبر (بیانات رهبری)؛
  - تحلیل نقاط هم‌پوشانی و ظرفیت‌های تلفیق XAI و جهاد تبیین؛
  - تدوین چارچوب مفهومی تلفیقی برای پیاده‌سازی عملیاتی جهاد تبیین مبتنی بر هوش مصنوعی توضیح‌پذیر؛
  - ترسیم نقشه راه پژوهشی و کاربردی برای توسعه مدل در حوزه پدافند شناختی.
- پرسش اصلی: چگونه می‌توان مدل مفهومی مبتنی بر XAI با رویکرد جهاد تبیین طراحی کرد تا هم شفافیت تصمیم‌های هوش مصنوعی را افزایش دهد و هم به ابزاری کارآمد برای رفع شبهات، تبیین حقایق و مقابله با جنگ شناختی تبدیل شود؟

پرسش‌های فرعی: ۱. مؤلفه‌ها و ابعاد کلیدی XAI در ادبیات موضوع کدام‌اند؟ ۲. مؤلفه‌ها و شاخص‌های جهاد تبیین به‌عنوان راهبرد روشنگری چیست؟ ۳. نقاط هم‌پوشانی میان قابلیت‌های XAI و الزامات جهاد تبیین کدام‌اند؟ ۴. چارچوب مفهومی ترکیبی XAI و جهاد تبیین چگونه باید ترسیم شود؟ ۵. نقشه راه پژوهشی و عملیاتی برای توسعه این مدل در پدافند شناختی چیست؟ این پژوهش از منظرهای زیر دارای نوآوری و ارزش افزوده است:

- نوآوری نظری: ارائه نخستین چارچوب مفهومی که هوش مصنوعی توضیح‌پذیر را با مبانی دینی-سیاسی جهاد تبیین پیوند می‌زند.

- ارزش افزوده اجتماعی و فرهنگی: کمک به تقویت تاب‌آوری شناختی جامعه در برابر جنگ نرم و عملیات روانی از طریق طراحی سامانه‌های فناورانه با قابلیت تبیین به‌موقع، مستدل و شفاف.

- ارزش افزوده راهبردی و بومی: حرکت به‌سوی توسعه هوش مصنوعی مسئولیت‌پذیر، اخلاق‌مدار و مبتنی بر ارزش‌های اسلامی-ایرانی. این پژوهش، به‌جای الگوبرداری صرف از چارچوب‌های غربی، به طراحی و توسعه هوش مصنوعی پاسخ‌گو به نیازها و اولویت‌های راهبردی ملی می‌پردازد و سهمی در بومی‌سازی فلسفه و کاربردهای هوش مصنوعی ایفا می‌کند.

## مبانی نظری پژوهش

### هوش مصنوعی توضیح‌پذیر (XAI)؛ مفهوم، تاریخچه و ضرورت

هوش مصنوعی توضیح‌پذیر یا تبیینی، به مجموعه‌ای از روش‌ها، تکنیک‌ها و ابزارهایی اطلاق می‌شود که به سامانه‌های هوشمند امکان ارائه توضیحاتی قابل فهم برای انسان درباره فرآیندهای تصمیم‌گیری، عملکرد و خروجی‌های خود را می‌دهد. ریشه‌های اولیه این مفهوم را می‌توان در سیستم‌های خبره دهه ۱۹۷۰ میلادی، همچون مدل MYCIN، جست‌وجو کرد؛ سامانه‌هایی که تلاش می‌کردند منطق استدلال خود را برای کاربران شفاف سازند (Das & Rad, 2020).

با این حال، نیاز به XAI در دهه‌های اخیر و هم‌زمان با رشد تصاعدی مدل‌های یادگیری ماشین -به‌ویژه شبکه‌های عصبی عمیق- شدت یافته است. این مدل‌ها، با وجود دستیابی به عملکردی بی‌سابقه در بسیاری از وظایف، به دلیل پیچیدگی ذاتی و تعداد بالای پارامترها، به پدیده‌ای موسوم به «جعبه سیاه» تبدیل شده‌اند؛ به این معنا که فرآیند تصمیم‌گیری آن‌ها برای انسان مبهم و غیرقابل درک باقی مانده است (Arrieta et al, 2020).

این عدم شفافیت در حوزه‌های حساسی مانند پزشکی، قضاوت قضایی، مالی و نظامی، می‌تواند به کاهش اعتماد، عدم پذیرش عمومی و بروز چالش‌های جدی قانونی و اخلاقی منجر شود. از این‌رو، ضرورت توسعه XAI برای حفظ اعتماد، افزایش پذیرش و امکان نظارت مؤثر انسانی بر سامانه‌های هوشمند بیش از پیش احساس می‌شود. XAI با تقویت شفافیت، قابلیت تفسیر و پاسخ‌گویی، در پی پر کردن شکاف میان عملکرد بالای هوش مصنوعی و درک انسانی از آن است (Gunning & Aha, 2019).

### ابعاد و اصول بنیادین XAI

برای طراحی و ارزیابی سامانه‌های XAI، مجموعه‌ای از ابعاد و اصول کلیدی مطرح شده‌اند که عبارت‌اند از:

- شفافیت: قابلیت مشاهده و درک فرآیندهای داخلی سیستم هوشمند توسط کاربران، از جمله ساختار مدل، الگوریتم‌های به‌کاررفته و نحوه پردازش داده‌ها.
- قابلیت تفسیر: توانایی سیستم در ارائه معانی قابل فهم از تصمیمات خود؛ این بُعد به درک «چگونگی» اتخاذ یک تصمیم یا رسیدن به یک خروجی خاص اشاره دارد.
- قابلیت توضیح: توانایی سیستم در ارائه «چرایی» یک تصمیم؛ یعنی ارائه دلایل و شواهد مستند و معقول که منجر به خروجی خاصی شده‌اند (Mustafa et al., 2025).
- اعتماد: ایجاد اطمینان در کاربران نسبت به صحت، دقت و قابلیت اتکای تصمیمات سیستم؛ که از طریق شفافیت و قابلیت توضیح حاصل می‌شود.
- انصاف: تضمین عدم تبعیض در تصمیم‌گیری‌های سیستم بر اساس ویژگی‌های حساس (مانند نژاد، جنسیت) و جلوگیری از سوگیری‌های ناخواسته موجود در داده‌ها یا الگوریتم‌ها.
- پاسخ‌گویی: قابلیت انتساب مسئولیت در قبال تصمیمات اتخاذشده توسط سیستم هوشمند؛ که مستلزم درک کامل از منطق عملکرد آن است (Al-Shammari, 2025).
- عقلانیت: پیروی از اصول عقلایی و منطقی در فرآیند تصمیم‌گیری، که به درک‌پذیری و پذیرش انسانی کمک می‌کند.

– محدوده اطمینان: ارائه میزان اطمینان یا عدم قطعیت سیستم نسبت به تصمیمات اتخاذشده، که به کاربران در ارزیابی اعتبار خروجی‌ها یاری می‌رساند (Gunning & Aha, 2019).

جدول ۱: ابعاد و اصول بنیادین XAI				
منبع: یافته‌های محقق				
منبع + بیانات رهبر انقلاب	ارتباط مفهومی با جهاد تبیین	شاخص کلیدی	تعریف	بعد
(Mustafa et al, 2025)	ایجاد پایه برای صراحت و صداقت	ساختار مدل، پردازش داده	مشاهده فرآیندهای داخلی	شفافیت
(Dwivedi et al, 2023)	امکان روشن‌نگری و تشریح فرآیندها	منطق فرآیند	درک "چگونگی" تصمیم	قابلیت تفسیر
(Das & Rad, 2020)	توانمندی برای پاسخگویی به شبهات و استدلال	دلایل مستند	ارائه "چرا" تصمیم	قابلیت توضیح
(Islam et al, 2021)	پیش‌نیاز اقناع عمومی و پذیرش تبیین	صحت خروجی	اطمینان کاربر	اعتماد
(AI Momani, 2025)	تضمین عدالت در تبیین و پرهیز از جانبداری	کنترل سوگیری	عدم تبعیض	انصاف
(AlShammari, 2025)	امکان مسئولیت‌پذیری در قبال محتوای تولیدشده	درک کامل منطق	انتساب مسئولیت	پاسخگویی
(Gunning & aha, 2019)	همسویی با استدلال و منطق به‌عنوان رکن تبیین	پذیرش انسانی	پیروی از اصول منطقی	عقلانیت
(AI-Shammari, 2025)	نشان‌دهنده صداقت در بیان میزان قطعیت اطلاعات	اعتبار خروجی	بیان عدم قطعیت	محدوده اطمینان

### دسته‌بندی روش‌های هوش مصنوعی توضیح‌پذیر (XAI)

روش‌های XAI را می‌توان در سه دسته اصلی طبقه‌بندی کرد:

۱. ذاتی (Intrinsic) در مقابل پس‌پردازشی (Post-hoc) در این دسته، مدل‌های ذاتاً تفسیرپذیر مانند درخت‌های تصمیم، به‌گونه‌ای طراحی شده‌اند که منطق تصمیم‌گیری آن‌ها به‌صورت شفاف قابل درک است. در مقابل، روش‌های پس‌پردازشی برای تفسیر مدل‌های «جعبه سیاه» به‌کار می‌روند. این روش‌ها-مانند تکنیک‌های مدل-آگنوستیک-با تقریب‌زدن رفتار مدل‌های پیچیده از طریق مدل‌های ساده و محلی، امکان درک تصمیم‌های الگوریتم را فراهم می‌سازند

(Islam et al, 2021). از جمله این روش‌ها می‌توان به SHAP و LIME اشاره کرد. SHAP، مبتنی بر نظریه بازی‌های تعاونی، با محاسبه مقادیر شپلی، سهم منصفانه و افزودنی هر ویژگی را در پیش‌بینی مدل محاسبه می‌کند و بدین‌وسیله تبیینی دقیق از نقش هر ویژگی ارائه می‌دهد (AI Momani, 2025).

۲. محلی (Local) در مقابل جهانی (Global) روش‌های تبیین‌پذیری می‌توانند محلی باشند؛ یعنی توضیحی برای یک تصمیم خاص یا یک داده ورودی مشخص ارائه دهند. در مقابل، روش‌های جهانی تصویری کلی از نحوه عملکرد کل مدل ارائه می‌کنند، مانند تحلیل اهمیت ویژگی‌ها یا درخت‌های تصمیم ساده‌شده که رفتار عمومی مدل را نمایش می‌دهند (Belghachi, 2023).

۳. مبتنی بر مدل (Model-specific) در مقابل مستقل از مدل (Model-agnostic) روش‌های مبتنی بر مدل، برای معماری یا الگوریتم خاصی طراحی شده‌اند و از ساختار درونی همان مدل بهره می‌برند؛ مانند Grad-CAM در شبکه‌های کانولوشنی یا روش‌های خاص برای درخت‌های تصمیم. در مقابل، روش‌های مستقل از مدل، قابل اعمال بر هر الگوریتمی هستند و بدون نیاز به دانستن جزئیات درونی مدل، صرفاً با استفاده از ورودی و خروجی، توضیحاتی ارائه می‌دهند (Zhang & Ogasawara, 2023).

### چالش‌ها و روندهای نوظهور در XAI

اصلی‌ترین چالش در حوزه XAI، تضاد ذاتی میان دقت و تفسیرپذیری است؛ به طوری که افزایش پیچیدگی مدل‌ها معمولاً با کاهش قابلیت تبیین همراه است. سایر چالش‌ها عبارت‌اند از:

- دشواری در ارزیابی کیفیت توضیحات؛

- تفاوت در درک‌پذیری توضیحات بین کاربران متخصص و عمومی؛

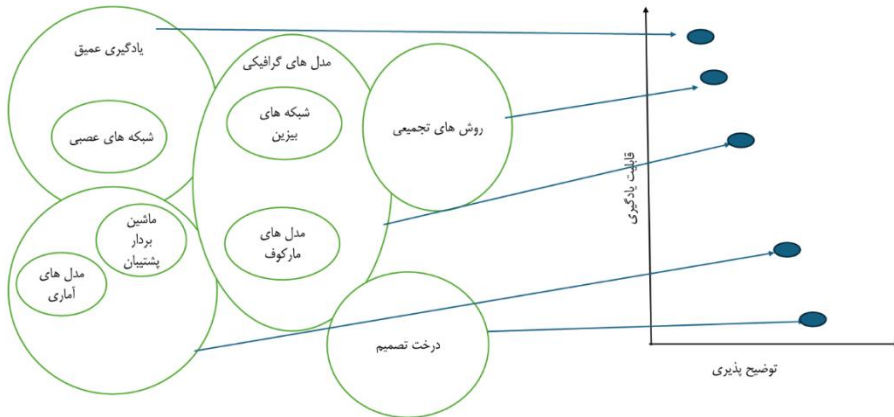
- امکان فریبندگی یا گمراه‌کنندگی توضیحات (Dwivedi et al, 2023).

یکی از روندهای نوظهور در این حوزه، تمرکز بر طراحی چارچوب‌های استاندارد برای سنجش کیفیت توضیحات است. پژوهش‌های اخیر نشان داده‌اند که صرف ارائه توضیح کافی نیست؛ بلکه این توضیح باید از منظر ویژگی‌هایی مانند دقت، پایداری، سازگاری و میزان کمک به تصمیم‌گیری

کاربر نیز ارزیابی شود (Machlev et al, 2022). همچنین، توجه به جنبه‌های تعاملی و بصری‌سازی توضیحات به‌عنوان عاملی مؤثر در درک‌پذیری و پذیرش نتایج مدل مطرح شده است. از این رو، توسعه رویکردهای ترکیبی که ضمن حفظ دقت مدل، شفافیت و تفسیرپذیری را نیز تضمین کنند، مسیر اصلی پیشرفت در حوزه XAI به‌شمار می‌رود. با این حال، یکی از چالش‌های مهم آن است که توضیحات تولیدشده توسط این روش‌ها برای کاربران غیرمتخصص به‌راحتی قابل درک نیستند (Babiker et al, 2024).

مدل‌های یادگیری عمیق، به‌دلیل ماهیت غیرشفاف خود، در حوزه‌های حساسی مانند پزشکی، دین، فرهنگ و قضاوت، با چالش‌های جدی مواجه‌اند؛ زیرا کاربران در این حوزه‌ها نیاز دارند بدانند چرا مدل به یک نتیجه خاص رسیده است. در این زمینه، XAI با ترکیب ماژول‌های چندگانه می‌تواند این ضعف را پوشش دهد و سامانه‌های هوشمند را به سطح بالاتری از شفافیت و تفسیرپذیری برساند (Ennab & Mcheick, 2025).

شکل (۱): عملکرد یادگیری در مقابل قابلیت توضیح برای چندین دسته از تکنیک‌های یادگیری (Gunning & Aha 2019)



## جهاد تبیین، چارچوبی راهبردی-فرهنگی

### مفهوم‌شناسی و مبانی اسلامی جهاد تبیین

جهاد تبیین به‌عنوان یک راهبرد کلان، توسط رهبر معظم انقلاب اسلامی طراحی و تبیین شده است. این رویکرد، برابرنهادی (آنتی‌تز) در مقابل جنگ‌های شناختی، روانی و فرهنگی دشمن محسوب می‌شود. مفهوم جهاد تبیین از ترکیب دو رکن اساسی شکل گرفته است:

- جهاد در لغت به معنای بذل و تلاش حداکثری است و در اصطلاح، «بذل الوسع فی طلب المطلوب الشرعی» تعریف می‌شود؛

- تبیین به معنای آشکارسازی، روشن کردن و زدودن ابهامات است. واژه «تبیان» در آیه شریفه ﴿وَنَزَّلْنَا عَلَيْكَ الْكِتَابَ تِبْيَانًا لِّكُلِّ شَيْءٍ﴾ (نحل: ۸۹) بر این معنا تأکید دارد (طباطبایی و همکاران، ۱۴۰۲).

بر این اساس، جهاد تبیین به‌عنوان تلاشی حداکثری، سازمان‌یافته و هوشمندانه برای روشنگری، کشف حقایق و مقابله فعالانه با تحریف و ابهام‌افکنی تعریف می‌شود. این مفهوم، ریشه در آموزه‌های امر به معروف و نهی از منکر و مسئولیت همگانی در برابر حقیقت دارد.

اهداف راهبردی جهاد تبیین عبارت‌اند از:

- روشنگری و افشای حقیقت؛
- مقابله با تحریف و خنثی‌سازی شبهات؛
- اقناع عمومی و تقویت انسجام اجتماعی.
- برای تحقق این اهداف، ارکان زیر به‌عنوان اصول عملیاتی استخراج شده‌اند:
- صراحت و صداقت: بیان بی‌پرده و راستین حقایق؛
- استدلال و منطق: استفاده از برهان‌های متقن برای اقناع مخاطب؛
- هوشمندی و زمان‌شناسی: اقدام به‌موقع و متناسب با بافتار اجتماعی؛
- پاسخ‌گویی فعال: پیش‌دستی در رفع شبهات، نه واکنش‌پذیری منفعل؛
- ساده‌سازی بدون تحریف: بیان مفاهیم پیچیده به زبان قابل فهم برای عموم؛
- مخاطب‌شناسی: شناخت دقیق نیازها و سطح فهم مخاطبان؛

- روایت‌سازی صحیح: تولید روایت‌های جذاب و معتبر در برابر روایت‌سازی دشمن؛  
- مسئولیت‌پذیری نخبگان: ایفای نقش پیشتاازانه توسط فرهیختگان و متخصصان (کاوند و محمدی مقدم، ۱۴۰۱؛ ملک‌زاده، ۱۴۰۳).

در عصر حاضر، فضای سایبری به عرصه اصلی جهاد تبیین تبدیل شده است. این فضا، به دلیل ویژگی‌هایی چون سرعت، گستردگی و ناشناختگی، هم تهدیدی جدی در قالب جنگ شناختی و هم فرصتی بزرگ برای روشنگری محسوب می‌شود. تحقق جهاد تبیین در این بستر، مستلزم بهره‌گیری هوشمندانه از ابزارهای فناورانه برای تولید و انتشار محتوا، رصد و شناسایی تهدیدات، و مقابله با عملیات روانی است (عراقی و همکاران، ۱۴۰۱).

### جنگ شناختی؛ زمینه و ضرورت پیوند

جنگ شناختی، نسل جدیدی از تهدیدات راهبردی است که هدف آن نه زیرساخت‌های فیزیکی، بلکه ذهن، ادراک و شناخت انسان‌هاست. در این نوع جنگ، از طریق دست‌کاری اطلاعات، تلاش می‌شود تا ادراکات، باورها و در نهایت رفتار افراد و جوامع تغییر یابد. ابزارهای این جنگ شامل اخبار جعلی، شبکه‌های اجتماعی و ربات‌ها، پروپاگاندا، عملیات روانی، مهندسی اجتماعی و بحران هویت است (Fenstermacher et al, 2023).

پاسخ مؤثر به این تهدیدات، «پدافند شناختی» است؛ مجموعه‌ای از اقدامات هدف‌مند برای تقویت تاب‌آوری فکری و شناختی جامعه. این اقدامات شامل افزایش سواد رسانه‌ای، تولید محتوای روشنگرانه، و توسعه فناوری‌های دفاعی مانند سامانه‌های شناسایی اخبار جعلی و تحلیل‌گرهای جریان‌های تحریفی است (محبوب و شکوری، ۱۴۰۱).

در این چارچوب، پیوند میان هوش مصنوعی توضیح‌پذیر (XAI) و راهبرد جهاد تبیین، ضرورتی دوگانه و مکمل می‌یابد؛ چرا که XAI با ایجاد شفافیت در تصمیم‌گیری‌های الگوریتمی، می‌تواند به ابزار مؤثری در مقابله با جنگ شناختی و تقویت پدافند غیرعامل تبدیل شود.

## پیشینه پژوهش

## پیشینه تحقیقات در حوزه XAI

مطالعات بین‌المللی گسترده‌ای به جنبه‌های فنی هوش مصنوعی توضیح‌پذیر، از جمله معیارهای ارزیابی، جلوگیری از سوگیری و کاربردهای آن در حوزه‌های حساس مانند پزشکی، امنیت، قضاوت و مالی پرداخته‌اند (Bhavaneetharan, 2025).

در سال‌های اخیر، پژوهش‌هایی با رویکردهای فرهنگی و دینی نیز ظهور کرده‌اند که به بررسی ظرفیت‌های XAI در زمینه‌های اسلامی پرداخته‌اند. از جمله مهم‌ترین این مطالعات می‌توان به موارد زیر اشاره کرد:

- بررسی پتانسیل XAI در شفاف‌سازی مدل‌های پردازش متون اسلامی و قرآنی (Mustafa, 2024)؛
- ارائه چارچوب‌هایی برای توسعه ربات‌های اجتماعی سازگار با ارزش‌های اسلامی (AlShammari, 2024).
- تحلیل چالش‌های اخلاقی هوش مصنوعی در فقه اسلامی و بررسی الزامات شرعی در طراحی سامانه‌های هوشمند (Al Momani, 2024).

جدول (۲): مطالعات جهانی مرتبط با ادبیات تحقیق. منبع: یافته‌های محقق

عنوان	منبع	جغرافیای تحقیق	یافته‌های کلیدی	ارتباط با پژوهش حاضر
Interpreting Arabic transformer models: A study on XAI interpretability for Quranic semantic-search	(Mustafa et al, 2024)	بین‌المللی (انگلستان)	استفاده از تکنیک‌های XAI (SHAP و LIME) برای تفسیر عملکرد مدل‌های زبانی در جستجوی معنایی آیات قرآن.	تبیین فناوری‌های معارف دینی با ارائه نمونه عینی برای تبیین معانی قرآن و پاسخ به شبهات از طریق شفافیت الگوریتم‌ها.
Culturally and Computationally Adaptive Social Robots for Arabophobes Islamic Societies	AlShammari (2025)	بحرین	ارائه چارچوب مفهومی با پنج ماژول برای طراحی ربات‌های اجتماعی سازگار با ارزش‌های اسلامی.	تبیین الگوی بومی فناوری مبتنی بر فرهنگ و ارزش‌های اسلامی به جای بپذیرش منفعلانه فناوری‌های وارداتی.
Explainable Artificial Intelligence Approaches: A Survey	(Islam et al, 2021)	قطر	مروری جامع بر روش‌های XAI، تحلیل مزایا و معایب، و تأکید بر لزوم حرکت به سمت هوش مصنوعی مسئولیت‌پذیر.	تبیین ضرورت اخلاقی و فلسفی توضیح‌پذیری به‌عنوان شرط لازم برای تحقق هوش مصنوعی مسئولیت‌پذیر.
Ethical challenges for using artificial intelligence in understanding Islamic jurisprudence	Al Momani (2025)	اردن	شناسایی چالش‌های کلیدی مانند سوگیری داده و فقدان زمینه فرهنگی؛ نتیجه‌گیری که هوش مصنوعی تنها ابزار است و جایگزین اجتهاد نمی‌شود.	تبیین محدود و نفوذ فناوری از منظر الهیاتی و فقهی با تعریف مرزین ابزار و مفتی.
A Review on Explainable Artificial Intelligence Methods, Applications, and Challenges	Belghachi (2023)	الجزایر	بررسی روش‌ها، کاربردها و چالش‌های اساسی XAI مانند تعادل دقت، تفسیرپذیری، سوگیری و ملاحظات فرهنگی.	تبیین چارچوب جهانی XAI و نشان دادن اینکه تلاش برای شفافیت یک هدف جهانی است.

## پیشینه تحقیقات در حوزه جهاد تبیین

در حوزه داخلی، پژوهش‌های متعددی به بررسی مفهوم جهاد تبیین و ابعاد مختلف آن پرداخته‌اند. این مطالعات، با رویکردهای نظری، راهبردی و کاربردی، به تبیین جایگاه جهاد تبیین در مقابله با جنگ شناختی و تحریف‌های رسانه‌ای پرداخته‌اند. مهم‌ترین پژوهش‌ها در این زمینه عبارت‌اند از:

- محمودی و صادقی (۱۴۰۱): این پژوهش به بررسی جایگاه جهاد تبیین در فضای مجازی پرداخته و بر ضرورت بهره‌برداری از ظرفیت‌های این فضا برای مقابله با توطئه‌های دشمن تأکید کرده است. تمرکز اصلی آن بر ابعاد عملیاتی و لزوم آمادگی در عرصه سایبری است.
- امینی و همکاران (۱۴۰۴): با رویکرد کیفی، الگویی راهبردی برای مدیریت رسانه‌های اجتماعی در چارچوب جهاد تبیین ارائه داده‌اند. این پژوهش مؤلفه‌هایی چون ماهیت رسانه‌ها، کنشگران، راهبردها و تنظیم‌گری را شناسایی کرده و بر کنش فعال در فضای مجازی تأکید دارد.
- ملک‌زاده (۱۴۰۳): جهاد تبیین را به‌عنوان الگویی برای گسترش مقاومت اسلامی در سطح جهانی تحلیل کرده و بر نقش آن در تقابل با روایت‌های سلطه‌گرایانه تأکید نموده است.
- طباطبایی و همکاران (۱۴۰۳): این پژوهش، جهاد تبیین را تلاشی روشنگرانه برای مقابله با تحریف دشمنان و اعتلای دین اسلام دانسته و به مبانی فقهی و الزامات شرعی آن پرداخته است.
- غفاری و همکاران (۱۴۰۳): با بهره‌گیری از روش نظریه داده‌بنیاد، مدلی شش مؤلفه‌ای از علل، عوامل پشتیبان، چالش‌ها و نتایج جهاد تبیین را بر اساس بیانات مقام معظم رهبری استخراج کرده‌اند.
- قربانی (۱۴۰۳): با روش تحلیل محتوای کیفی، به بررسی نقش همکاری نخبگان، رسانه‌ها و مردم در تحقق جهاد تبیین پرداخته و بر اهمیت روایت‌سازی صحیح و اقدام مؤثر در شکست جنگ شناختی تأکید کرده است.
- نیکونهاد و بیوک (۱۴۰۳): در پژوهش خود، جهاد تبیین را به‌عنوان «پادزهر جنگ شناختی» معرفی کرده‌اند و بر نقش آن در ارتقای تاب‌آوری شناختی جامعه تأکید داشته‌اند.

شکل (۲): منظومه فکری مقام معظم رهبری در خصوص جهاد تبیین (اقتباس از غفاری و همکاران، ۱۴۰۳)



### پژوهش‌های مرتبط با جنگ شناختی و پدافند شناختی

پژوهش‌های متعددی به بررسی ماهیت جنگ شناختی و راهکارهای مقابله با آن پرداخته‌اند. در سطح بین‌المللی، مطالعاتی چون:

– (Deppe & Schaal, 2022): جنگ شناختی را به‌عنوان تلاش هدف‌مند برای دست‌کاری ادراک و رفتار انسان‌ها مفهوم‌سازی کرده‌اند و چارچوبی برای دفاع و حمله در این نوع جنگ‌ها در حوزه ناتو ارائه داده‌اند.

– (Shafie et al. 2024): ابزارهای جنگ شناختی مانند اخبار جعلی و شبکه‌های اجتماعی را تحلیل کرده‌اند و بر ضرورت پدافند شناختی و جلوگیری از انتشار اطلاعات کذب برای ارتقای تاب‌آوری روانی جوانان ایالات متحده تأکید داشته‌اند.

در حوزه داخلی نیز پژوهش‌های قابل توجهی انجام شده است:

- نیکونهاد و بیوک (۱۴۰۳): با رویکرد تحلیل محتوا، جنگ شناختی را نسل پنجم جنگ‌ها معرفی کرده و جهاد تبیین را در اندیشه مقام معظم رهبری به‌عنوان پادزهر این تهدیدات تبیین نموده‌اند. این پژوهش به قواعد، الزامات و روش‌های جهاد تبیین در مقابله با جنگ شناختی پرداخته و الگویی مفهومی ارائه داده است.

- بهشتی نژاد و رضازاده (۱۴۰۴): همبستگی میان جهاد تبیین در فضای مجازی و مؤلفه‌های پیشگیری از بحران -از جمله آگاه‌سازی، مصون‌بخشی و سواد رسانه‌ای- را بررسی کرده‌اند و بر ضرورت آگاه‌سازی عمومی برای ایجاد مصونیت شناختی تأکید داشته‌اند.

- عزیزی و پیرانی (۱۴۰۱): به ضرورت بهره‌برداری از ظرفیت‌های جهاد تبیین برای پیشگیری از آسیب‌های فضای مجازی در راستای تحقق حکمرانی مجازی پرداخته‌اند.

- رضایی و ابراهیمی (۱۴۰۳): رفتارشناسی نسل Z و نحوه تعامل مدیران با این نسل را یکی از مهارت‌های مدیریتی کلیدی دانسته‌اند و راهکارهای اجرایی برای کاهش تأثیرات منفی جنگ شناختی بر این گروه را احصاء کرده‌اند.

این مجموعه مطالعات، ضمن تبیین ماهیت جنگ شناختی و راهکارهای دفاعی کلی، کمتر به ابعاد فنی و نحوه به‌کارگیری فناوری‌هایی چون هوش مصنوعی توضیح‌پذیر (XAI) در مقابله با این تهدیدات پرداخته‌اند. از این رو، پژوهش حاضر با تمرکز بر تلفیق XAI و جهاد تبیین، می‌کوشد خلأ موجود در پیوند میان فناوری و راهبردهای شناختی را پوشش دهد.

## شکاف‌های پژوهشی

بر اساس مرور پیشینه تحقیق، سه شکاف اصلی در مطالعات پیشین شناسایی شده‌اند:

- شکاف تلفیقی: فقدان پیوند مؤثر میان حوزه فناورانه) هوش مصنوعی توضیح‌پذیر (XAI) - و حوزه فرهنگی-راهبردی (جهاد تبیین). اغلب پژوهش‌ها به‌صورت مجزا به هر یک از این حوزه‌ها پرداخته‌اند و تعامل میان آن‌ها مغفول مانده است.

- شکاف ارزشی: نبود چارچوب هنجاری و بومی برای توسعه XAI در بستر فرهنگی-دینی جمهوری اسلامی ایران؛ به‌گونه‌ای که ارزش‌ها، اصول اخلاقی و الزامات اسلامی در طراحی سامانه‌های هوشمند لحاظ نشده‌اند.

- شکاف مدل‌سازی: فقدان مدل عملیاتی و کاربردی برای طراحی و پیاده‌سازی سامانه‌های هوشمند تبیین‌گر که بتوانند در راستای اهداف جهاد تبیین، عملکردی شفاف، قابل توضیح و اقناع‌گرایانه داشته باشند.

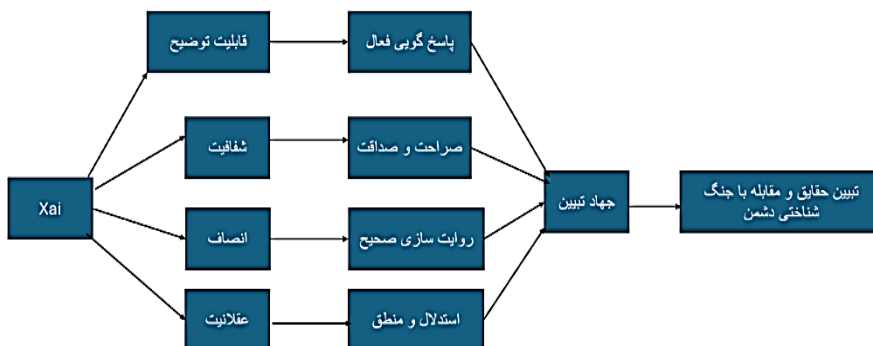
### مدل مفهومی مبتنی بر پیشنهاد تحقیق

این بخش، بر اساس مبانی نظری و مرور پیشینه پژوهش، چارچوبی مفهومی را معرفی می‌کند که روابط هم‌افزایانه میان مؤلفه‌های هوش مصنوعی توضیح‌پذیر (XAI) و راهبرد جهاد تبیین را تبیین می‌نماید. این چارچوب، امکان‌سنجی همگرایی میان فناوری‌های پیشرفته و راهبردهای فرهنگی را فراهم کرده و زمینه‌ساز طراحی سامانه‌های هوشمند بومی با کارکرد تبیینی است.

پژوهش‌های پیشین، در تبیین ابعاد، ضرورت و راهکارهای کلی جهاد تبیین در فضای مجازی موفق بوده‌اند؛ با این حال، تمرکز آن‌ها عمدتاً بر جنبه‌های مفهومی، راهبردی و اجتماعی-فرهنگی بوده و به‌طور خاص به چگونگی بهره‌برداری از ابزارهای فناورانه نوین، به‌ویژه هوش مصنوعی توضیح‌پذیر (XAI) برای تحقق اهداف جهاد تبیین پرداخته نشده است.

از این رو، مدل مفهومی پیشنهادی در این پژوهش، با هدف پر کردن شکاف‌های تلفیقی، ارزشی و مدل‌سازی، به‌دنبال ارائه چارچوبی نوآورانه برای توسعه سامانه‌های هوشمند تبیین‌گر در راستای پدافند شناختی و حکمرانی فرهنگی- رسانه‌ای است.

شکل (۳): مدل مفهومی ادبیات تحقیق



## روش‌شناسی تحقیق

این پژوهش از نظر هدف، در زمره تحقیقات کاربردی-توسعه‌ای قرار می‌گیرد و با هدف طراحی مدل مفهومی برای پدافند شناختی انجام شده است. از نظر ماهیت، پژوهش کیفی و با رویکرد تلفیقی در گردآوری و تحلیل داده‌ها صورت گرفته است. این رویکرد، با ترکیب تحلیل اسنادی-متنی (شامل بیانات مقام معظم رهبری و ادبیات نظری مرتبط) و داده‌های میدانی حاصل از مصاحبه‌های نیمه‌ساختاریافته با خبرگان، به افزایش اعتبار و عمق یافته‌ها کمک می‌کند.

استفاده از روش تأیید متقابل داده‌ها (Data Triangulation) در این پژوهش، موجب تقویت انسجام مفهومی و اعتبار نتایج شده و امکان بررسی هم‌زمان ابعاد نظری و تجربی موضوع را فراهم ساخته است (Creswell & Inoue, 2025).

جمع‌آوری داده‌ها در سه مرحله مجزا اما هم‌زمان انجام شده است:

### مرحله اول: مرور و تحلیل اسنادی

منبع داده‌ها	روش تحلیل	خروجی مورد انتظار
ادبیات Xai	مرور نظام‌مند (Systematic Review)	استخراج مؤلفه‌ها، اصول و روش‌های کلیدی XAI با تمرکز بر قابلیت پیوند فرهنگی-اجتماعی
بیانات مقام معظم رهبری	تحلیل تماتیک (Thematic Analysis)	شناسایی مفاهیم، مؤلفه‌ها و شاخص‌های جهاد تبیین از طریق تحلیل محتوای کیفی متون (khamenei.ir)

### مرحله دوم: مصاحبه با خبرگان

عنصر	توضیحات
جامعه آماری	متخصصان دو حوزه: هوش مصنوعی و داده‌کاوی (۵۰٪) - مطالعات فرهنگی و رسانه، جهاد تبیین و پدافند شناختی (۵۰٪)، تعداد ۱۰ نفر تا رسیدن به مرحله اشباع نظری
روش نمونه‌گیری	هدفمند و گلوله‌برفی، برای دسترسی به خبرگان با دانش عمیق
ابزار جمع‌آوری	مصاحبه‌های نیمه‌ساختاریافته مبتنی بر پرسشنامه اولیه حاصل از تحلیل اسنادی
ارزیابی	به دو صورت مجازی و رو در رو، ضبط مصاحبه‌ها با رضایت آگاهانه خبرگان، ترانسکرپت کامل مصاحبه‌ها (تبدیل صوت به متن) - کدگذاری محرمانه داده‌ها

## روش تحلیل داده‌ها: تحلیل تماتیک بازتابی

داده‌های سه‌گانه (ادبیات، بیانات رهبری، مصاحبه‌ها) با استفاده از رویکرد جدید تحلیل تماتیک بازتابی ۱ تحلیل شدند. این روش امکان توجه به معانی ضمنی داده‌ها و بررسی تأثیر دیدگاه‌های پژوهشگر ۲ بر تحلیل را فراهم می‌کند (Braun & Clarke, 2024). فرآیند شش مرحله‌ای به صورت زیر اجرا می‌شود:

مرحله	اقدامات اجرایی
۱. آشنایی با داده‌ها	خواندن مکرر ترانسکرپت‌ها، بیانات رهبری و یادداشت‌برداری از ادبیات برای درک عمیق محتوا، استفاده از نرم افزار Maxqda 2024
۲. تولید کدهای اولیه	اختصاص کدهای اولیه به بخش‌های کلیدی متن مرتبط با XAI و جهاد تبیین
۳. جستجوی تم‌ها	دسته‌بندی کدها در قالب تم‌های اصلی و فرعی برای کشف الگوهای معنایی
۴. بازبینی تم‌ها	بررسی اصلاحی تم‌ها برای اطمینان از همسویی با کل داده‌ها و ارتباط منطقی بین کدها
۵. تعریف و نام گذاری تم‌ها	ارائه تعریف دقیق برای هر تم و نام‌گذاری شفاف بر اساس ماهیت مفهومی
۶. نگارش تحلیل	تدوین یافته‌ها به صورت منسجم برای طراحی مدل مفهومی

## فرآیند طراحی مدل مفهومی

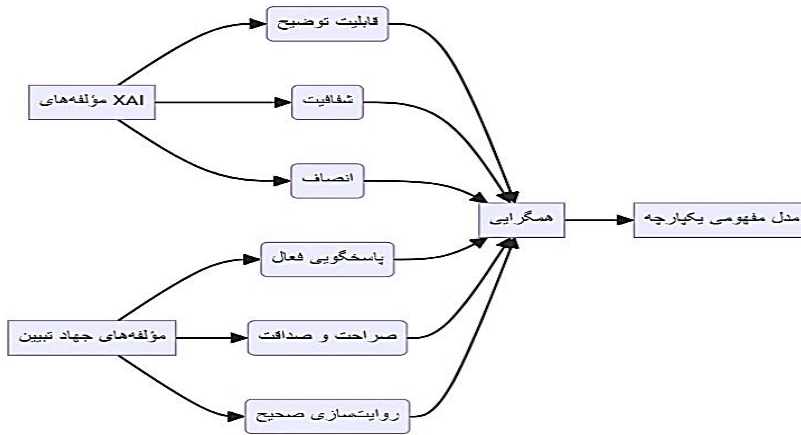
مدل مفهومی نهایی بر اساس یافته‌های تحلیل تماتیک در سه مرحله طراحی می‌شود:

مرحله ۱: استخراج مؤلفه‌ها که شامل تدوین فهرست نهایی مؤلفه‌های کلیدی XAI و جهاد تبیین حاصل از تحلیل‌های قبلی و اولویت‌بندی مؤلفه‌ها بر اساس تکرار در داده‌ها و اهمیت نظری است.

مرحله ۲: ایجاد ماتریس همگرایی و طراحی گرافیکی مدل آن برای نمایش روابط متقابل و هم‌پوشانی میان مؤلفه‌های دو حوزه:

1. Reflexive Thematic Analysis
2. Reflexivity

شکل (۴): روابط و هم پوشانی مولفه های حوزه جهاد تبیین و Xai



مرحله ۳: ترسیم مدل نهایی که شامل طراحی نمودار گرافیکی روابط بین مؤلفه‌ها و ارائه توضیحات متنی برای هر بخش از مدل و اعتبارسنجی اولیه مدل توسط خبرگان است.

### معیارهای اعتبار و پایایی پژوهش

محقق برای تضمین کیفیت یافته‌های کیفی، از معیارهای لینکلن و گوبا با رویکرد به روز شده بازتابی استفاده کرده است که شامل موارد ذیل است (Bang, 2024):

معیار	راهکار اجرایی
اعتبار	بازبینی مدل توسط خبرگان، ارائه نقل قول‌های مستقیم از مصاحبه‌ها در متن
قابلیت انتقال	توصیف دقیق بافتار پژوهش (نمونه‌گیری، روش مصاحبه، تحلیل)، ارائه جزئیات برای امکان تکرار پژوهش
قابلیت اتکا	مستندسازی دقیق تمام مراحل (کدگذاری، تحلیل)، امکان بازبینی فرآیند توسط پژوهشگران دیگر، محاسبه ضریب هولستی <sup>۱</sup> برای پایایی کدگذاری بین کدگذارها (حداقل ۰.۸).
تأییدپذیری	تفکیک داده‌ها از تفسیرهای شخصی، ارائه شواهد کافی برای هر نتیجه‌گیری، ذکر محدودیت‌های پژوهش

### 1. Holsti Index

## تحلیل یافته‌ها و طراحی مدل مفهومی

### ۱. تجزیه و تحلیل تم‌ها

تحلیل تماتیک ادبیات تحقیق: در این بخش، نتایج حاصل از مرور نظام‌مند ادبیات در حوزه‌های هوش مصنوعی توضیح‌پذیر (XAI) و مباحث مرتبط با آن ارائه شده است. هدف این تحلیل، استخراج مؤلفه‌ها، اصول و روش‌های کلیدی XAI است که به‌عنوان لایه فناورانه مدل مفهومی عمل می‌کنند. جدول (۳) خلاصه‌ای از این تحلیل را نشان می‌دهد.

جدول (۳): نمونه خلاصه تحلیل تماتیک ادبیات تحقیق

تم اصلی	تم‌های فرعی	مفاهیم	منابع/نمونه مراجع داخل متن
اصول و ابعاد XAI	شفافیت؛ قابلیت تفسیر؛ قابلیت توضیح؛ اعتماد؛ انصاف؛ پاسخگویی؛ عقلانیت؛ محدوده اطمینان	مجموعه قابل فهم، مبانی نظری، سامانه تبیینی، قابلیت شفافیت، قابل اعتماد و پاسخگو باشد؛ ستون‌های ارزیابی طراحی و اخلاق	Mustafa et al, (2024); Gunning & Aha (2019); Belghachi (2023); Ennab & Mcheick (2025); Machlev & Heistrene (2022)
روش‌ها و چالش‌ها	مدل‌های ذاتی، روش‌های پس از-آموزش، مدل محلی در مقابل جهانی؛ مدل-وابسته/مدل-آگنوستیک	دسته‌بندی فنی روش‌ها و فرآورده‌های عملی؛ چالش‌های اساسی (تعادل دقت و تبیین، ارزیابی کیفیت توضیحات، مخاطب‌محوری)	Das & Rad (2020); Arrieta et al, (2020); Islam et al. (2021)
کاربردها و پدافند	تشخیص اخبار جعلی؛ تبیین تصمیمات مدل‌های زبانی؛ ربات‌های اجتماعی؛ کاربرد در حوزه‌های حساس (پزشکی، مالی، نظامی، مذهبی)	نمونه‌های کاربردی، توانایی در شناسایی، توضیح علت‌رد/پذیرش محتوا، کاربردی، کاربرپسند، مصداق در متن: جستجوی معنای قرآنی، شفاف‌سازی مدل‌های (NLP)	Dwivedi et al, (2023); Zhang, & Ogasawara (2023); Machlev et al, (2022); Babiker et al (2024)

تم اصلی	تم‌های فرعی	مفاهیم	منابع / نمونه مراجع داخل متن
تطبیق فرهنگی و پایگاه دانش	پایگاه بومی سازی؛ دانش؛ نقش ارزش های دینی؛ سازگاری زبانی و فرهنگی	الزام به طراحی توضیحات، رابط‌های کاربری با زمینه، فرهنگی/دینی، تاکید بر الگوی بومی	AlShammari (2025); Al Momani (2025)
حاکمیت، نظارت و مسئولیت‌پذیری	پاسخگویی؛ انتساب مسئولیت؛ کمیته نظارتی انسانی	نیاز به چارچوب حکمرانی برای تضمین انصاف، مشروعیت و امکان حسابرسی خروجی‌های XAI، حاکمیت یک پارچه سایبری، قدرت نرم، فیدبک یکپارچه	Babiker et al, (2025) Bhavaneetharan (2025) Gunning & Aha (2019)

یافته‌های حاصل از ادبیات، مبانی نظری و فنی هوش مصنوعی تبیینی را برای مدل تحقیق فراهم می‌کنند. اصول و ابعاد XAI (مانند قابلیت توضیح، شفافیت و انصاف) به‌عنوان ستون‌های اصلی طراحی مدل عمل می‌کنند، در حالی که روش‌ها و چالش‌ها (مانند تعادل دقت-تفسیرپذیری) به محقق می‌کنند تا از نظر فنی، رویکردی واقع‌بینانه داشته باشد. مفاهیمی چون تطبیق فرهنگی و حاکمیت نیز، که در ادبیات ظهور کرده‌اند، زمینه را برای پیوند XAI با گفتمان جهاد تبیین فراهم می‌کنند.

تحلیل تماتیک بیانات رهبر معظم انقلاب: در این بخش، محقق کار خود را با تجزیه تماتیک بیانات رهبر معظم انقلاب در خصوص جهاد تبیین ادامه داد. این کار به پژوهش اجازه می‌دهد که مولفه‌های اصلی و مرتبط جهاد تبیین و هوش مصنوعی توضیح‌پذیر را استخراج کند. جدول (۴) نمونه‌ای از تحلیل تماتیک بیانات رهبر معظم انقلاب در بازه ۱۳۹۵ تا نیمه اول ۱۴۰۴ را نشان می‌دهد.

جدول (۴): نمونه تحلیل تماتیک بیانات رهبر معظم انقلاب در خصوص جهاد تبیین

منبع-تاریخ	ارتباط با Xai	مفاهیم	تم‌های فرعی	تم اصلی	ردیف
۱۹/۱۱/۱۴۰۰ (دیدار با نیروی هوایی)؛ ۱۹/۱۰/۱۴۰۱ (دیدار با مردم) ۰۳/۱۱/۱۴۰۰ (دیدار با مداحان)؛ ۱۹/۱۰/۱۴۰۱ (دیدار با مردم قم) ۰۳/۱۱/۱۴۰۰ (دیدار با مداحان)؛ ۱۲/۰۴/۱۳۹۵ (دیدار با دانشجویان)	۱- درک منطق مقابله با شبهات ۲- تمایز تلاش هدف‌دار از عمومی	تلاش هدف‌دار در مقابل دشمن؛ تبیین حقایق و رفع شبهات؛ جهاد با زبان (السنه) بر اساس وصیت امیرالمؤمنین، جهاد یعنی تلاش با هدف مقابله با دشمن؛ نه هر تلاشی جهاد است؛ مکمل جهاد نظامی فریضه قطعی و فوری؛ علاج پروپاگاندا دشمن؛ خنثی کردن اغوا و دروغ‌پردازی	ماهیت جهاد تبیین به‌عنوان مقابله با دشمن؛ تمایز جهاد تبیین از تلاش‌های عادی؛ فریضه بودن جهاد تبیین	تعریف و ماهیت	۱
۱۴/۰۳/۱۴۰۲ (سالگرد ارتحال امام)؛ ۱۹/۱۲/۱۴۰۰ (دیدار با خبرنگاران)	۱- پاسخگویی فوری به پروپاگاندا ۲- تقویت اعتماد و امید عمومی	تقویت امید و ایمان؛ مقابله با یأس آفرینی؛ حفظ مردم‌سالاری دینی	نقش در حفظ انقلاب و جامعه؛ آگاه‌سازی و تاب‌آوری روانی جامعه؛ لزوم توجه به نسل جوان	اهمیت و ضرورت	۲

منبع-تاریخ	ارتباط با Xai	مفاهیم	تم‌های فرعی	تم اصلی	ردیف
۱- ابزارهای پس از تحلیل برای تبیین ۲- شفافیت بدون سوگیری ۱۹/۱۱/۱۴۰۰ (دیدار با نیروی هوایی)؛ ۰۸/۰۸/۱۴۰۱ (دیدار با دست‌اندرکاران کنگره شهادای قم) ۰۵/۰۷/۱۴۰۰ (اربعین حسینی)؛ ۱۲/۰۴/۱۳۹۵ (دیدار با دانشجویان)	۱- ابزارهای پس از تحلیل برای تبیین ۲- شفافیت بدون سوگیری	استفاده از رسانه‌ها، فضای مجازی، هنر (فیلم، شعر، مستند)؛ هیئت‌ها و خطبه‌ها صراحت و شفافیت؛ اجتناب از دروغ و تهمت؛ منطق محکم و عاطفه انسانی	ابزارهای تبیین شیوه‌های اخلاقی و منطقی	روش‌ها و ابزارها	۴
۱۶/۱۱/۱۴۰۲ (دیدار با فرماندهان نیروی هوایی)؛ ۰۵/۰۵/۱۴۰۱ (دیدار با ائمه جمعه)؛ ۲۹/۰۶/۱۴۰۲ (دیدار با پیشکسوتان دفاع مقدس)؛ ۲۸/۱۱/۱۴۰۰ (ارتباط با مردم آذربایجان)	۱. مخاطب‌شناسی نخبگان ۲. همخوانی با مخاطبان عمومی	وظیفه خواص برای صریح بودن و شبهه‌زدایی؛ نقش مداحان، ائمه جمعه، دانشجویان وظیفه همه (بسیجیان، هسته‌های مقاومت، جوانان)؛ تبیین در مساجد و هیئت‌ها	۴- مسئولین، خواص و نخبگان ۴-۲ مسئولیت عمومی و جوانان	مخاطبان و مسئولیت‌ها	۴
۱۳/۱۰/۱۴۰۲ (دیدار با مداحان)؛ ۰۵/۰۹/۱۴۰۱ (دیدار با بسیجیان) ۲۹/۰۶/۱۴۰۲ (دیدار با پیشکسوتان دفاع مقدس)؛ ۱۲/۰۴/۱۳۹۵ (دیدار با دانشجویان)	۱. تبیین روایت‌های سیاسی ۲. تاب‌آوری روایی فرهنگی	تبیین غزه، انتخابات، دشمنی با نظام؛ روایت درست پیشرفت‌ها و حماسه‌ها تبیین دفاع مقدس، سبک زندگی اسلامی، عدالت و استقلال	۵-۱ مسائل روز و سیاسی ۵-۲ مسائل تاریخی و فرهنگی ۵-۳ شبهات دینی و مذهبی	موضوعات کلیدی تبیین	۵

منبع-تاریخ	ارتباط با Xai	مفاهیم	تم‌های فرعی	مفهوم اصلی	ردیف
۱. مستند به منبع دینی ۲. ردیابی تاریخی	۱. مستند به منبع دینی ۲. ردیابی تاریخی	جهاد تبیین فاطمه زهرا و زینب کبری؛ احیای امر اهل بیت، امام راحل با جهاد تبیین حکومت سلطنتی را ساقط کرد؛ ادامه نهضت اسلامی	۱-۶ ریشه در اهل بیت ۲-۶ ارتباط با امام خمینی و انقلاب	ارتباط تاریخی و دینی	۶
۱. شناسایی شبهات دشمن ۲. جلوگیری از سوء استفاده	۱. شناسایی شبهات دشمن ۲. جلوگیری از سوء استفاده	دروغ‌پردازی، شبهه مانند ویروس؛ اغوای خواص و توده‌ها، خطر دنیا طلبی؛ عدم تبیین منجر به سوء استفاده از دین	۱-۷ چالش‌های دشمن ۲-۷ هشدارهای داخلی	چالش‌ها و هشدارها	۷
۱. تفویت قدرت نرم ۲. حضور عمومی آحاد ملت	۱. تفویت قدرت نرم ۲. حضور عمومی آحاد ملت	ماندگار کردن انقلاب؛ ساقط کردن توطئه‌ها؛ تقویت قدرت نرم، حضور مردم در صحنه؛ مقابله با جنگ نرم؛ روایت درست تاریخ مانند عاشورا	۱-۸ اهداف نهایی ۲-۸ پیامدهای عملی	پیامدها و اهداف نهایی	۸

## ۲. تفسیر تم‌های احصاء شده و پیوند آن‌ها با مبانی نظری

تعریف جهاد تبیین: جهاد تبیین یک فریضه قطعی، هوشمندانه و هدفمند برای مقابله با پروپاگاندای دشمن است.

پیوند با Xai: همسویی با مفهوم قابلیت توضیح (Explainability)؛ هر دو بر تبیین مستدل حقایق و خنثی‌سازی شبهات با استدلال‌های مستند تأکید دارند.

اهمیت و ضرورت: ابزاری برای مقابله با یأس، تقویت امید عمومی، حفظ مردم‌سالاری دینی و افزایش تاب‌آوری روانی جامعه.

پیوند باXai: ارتباط مستقیم با اعتماد (Trust)؛ هر دو با شفاف‌سازی و روشن‌گری، اعتماد مخاطب (جامعه/کاربر) را افزایش می‌دهند.

روش‌ها و ابزارها: تأکید بر صراحت، شفافیت، منطق محکم و استفاده از رسانه‌ها و فضای مجازی.

پیوند باXi: تطابق با شفافیت (Transparency) و قابلیت تفسیر (Interpretability)؛ سیستم‌های XAI با نمایش فرآیند تحلیل و ارائه توضیحات قابل فهم، به تحقق این اهداف کمک می‌کنند.

مخاطبان و مسئولیت‌ها: مسئولیت دوگانه "خواص" (نخبگان) برای شبهه‌زدایی و "عموم مردم" برای مشارکت در تبیین.

پیوند باXai: تأکید بر مخاطب‌محوری؛ توضیحات XAI باید متناسب با سطح درک مخاطب (فنی برای متخصصان، ساده برای عموم) ارائه شود.

موضوعات کلیدی: پوشش گسترده مسائل سیاسی، تاریخی، فرهنگی و شبهات دینی.

پیوند باXi: نیاز به انعطاف‌پذیری کاربردی؛ مدل XAI باید در حوزه‌های مختلف (تحلیل سیاسی، فرهنگی و...) با الگوریتم‌های متنوع کاربرد داشته باشد.

ارتباط تاریخی و دینی: ریشه‌دار در نهضت اهل بیت و انقلاب اسلامی با تأکید بر اصالت دینی.

پیوند باXai: لزوم بومی‌سازی (Localization) و اخلاق‌مداری (Ethical AI)؛ سیستم‌ها باید با ارزش‌های دینی و انقلابی همخوانی داشته باشند.

چالش‌ها و هشدارها: مقابله با دروغ‌پردازی دشمن و نیاز به اقدام سریع برای پیشگیری از شبهات.

پیوند باXai: ارتباط با پاسخگویی فعال و کنترل سوگیری (Bias Control)؛ سیستم باید شبهات را سریع شناسایی کرده و از تحریف ناشی از سوگیری پرهیز کند.

پیامدها و اهداف نهایی: تقویت قدرت نرم و ماندگاری انقلاب به‌عنوان اهداف بلندمدت.

پیوند باXi: نقش راهبردی XAI به‌عنوان ابزاری برای تحقق اهداف ملی، فراتر از یک راهکار فنی صرف.

این تحلیل هشت‌گانه، جهاد تبیین را به‌عنوان یک فریضه هوشمندانه، چندبعدی و راهبردی معرفی می‌کند و نشان می‌دهد چگونه مفاهیم کلیدی XAI - نظیر قابلیت توضیح، اعتماد، شفافیت، مخاطب‌محوری، انعطاف‌پذیری، اخلاق‌مداری و کنترل سوگیری - می‌توانند به‌صورت سیستمی در تحقق اهداف جهاد تبیین نقش‌آفرینی کنند. این پیوند، زمینه‌ساز طراحی سامانه‌های هوشمندی است که نه تنها اطلاعات ارائه می‌دهند، بلکه با استدلال‌های مستند و منطقی، شبهات را خنثی کرده و اعتماد عمومی را تقویت می‌نمایند.

### ۳. تحلیل تماتیک مصاحبه با خبرگان

در این بخش، یافته‌های حاصل از تحلیل تماتیک مصاحبه‌های نیمه‌ساختاریافته با متخصصان دو حوزه هوش مصنوعی و مطالعات فرهنگی - رسانه‌ای ارائه شده است. تحلیل‌ها بر اساس ۱۲ تم اصلی دسته‌بندی شده‌اند که نمایانگر دیدگاه‌های عملی و راهبردی خبرگان در خصوص طراحی مدل مفهومی هوش مصنوعی تبیینی با رویکرد جهاد تبیین هستند.

جدول (۵): نمونه تحلیل تماتیک مصاحبه با خبرگان

تم اصلی	تم‌های فرعی	مفاهیم کلیدی	نقل قول
نقش و قابلیت‌های XAI	شفافیت؛ ارائه شواهد؛ محدوده اطمینان	نمایش چرایی/چگونگی تصمیم، مقدار عدم قطعیت، قابلیت پیگیری	XAI باید بتواند بگوید کدام دلایل منجر به رد یک ادعا شد صرف گفتن غلط است کافی نیست. (خبره اول)
سطح‌بندی توضیحات و مخاطب‌شناسی	لایه‌های توضیح؛ قالب‌های بصری؛ تطبیق زبان	توانایی تشخیص، تخصص محور، عام محور، ارجاع صحیح به متون دینی و قرآنی	برای مردم معمولی یک خلاصه ساده، برای روحانیون یک گزارش مستدل و منابع قرآنی لازم است. (خبره سوم)
تلفیق انسان در حلقه (HITL)	نقاط سرنوشت‌ساز انسان؛ اشتراک‌گذاری تصمیم	چه مواردی باید حتماً انسان تایید کند؛ مکانیزم ارجاع	هر تصمیمی که به احکام دینی نزدیک است باید توسط یک کمیته فقهی بازبینی شود؛ AI فقط پیش‌نویس ارائه دهد (نفر دوم)

تم اصلی	تم‌های فرعی	مفاهیم کلیدی	نقل قول
حاکمیت، مسئولیت و شفافیت نهادی	کمیته فقهی-فنی؛ مستندسازی؛ ممیزی	چارچوب‌های نظارتی، لاگینگ، نسخه‌نگاری، گزارش شفافیت	لاگ توضیحات باید ذخیره شود تا در صورت اختلاف قابل بررسی باشد؛ بدون این امکان، مسئولیت‌پذیری فراهم نمی‌شود. (خبیره پنجم)
پایگاه دانش و تطبیق فرهنگی	منابع دینی؛ واژگان بومی؛ ترجمه مفاهیم	ساخت بانک شبهات، منابع معتبر فقهی، چندزبانه بودن	پایگاه دانش باید شامل فتاوی معتبر، تفسیرهای مرجع و مثال‌های تاریخی باشد؛ غیر از این توضیح فاقد مشروعیت خواهد بود. (خبیره اول)
معیارهای ارزیابی و KPIها	دقت تشخیص؛ فهم‌پذیری؛ تأثیر عملیات	پوشش شایعه‌های واقعی در تشخیص، امتیاز درک کاربران، کاهش شیوع	ما باید نشان دهیم پس از مداخله، میزان اشتراک یک شایعه حداقل ۳۰٪ کاهش یافته است و باز خورد کاربر را دریافت کنیم؛ این KPI عملیاتی است (خبیره هفتم)

در نهایت، از تحلیل مصاحبه‌های نیمه‌ساختاریافته با خبرگان، ۱۲ تم اصلی استخراج شد که در ادامه به صورت فشرده تفسیر می‌شوند:

جهاد تبیین به عنوان یک فریضه هوشمندانه، هدف‌مند و فوری (تم ۱)، با تأکید بر شفافیت، صراحت و بهره‌گیری از ابزارهای نوین ارتباطی (تم ۳)، در راستای مقابله با یأس‌آفرینی دشمن، تقویت امید عمومی و افزایش تاب‌آوری روانی جامعه (تم ۲) طراحی شده است. این مفهوم، مسئولیتی دوگانه برای نخبگان (در زمینه شبهه‌زدایی) و عموم مردم (در زمینه مشارکت فعال در تبیین) تعریف می‌کند (تم ۴) و با پوشش موضوعات متنوع سیاسی، تاریخی، فرهنگی و دینی (تم ۵)، به دنبال مقابله با دروغ‌پردازی، اغواگری و شبهه‌افکنی دشمن از طریق اقدام سریع و پیشگیرانه است (تم ۷).

ریشه‌های دینی و انقلابی این راهبرد (تم ۶) و اهداف بلندمدت آن - از جمله تقویت قدرت نرم، ماندگاری انقلاب اسلامی و حفظ مردم‌سالاری دینی - بر اهمیت راهبردی آن تأکید دارند (تم ۸). در عین حال، چالش‌هایی نظیر سوگیری اطلاعاتی، پیچیدگی مخاطبان و ضرورت بومی‌سازی ابزارهای ارتباطی و فناورانه (تم‌های ۹ تا ۱۲)، بر نیاز به طراحی سامانه‌های هوشمند تبیین‌گر تأکید می‌ورزند.

پیوند این ابعاد با هوش مصنوعی توضیح‌پذیر (XAI) از طریق مفاهیم زیر تحقق می‌یابد:

- قابلیت توضیح برای تبیین مستدل حقایق؛
- اعتمادسازی برای تقویت باور عمومی؛
- شفافیت و قابلیت تفسیر برای ارائه منطق مستند؛
- مخاطب‌محوری در تطبیق سطح توضیحات با درک نخبگان و عموم؛
- انعطاف‌پذیری کاربردی در حوزه‌های مختلف؛
- بومی‌سازی و اخلاق‌مداری برای هم‌خوانی با ارزش‌های دینی و انقلابی؛
- پاسخ‌گویی فعال برای خنثی‌سازی سریع شبهات؛
- کنترل سوگیری برای جلوگیری از تحریف اطلاعات.

این چارچوب، هوش مصنوعی توضیح‌پذیر را از یک ابزار صرفاً فنی به ابزاری راهبردی برای تحقق اهداف غایی جهاد تبیین تبدیل می‌کند؛ سامانه‌ای که نه تنها اطلاعات ارائه می‌دهد، بلکه با استدلال‌های مستند و اقناع‌گرایانه، شبهات را خنثی کرده و اعتماد عمومی را تقویت می‌نماید.

#### ۴. تفسیر و پیوند یافته‌های مصاحبه‌ها با مبانی نظری و بیانات رهبری

۱. قابلیت‌های هوش مصنوعی قابل توضیح (XAI) همچون شفافیت، ارائه شواهد و بیان محدوده اطمینان، با ضرورت «استدلال و منطق» در جهاد تبیین هم‌راستا هستند؛ سامانه باید افزون بر ارائه پاسخ صحیح، دلایل و نحوه دستیابی به آن را نیز توضیح دهد تا بستری مؤثر برای پاسخ‌گویی به شبهات فراهم گردد.

۲. سطح‌بندی توضیحات با اصل «مخاطب‌شناسی» در بیانات رهبری پیوند خورده است؛ مدل باید توانایی ارائه محتوا از خلاصه‌های ساده برای عموم تا گزارش‌های فنی برای متخصصان را داشته باشد.

۳. مکانیسم «انسان در حلقه تصمیم‌گیری» (HITL) به معنای آن است که تصمیم‌گیری‌های نهایی یا حساس توسط انسان و نه صرفاً هوش مصنوعی انجام شود؛ این رویکرد بر نقش انسان در تصمیمات حساس، به‌ویژه فقهی، تأکید دارد و موجب تقویت مسئولیت‌پذیری نخبگان و حکمرانی هوش مصنوعی می‌شود.
۴. تحقق حاکمیت و شفافیت نهادی از طریق ایجاد چارچوب‌های نظارتی، ثبت لاگ‌ها و تشکیل کمیته‌های فقهی-فنی، اصل «پاسخ‌گویی» در XAI و «صراحت» در جهاد تبیین را عملیاتی می‌سازد.
۵. بهره‌گیری از پایگاه دانش بومی شامل فتاوی، منابع دینی و تاریخی برای بومی‌سازی مدل ضروری است؛ این امر از تحریف جلوگیری کرده و «روایت‌سازی صحیح» را تضمین می‌کند.
۶. معیارهای ارزیابی باید فراتر از شاخص‌های فنی، بر مؤلفه‌هایی چون «کاهش شایعات» و «افزایش آگاهی عمومی» تمرکز داشته باشند که با ماهیت «فریضه‌محور و هدف‌دار» جهاد تبیین هم‌خوانی دارند.
۷. مدیریت سوگیری از طریق تنوع‌بخشی به منابع داده، اصل «انصاف» در XAI و پرهیز از «جانبداری» در تبیین را محقق می‌سازد.
۸. زمان‌شناسی در پاسخ‌گویی با تولید خودکار پاسخ‌های اولیه در موج‌های خبری، اصل «هوشمندی» در جهاد تبیین و «پاسخ‌گویی فعال» در XAI را به‌ثمر می‌رساند.
۹. کاهش ریسک‌هایی چون ارائه توضیحات فریبنده، از طریق آزمون‌های پایداری و تعیین آستانه عدم قطعیت، موجب ارتقاء «اعتماد» و «صداقت» مدل می‌شود.
۱۰. رعایت اخلاق و حریم خصوصی با توجه به ملاحظات حقوقی در داده‌ها، به تقویت مسئولیت‌پذیری و حکمرانی هوش مصنوعی کمک می‌کند.
۱۱. آموزش سواد رسانه‌ای با تمرکز بر ارتقاء آگاهی عمومی، در راستای هدف «افزایش تاب‌آوری شناختی جامعه» در جهاد تبیین عمل می‌کند.
۱۲. پیاده‌سازی مرحله‌ای با آغاز از حداقل محصول قابل بهره‌برداری (MVP)، رویکردی عملیاتی برای تدوین نقشه راه پژوهش ارائه می‌دهد.

## ۵. یکپارچه‌سازی یافته‌ها: از تم‌های جزئی تا ابرتم‌های انتزاعی

در این پژوهش، یافته‌های کیفی حاصل از تحلیل تماتیک سه منبع (ادبیات، بیانات رهبری و مصاحبه‌های خبرگان) از طریق فرآیند انتزاع معنایی، به ده ابرتم کلیدی در سه لایه سلسله‌مراتبی تقلیل یافت تا از پیچیدگی غیرضروری کاسته و بر مفاهیم بنیادین تکرار شده در تمام منابع تمرکز شود.

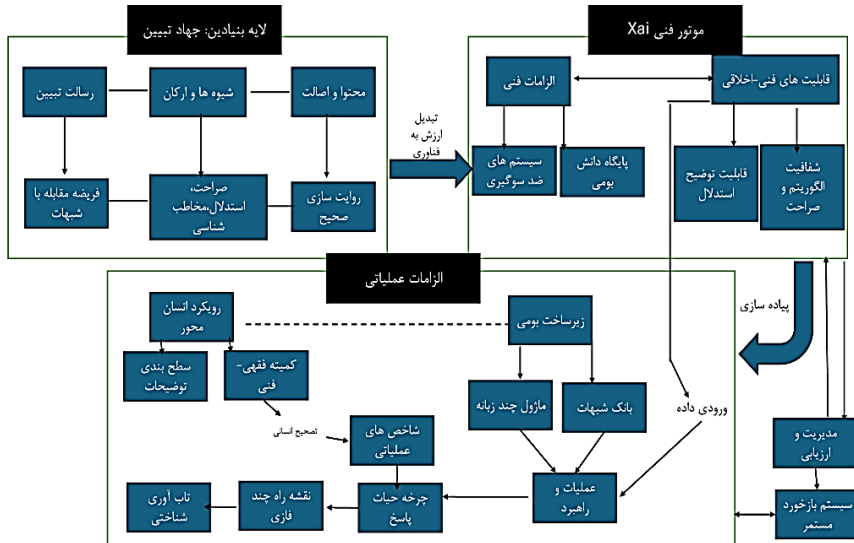
جدول (۶): لایه بندی ابرتم‌های انتزاعی

لایه	ابرتم‌ها
لایه بنیادین (جهاد تبیین)	رسالت تبیین (شامل تعریف و اهمیت)؛ شیوه‌ها و ارکان (روش‌ها و مخاطبان)؛ محتوا و اصالت (موضوعات و ریشه‌های دینی)
لایه میانی (XAI)	قابلیت‌های فنی-اخلاقی (اصول XAI و حاکمیت)؛ الزامات فنی (روش‌ها و تطبیق فرهنگی)
لایه بیرونی (الزامات عملیاتی)	رویکرد انسان‌محور (تلفیق انسان-در-حلقه و مخاطب‌شناسی)؛ زیرساخت بومی (پایگاه دانش و مدیریت سوگیری)؛ مدیریت و ارزیابی (حاکمیت نهادی و KPI ها)؛ عملیات و راهبرد (زمان‌شناسی و اولویت‌های اجرایی)

## ۶. تفسیر و تبیین مدل مفهومی نهایی

در نهایت تمامی تحلیل‌های این پژوهش در یک مدل مفهومی محقق محور خلاصه می‌شوند. مدل مفهومی «هوش مصنوعی تبیینی با رویکرد جهاد تبیین» با تلفیق داده‌های کیفی حاصل از بیانات رهبری، ادبیات تحقیق و مصاحبه‌های خبرگان طراحی شده و سه لایه مجزا اما پیوسته را شامل می‌شود: لایه بنیادین با محوریت اهداف و ارزش‌های جهاد تبیین (مانند رسالت تبیین، مقابله با شبهات و اصالت محتوا) که به‌عنوان قطب‌نمای اخلاقی، چرایی پژوهش را تعیین می‌کند؛ لایه میانی به‌عنوان موتور فنی XAI که با قابلیت‌هایی چون شفافیت الگوریتم، سیستم‌های ضد سوگیری و پایگاه دانش بومی، ارزش‌های لایه بنیادین را به فناوری تبدیل می‌کند و لایه بیرونی که الزامات عملیاتی شامل رویکرد انسان‌محور (مانند نظارت کمیته‌های فقهی-فنی)، زیرساخت بومی و مدیریت ارزیابی را فراهم می‌کند تا تضمین نماید سیستم نه تنها در پیاده‌سازی واقعی کاربردی باشد، بلکه با ترکیب هوشمندانه نظارت انسانی و فناوری، مسئولیت‌پذیری، انطباق با ارزش‌های دینی و اثربخشی در خنثی‌سازی شبهات و تقویت اعتماد عمومی را محقق سازد.

شکل (۵): مدل مفهومی هوش مصنوعی Xai با رویکرد جهاد تبیین



### بحث و نتیجه‌گیری پژوهش

این پژوهش با هدف طراحی مدل مفهومی «هوش مصنوعی تبیینی با رویکرد جهاد تبیین» انجام شده و با اتکا به سه منبع اصلی داده (بیانات رهبری، ادبیات تحقیق و دیدگاه خبرگان) نشان می‌دهد که چنین مدلی صرفاً فنی نیست، بلکه باید تلفیقی از ارزش‌های راهبردی، قابلیت‌های فناورانه و الزامات عملیاتی باشد.

بر اساس نتایج حاصل از تحلیل تلفیقی، مؤلفه‌های بنیادین جهاد تبیین - مطابق با تحلیل بیانات رهبری - شامل سه محور اصلی اند:

- رسالت و ماهیت (فریضه بودن و مقابله با دشمن)،
- شیوه‌ها و ارکان (استدلال، صراحت و مخاطب‌شناسی)،
- محتوا و اصالت (روایت‌سازی صحیح).

در مقابل، مرور ادبیات تخصصی نشان می‌دهد که مؤلفه‌های کلیدی هوش مصنوعی قابل توضیح (XAI) شامل قابلیت توضیح، شفافیت، انصاف و پاسخ‌گویی هستند که به‌عنوان ابزارهای فناورانه در مدل ایفای نقش می‌کنند.

مدل مفهومی یکپارچه طراحی شده در فصل چهارم، با ایجاد سه لایه مجزا و در عین حال مرتبط - لایه بنیادین جهاد تبیین، لایه میانی XAI، و لایه بیرونی الزامات عملیاتی - این ترکیب را محقق ساخته است. این مدل نشان می‌دهد که برای دستیابی به یک چارچوب عملیاتی مؤثر، لازم است مفاهیم نظری با الزامات اجرایی نظیر حلقه انسانی در تصمیم‌گیری، بهره‌گیری از پایگاه دانش بومی و سازوکارهای ارزیابی عملکردی تکمیل شوند.

محقق پیشنهاد می‌کند که ارتقای کارآمدی جهاد تبیین در عصر دیجیتال، مستلزم گذار پارادایمی به «جهاد تبیین هوشمند» است؛ رویکردی نوین که با بهره‌گیری از ظرفیت‌های هوش مصنوعی، الگوی سنتی «پخش پیام» را به سمت الگوی «گفت‌وگو و تعامل» سوق می‌دهد. این تحول مفهومی در سه بُعد کلیدی قابل تبیین است:

۱. از مونولوگ به دیالوگ: استفاده از هوش مصنوعی برای شنیدن واقعی صدای جامعه از طریق تحلیل داده‌های کلان و پردازش زبان طبیعی، با هدف شناخت دقیق دغدغه‌ها، نیازها و ارزش‌های مخاطبان.
۲. از پیام یکسان به محتوای شخصی‌سازی شده: تولید محتوا متناسب با خرده‌فرهنگ‌ها، گروه‌های اجتماعی و سطوح مختلف مخاطبان، با بهره‌گیری از قابلیت‌های تطبیقی و تحلیلی هوش مصنوعی.
۳. از واکنش کند به پیش‌بینی و پاسخ سریع: به‌کارگیری هوش مصنوعی برای رصد مستمر روندهای گفتمانی، پیش‌بینی بحران‌های احتمالی و ارائه پاسخ‌های دقیق، هدفمند و به‌موقع در مواجهه با تحولات رسانه‌ای.

این تحقیق تأکید می‌کند که ارتقای جهاد تبیین در عصر دیجیتال، مستلزم گذار از سطحی صرفاً انسانی به فرآیندی داده‌محور و هوشمند است. در این چارچوب، «منطق و استدلال» به مؤلفه «توضیح‌پذیری» در هوش مصنوعی ترجمه می‌شود، «مخاطب‌شناسی» به قابلیت «تطبیق پویا» در تولید محتوا بدل می‌گردد، و «مسئولیت‌پذیری نخبگان» در قالب مکانیسم «انسان در حلقه تصمیم‌گیری» (Human-in-the-Loop) در سیستم‌های هوشمند بازتعریف می‌شود.

در این پارادایم، نقش انسان نه تنها حذف نمی‌شود، بلکه توانمندتر می‌گردد؛ هوش مصنوعی جایگزین انسان نیست، بلکه تسهیل‌گر و مکمل او در فرآیند پر حجم و پیچیده تبیین است. چراکه هیچ مدل یادگیری ماشین، حتی در پیشرفته‌ترین سطوح زبانی، قادر به جایگزینی جایگاه اجتهاد و تشخیص‌های فقهی نیست.

برای عملیاتی سازی مدل مفهومی یکپارچه-XAI جهاد تبیین، نقشه راهی سه‌فازی با بهره‌گیری از نظرسنجی خبرگان طراحی شده است. این نقشه راه، چارچوبی زمان‌بندی شده برای تبدیل ایده‌های نظری به یک سیستم ملی کارآمد ارائه می‌دهد و بر چهار اصل بنیادین استوار است:

۱. تدریجی بودن: حرکت مرحله‌به‌مرحله از پایه‌ریزی اولیه تا گسترش ملی، با مدیریت ریسک و کنترل مخاطرات اجرایی.

۲. تلفیق نهادی: ایجاد هم‌افزایی میان نهادهای فنی، حاکمیتی و دینی به منظور تضمین انسجام ساختاری و مشروعیت اجتماعی مدل.

۳. یادگیری مبتنی بر شواهد: به‌روزرسانی مستمر مدل بر اساس داده‌های واقعی، بازخوردهای عملیاتی و تجربیات میدانی.

۴. انعطاف‌پذیری: قابلیت تطبیق با تحولات فناورانه، تغییرات محیطی و نیازهای متغیر ملی در حوزه تبیین و حکمرانی دیجیتال.

فاز ۱: پایه‌ریزی (۶ ماه) ایجاد زیرساخت‌های بنیادین

هدف: طراحی و استقرار چارچوب‌های نظری، فنی و حاکمیتی اولیه

اقدام کلیدی	مسئول اجرایی	خروجی مورد نیاز	شاخص موفقیت
تدوین چارچوب فقهی-فنی	شورای عالی فضای مجازی + مجمع فقهی	سند منشور حکمرانی XAI بومی	تصویب سند توسط ۳ نهاد مرتبط
توسعه پایگاه دانش بومی	سازمان تبلیغات اسلامی + وزارت ارتباطات	بانک اطلاعاتی شبهات (۱۰,۰۰۰ مورد اولیه)	پوشش ۹۰٪ شبهات پرتکرار
طراحی معماری فنی اولیه	پارک فناوری اطلاعات + دانشگاه‌ها	مستندات فنی + نمونه اولیه (MVP)	تایید معماری توسط ۵ خبره فنی
تشکیل کمیته‌های فقهی-فنی	نهاد نمایندگی مقام معظم رهبری	ساختار سازمانی + آیین‌نامه کار+تاییدیه شورای نگهبان	فعال‌سازی ۳ کمیته استانی آزمایشی

فاز ۲: آزمایشی (۴ ماه) - پیاده‌سازی و ارزیابی در محیط کنترل شده  
 هدف: تست عملکردی مدل در یک استان نمونه و اصلاح نقاط ضعف

شاخص موفقیت	خروجی مورد نیاز	مسئول اجرایی	اقدام کلیدی
پوشش ۱۰۰٪ شهرهای استان نمونه	نسخه آزمایشی سیستم تبیین هوشمند	استانداری استان + مرکز ملی فضای مجازی	استقرار MVP در استان نمونه
رضایت حداقل ۶۰٪ کاربران	۵۰۰ کاربر آموزش دیده (عمومی و نخبگان)	آموزش و پرورش استان + رسانه ملی + سازمان فنی و حرفه ای استان	آموزش کاربران پایلوت
شناسایی و رفع ۸۰٪ خطاهای سیستم	گزارش ارزیابی چندوجهی (فنی + شرعی + کاربردی)	پارک فناوری اطلاعات + دانشگاه‌ها + مرکز تحقیقات استراتژیک	ارزیابی عملکردی
افزایش ۲۰٪ دقت تشخیص شبهات	نسخه بهبود یافته MVP	شرکت‌های دانش بنیان + دانشگاه‌ها	بهینه‌سازی الگوریتم‌ها

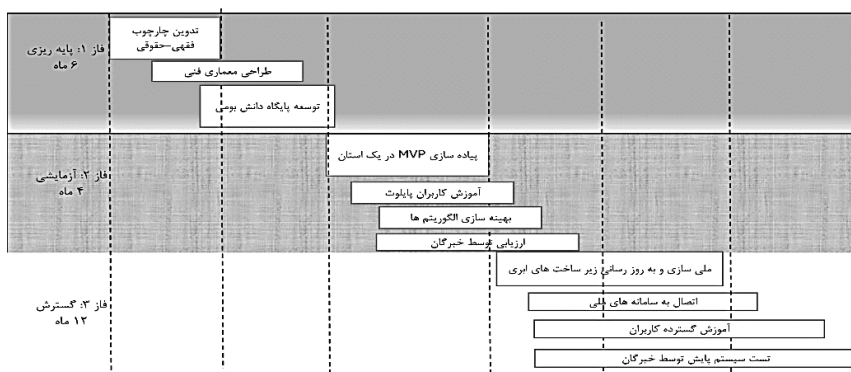
فاز ۳: گسترش (۱۲ ماه) - ملی‌سازی و یکپارچه‌سازی

هدف: مقیاس‌گذاری سیستم به سطح ملی و اتصال به زیرساخت‌های حیاتی کشور.

شاخص موفقیت	خروجی مورد نیاز	مسئول اجرایی	اقدام کلیدی
ظرفیت پردازش ۱ میلیون درخواست در روز	مرکز داده ملی + پلتفرم پردازش توزیع شده	وزارت ICT + شرکت‌های ارائه‌دهنده خدمات ابری	توسعه زیرساخت ابری ملی
تبادل داده بی‌درنگ با سامانه‌ها	یکپارچه‌سازی با ۱۰ سامانه ملی	سازمان تبلیغات اسلامی + مرکز رسانه‌های دیجیتال + وزارت ict	اتصال به سامانه‌های ملی
پوشش ۹۵٪ استان‌ها	۱۰,۰۰۰ کاربر آموزش دیده در سراسر کشور	رسانه ملی + آموزش و پرورش + نهادهای دینی + بسیج مردمی	آموزش گسترده کاربران
کاهش ۳۰٪ شایعات در مناطق تحت پوشش	داشبورد مدیریتی + گزارش‌های ماهانه	مرکز ملی فضای مجازی + مرکز تحقیقات استراتژیک	راه‌اندازی سیستم پایش عملکرد

این نقشه راه با هدف مدیریت ریسک و تضمین کارآمدی طراحی شده است؛ به گونه‌ای که فازبندی از پایه‌ریزی تا گسترش، امکان رفع نقاط ضعف پیش از مقیاس‌گذاری ملی را فراهم کرده و اجرای آزمایشی در یک استان نمونه، ریسک‌های فنی و اجتماعی را کاهش می‌دهد. توسعه زیرساخت ابری ملی و استقرار سیستم پایش عملکرد نیز پایداری و به‌روزرسانی مستمر مدل را تضمین می‌کند. علاوه بر این، تعریف دقیق مسئولیت‌ها مانع از تداخل وظایف می‌شود و کمیته‌های فقهی-فنی، هم‌سویی فناوری با ارزش‌های اسلامی و هم‌افزایی نهادی را تقویت می‌کنند. در نهایت، شاخص‌های قابل اندازه‌گیری و گزارش‌های چندوجهی، اعتبار علمی و عملیاتی سیستم را تضمین کرده و بر نتایج عینی تمرکز دارند.

شکل (۶): نقشه راه پیاده‌سازی مدل مفهومی



در نهایت، این پژوهش نشان می‌دهد که هوش مصنوعی نه تنها یک ابزار برای تسهیل امور، بلکه یک نقطه عطفی راهبردی است که می‌تواند با تبیین هوشمندانه حقایق، به تقویت بنیان‌های فکری و شناختی جامعه کمک کند.

## فهرست منابع

- امینی، س. ج.، محمدیان، م.، و سلیم، ح. (۱۴۰۴). الگوی راهبردی مدیریت رسانه‌های اجتماعی در عرصه جهاد تبیین. دین و ارتباطات (دانشگاه امام صادق/ نامه صادق).  
([https://rc.isu.ac.ir/article\\_77669.html](https://rc.isu.ac.ir/article_77669.html))
- بهشتی‌نژاد، س. م.، و رضازاده، ر. (۱۴۰۴). نقش جهاد تبیین در فضای مجازی برای پیشگیری از بحران. مدیریت بحران‌های اجتماعی، ۱۷(۱)، ۹۵-۶۳. ([https://cmj.ihu.ac.ir/article\\_209951.html](https://cmj.ihu.ac.ir/article_209951.html))
- عراقی، ع.، بیدگلی، م.، و رجبی ده‌برزویی، ا. (۱۴۰۱). واکاوی اهداف جنگ شناختی دشمن و راهکارهای تاب‌آوری مقابله با آن با تأکید بر آموزه‌های قرآن. فصلنامه علمی مطالعات دفاع مقدس، ۸(۴)، ۱۴۳-۱۶۲. (<https://dor.isc.ac/dor/20.1001.1.37883674.1401.8.4.7.6>)
- عزیزی، م.، و پیرانی، ش. (۱۴۰۱). کاربست نرم‌افزاری فضای مجازی در جهاد تبیین. فصلنامه علمی پژوهش‌های انقلاب اسلامی، ۱۱(۴)، ([https://www.roir.ir/article\\_168676.html](https://www.roir.ir/article_168676.html))
- طباطبایی، س. ف.، محمدحسین‌زاده، ع.، و اکبرپور فرسنگی، س. (۱۴۰۲). جهاد تبیین از منظر آیت‌الله خامنه‌ای و مبانی فقهی آن. فصلنامه علمی مطالعات بیداری اسلام، ۱۲(۳)، ۱۴۳-۱۶۸. ([https://www.iabaj.ir/article\\_180733.html](https://www.iabaj.ir/article_180733.html))
- غفاری، س. ا.، مهدیان، ح.، سیدی، س.، و ترکاشوند، ج. (۱۴۰۳). مدل نظام‌مند جهاد تبیین در اندیشه مقام معظم فرماندهی کل قوا (حفظ‌الله تعالی). فصلنامه مدیریت نظامی، ۲۴(۹۳)، ۲۷-۱. ([https://jmm.iranjournals.ir/article\\_713680.html](https://jmm.iranjournals.ir/article_713680.html))
- محبوب، ح.، و شکوری، س. (۱۴۰۱). جنگ‌شناختی مدرن: از شناخت در رزم تا عرصه جنگ‌شناختی. مطالعات منابع انسانی، ۱۲(۲)، ۱۵۶-۱۸۰. (<https://doi.org/10.22034/jhrs.2022.158895>)
- محمودی، م.، و صادقی، س. (۱۴۰۱). جهاد تبیین در فضای مجازی از منظر رهبر انقلاب. در سومین کنفرانس ملی پدافند سایبری، مراغه. (<https://civilica.com/doc/1493906>)
- ملک‌زاده، م. (۱۴۰۳). کارکردها و پیامدهای جهاد تبیین در بین‌المللی شدن الگوی مقاومت از دیدگاه آیت‌الله خامنه‌ای. فصلنامه مدیریت نظامی، ۲۴(۱۰)، ۱۷۶-۱۵۷. ([https://mag.rcipt.ir/article\\_181105.html](https://mag.rcipt.ir/article_181105.html))

- نیکونهاد، ا.، و بیوک، م. (۱۴۰۳). طراحی و تصویرسازی مطلوب از «جهاد تبیین» در تقابل با پدیده جنگ شناختی مبتنی بر اندیشه‌های رهبری معظم انقلاب. آینده‌پژوهی انقلاب اسلامی، (۱۱)۵، ۳۶-۱۱. (<https://dor.isc.ac/dor/20.1001.1.27173674.1403.5.1.1.3>)
- رضایی، د.، و ابراهیمی، ح. (۱۴۰۳). شناسایی عوامل، چالش‌ها و راهکارهای مقابله با جنگ شناختی با تمرکز بر نسل زد در ایران. فصلنامه مدیریت دفاع هوایی، ۳(۲)، ۲۷-۴۲. ([https://jadm.khadu.ir/article\\_125.html?lang=fa](https://jadm.khadu.ir/article_125.html?lang=fa))
- فیش‌های بیانات رهبر معظم انقلاب در خصوص جهاد تبیین. (۱۳۹۵-۱۴۰۴). پایگاه اطلاع‌رسانی دفتر حفظ و نشر آثار حضرت آیت‌الله‌العظمی خامنه‌ای. (<https://farsi.khamenei.ir/newspart-index-old?tid=26000#128187>)
- Al Momani, I. (2025). Ethical challenges for using artificial intelligence in understanding Islamic jurisprudence. *Salud, Ciencia y Tecnología-Serie de Conferencias*, (4), 1519. (<http://dx.doi.org/10.56294/sctconf20251519>).
- Al-Shammari, E. T. (2025, April). Culturally and computationally adaptive social robots for Arabophone Islamic societies. In 2024 International Conference on IT Innovation and Knowledge Discovery (ITIKD) (pp. 1–6). IEEE. (<https://doi.org/10.1109/ITIKD63574.2025.11005245>).
- Arrieta, A. G. J., et al. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. (<https://doi.org/10.1016/j.inffus.2019.12.012>).
- Babiker, A., et al. (2025). Attitude towards AI: Potential influence of conspiracy belief, XAI experience and locus of control. *International Journal of Human-Computer Interaction*, 41(13), 7939–7951. (<https://doi.org/10.1080/10447318.2024.2401249>).
- Bang, T. C. (2024). Ensuring credibility and trustworthiness in qualitative inquiries. In *Applied Linguistics and Language Education Research Methods: Fundamentals and Innovations* (pp. 70–85). IGI Global. (<https://doi.org/10.4018/979-8-3693-2603-9.ch006>).
- Belghachi, M. (2023). A review on explainable artificial intelligence methods, applications, and challenges. *Indonesian Journal of Electrical*

- Engineering and Informatics (IJEEI), 11(4), 1007–1024. (<http://dx.doi.org/10.52549/ijeei.v11i4.5151>).
- Bhavaneetharan, L. (2025, February). A comprehensive review of bias detection and XAI techniques in NLP. In 2025 5th International Conference on Advanced Research in Computing (ICARC) (pp. 1–6). IEEE. (<https://doi.org/10.1109/ICARC64760.2025.10962966>).
- Braun, V., & Clarke, V. (2024). Supporting best practice in reflexive thematic analysis reporting in palliative medicine: A review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). *Palliative Medicine*, 38(6), 608–616. (<https://doi.org/10.1177/02692163241234800>).
- Creswell, J. W., & Inoue, M. (2025). A process for conducting mixed methods data analysis. *Journal of General and Family Medicine*, 26(1), 4–11. (<https://doi.org/10.1002/jgf2.736>).
- Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (XAI): A survey. arXiv preprint arXiv:2006.11371. (<https://doi.org/10.48550/arXiv.2006.11371>).
- Deppe, C., & Schaal, G. S. (2024). Cognitive warfare: A conceptual analysis of the NATO ACT cognitive warfare exploratory concept. *Frontiers in Big Data*, 7, 1452129. (<https://doi.org/10.3389/fdata.2024.1452129>).
- Dwivedi, R., et al. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys*, 55(9), 1–33. (<https://doi.org/10.1145/3561048>).
- Ennab, M., & Mcheick, H. (2025). Advancing AI interpretability in medical imaging: A comparative analysis of pixel-level interpretability and Grad-CAM models. *Machine Learning and Knowledge Extraction*, 7(1), 12. (<https://doi.org/10.3390/make7010012>).
- Fenstermacher, L., Uzcha, D., Larson, K., Vitiello, C., & Shellman, S. (2023, June). New perspectives on cognitive warfare. In *Signal Processing, Sensor/Information Fusion, and Target Recognition XXXII* (Vol. 12547, pp. 172–187). SPIE. (<https://doi.org/10.1117/12.2666777>).

- Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine*, 40(2), 44–58. (<https://doi.org/10.1609/aimag.v40i2.2850>).
- Islam, S. R., Eberle, W., Ghafoor, S. K., & Ahmed, M. (2021). Explainable artificial intelligence approaches: A survey. *arXiv preprint arXiv:2101.09429*. (<https://doi.org/10.48550/arXiv.2101.09429>)
- Machlev, R., et al. (2022). Explainable artificial intelligence (XAI) techniques for energy and power systems: Review, challenges and opportunities. *Energy and AI*, 9, 100169. (<http://dx.doi.org/10.1016/j.egyai.2022.100169>)
- Mustafa, A. M., et al. (2024). Interpreting Arabic transformer models: A study on XAI interpretability for Qur'anic semantic-search models. *Jordanian Journal of Computing and Information Technology*, 10. (<https://doi.org/10.5455/jjcit.71-1704878720>).
- Shafie, R. S., et al. (2024). Psychosemantics of “Us” and “Them” in social media warfare. *SEARCH Journal of Media and Communication Research*, 16(2), 65–81. (<http://dx.doi.org/10.58946/search-16.2.P5>).
- Stanek, D., Klaban, I., & Coufalíková, A. (2025, May). Explainable artificial intelligence: State of the art and beyond. In *2025 International Conference on Military Technologies (ICMT)* (pp. 1–7). IEEE. (<https://ieeexplore.ieee.org/document/11005246>).
- Zhang, H., & Ogasawara, K. (2023). Grad-CAM-based explainable artificial intelligence related to medical text processing. *Bioengineering*, 10(9), 1070. (<https://doi.org/10.3390/bioengineering10091070>)

## References

- Amini, S. J., Mohammadian, M., & Salim, H. (2025/1404 SH). Strategic model for social media management in the field of clarification jihad. \*Religion and Communications (Imam Sadiq University / Nameh Sadiq)\*. ([[https://rc.isu.ac.ir/article\\_77669.html](https://rc.isu.ac.ir/article_77669.html)])([https://rc.isu.ac.ir/article\\_77669.html](https://rc.isu.ac.ir/article_77669.html)))
- Beheshtinejad, S. M., & Reza-zadeh, R. (2025/1404 SH). The role of clarification jihad in cyberspace for crisis prevention. \*Social Crisis Management\*, 17(1), 63–95. ([[https://cmj.ihu.ac.ir/article\\_209951.html](https://cmj.ihu.ac.ir/article_209951.html)])([https://cmj.ihu.ac.ir/article\\_209951.html](https://cmj.ihu.ac.ir/article_209951.html)))
- Araqi, A., Bidgoli, M., & Rajabi Dehbarzouei, A. (2022/1401 SH). Analysis of the goals of enemy cognitive warfare and strategies for resilience with emphasis on Quranic teachings. \*Scientific Quarterly of Sacred Defense Studies\*, 8(4), 143–162. ([<https://dor.isc.ac/dor/20.1001.1.37883674.1401.8.4.7.6>])(<https://dor.isc.ac/dor/20.1001.1.37883674.1401.8.4.7.6>))
- Azizi, M., & Pirani, Sh. (2022/1401 SH). Software applications of cyberspace in clarification jihad. \*Scientific Quarterly of Islamic Revolution Research\*, 11(4). ([[https://www.roir.ir/article\\_168676.html](https://www.roir.ir/article_168676.html)])([https://www.roir.ir/article\\_168676.html](https://www.roir.ir/article_168676.html)))
- Tabatabaei, S. F., Mohammad-Hosseini-zadeh, A., & Akbarpour-Farsangi, S. (2023/1402 SH). Clarification jihad from the perspective of Ayatollah Khamenei and its jurisprudential foundations. \*Scientific Quarterly of Islamic Awakening Studies\*, 12(3), 143–168. ([[https://www.iabaj.ir/article\\_180733.html](https://www.iabaj.ir/article_180733.html)])([https://www.iabaj.ir/article\\_180733.html](https://www.iabaj.ir/article_180733.html)))
- Ghaffari, S. A., Mahdian, H., Seydi, S., & Tarkashvand, J. (2024/1403 SH). A systematic model of clarification jihad in the thought of the Supreme Leader. \*Military Management Quarterly\*, 24(93), 1–27. ([[https://jmm.iranjournals.ir/article\\_713680.html](https://jmm.iranjournals.ir/article_713680.html)])([https://jmm.iranjournals.ir/article\\_713680.html](https://jmm.iranjournals.ir/article_713680.html)))

- Mahjoob, H., & Shakouri, S. (2022/1401 SH). Modern cognitive warfare: from battlefield understanding to the cognitive warfare domain. *\*Human Resources Studies\**, 12(2), 156–180. ([<https://doi.org/10.22034/jhrs.2022.158895>](<https://doi.org/10.22034/jhrs.2022.158895>))
- Mahmoudi, M., & Sadeghi, S. (2022/1401 SH). Clarification jihad in cyberspace from the perspective of the Supreme Leader. In *\*Third National Conference on Cyber Defense\**, Maragheh. ([<https://civilica.com/doc/1493906>](<https://civilica.com/doc/1493906>))
- Malekzadeh, M. (2024/1403 SH). Functions and consequences of clarification jihad in the internationalization of the resistance model from Ayatollah Khamenei's perspective. *\*Military Management Quarterly\**, 24(10), 157–176. ([[https://mag.rcipt.ir/article\\_181105.html](https://mag.rcipt.ir/article_181105.html)]([https://mag.rcipt.ir/article\\_181105.html](https://mag.rcipt.ir/article_181105.html)))
- Nikonehad, A., & Biouk, M. (2024/1403 SH). Designing and visualizing an optimal clarification jihad in confrontation with cognitive warfare based on the Supreme Leader's thought. *\*Islamic Revolution Futures Studies\**, 5(1), 11–36. ([<https://dor.isc.ac/dor/20.1001.1.27173674.1403.5.1.1.3>](<https://dor.isc.ac/dor/20.1001.1.27173674.1403.5.1.1.3>))
- Rezaei, D., & Ebrahimi, H. (2024/1403 SH). Identifying factors, challenges, and strategies to counter cognitive warfare with focus on Generation Z in Iran. *\*Air Defense Management Quarterly\**, 3(2), 27–42. ([[https://jadm.khadu.ir/article\\_125.html?lang=fa](https://jadm.khadu.ir/article_125.html?lang=fa)]([https://jadm.khadu.ir/article\\_125.html?lang=fa](https://jadm.khadu.ir/article_125.html?lang=fa)))
- Extracts from the Supreme Leader's statements on clarification jihad (2016–2025/1395–1404 SH). Official Website of the Office for Preservation and Publication of Ayatollah Khamenei's Works. ([<https://farsi.khamenei.ir/newspart-index-old?tid=26000#128187>](<https://farsi.khamenei.ir/newspart-index-old?tid=26000#128187>))

- Al Momani, I. (2025). Ethical challenges for using artificial intelligence in understanding Islamic jurisprudence. \*Salud, Ciencia y Tecnología-Serie de Conferencias\*, (4), 1519. ([<http://dx.doi.org/10.56294/sctconf20251519>])(<http://dx.doi.org/10.56294/sctconf20251519>))
- Al-Shammari, E. T. (2025, April). Culturally and computationally adaptive social robots for Arabophone Islamic societies. In \*2024 International Conference on IT Innovation and Knowledge Discovery (ITIKD)\* (pp. 1–6). IEEE. ([<https://doi.org/10.1109/ITIKD63574.2025.11005245>])(<https://doi.org/10.1109/ITIKD63574.2025.11005245>))
- Arrieta, A. G. J., et al. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. \*Information Fusion\*, 58, 82–115. ([<https://doi.org/10.1016/j.inffus.2019.12.012>])(<https://doi.org/10.1016/j.inffus.2019.12.012>))
- Babiker, A., et al. (2025). Attitude towards AI: Potential influence of conspiracy belief, XAI experience and locus of control. \*International Journal of Human–Computer Interaction\*, 41(13), 7939–7951. ([<https://doi.org/10.1080/10447318.2024.2401249>])(<https://doi.org/10.1080/10447318.2024.2401249>))
- Bang, T. C. (2024). Ensuring credibility and trustworthiness in qualitative inquiries. In \*Applied Linguistics and Language Education Research Methods: Fundamentals and Innovations\* (pp. 70–85). IGI Global. ([<https://doi.org/10.4018/979-8-3693-2603-9.ch006>])(<https://doi.org/10.4018/979-8-3693-2603-9.ch006>))
- Belghachi, M. (2023). A review on explainable artificial intelligence methods, applications, and challenges. \*Indonesian Journal of Electrical Engineering and Informatics (IJEI)\*, 11(4), 1007–1024. ([<http://dx.doi.org/10.52549/ijeel.v11i4.5151>])(<http://dx.doi.org/10.52549/ijeel.v11i4.5151>))
- Bhavaneetharan, L. (2025, February). A comprehensive review of bias detection and XAI techniques in NLP. In \*2025 5th International Conference on Advanced Research in Computing (ICARC)\* (pp. 1–6). IEEE. ([<https://doi.org/10.1109/ICARC64760.2025.10962966>])(<https://doi.org/10.1109/ICARC64760.2025.10962966>))

- Braun, V., & Clarke, V. (2024). Supporting best practice in reflexive thematic analysis reporting in palliative medicine: A review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). *\*Palliative Medicine\**, 38(6), 608–616. ([<https://doi.org/10.1177/02692163241234800>](<https://doi.org/10.1177/02692163241234800>))
- Creswell, J. W., & Inoue, M. (2025). A process for conducting mixed methods data analysis. *\*Journal of General and Family Medicine\**, 26(1), 4–11. ([<https://doi.org/10.1002/jgf2.736>](<https://doi.org/10.1002/jgf2.736>))
- Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (XAI): A survey. arXiv preprint arXiv:2006.11371. ([<https://doi.org/10.48550/arXiv.2006.11371>](<https://doi.org/10.48550/arXiv.2006.11371>))
- Deppe, C., & Schaal, G. S. (2024). Cognitive warfare: A conceptual analysis of the NATO ACT cognitive warfare exploratory concept. *\*Frontiers in Big Data\**, 7, 1452129. ([<https://doi.org/10.3389/fdata.2024.1452129>](<https://doi.org/10.3389/fdata.2024.1452129>))
- Dwivedi, R., et al. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *\*ACM Computing Surveys\**, 55(9), 1–33. ([<https://doi.org/10.1145/3561048>](<https://doi.org/10.1145/3561048>))
- Ennab, M., & Mcheick, H. (2025). Advancing AI interpretability in medical imaging: A comparative analysis of pixel-level interpretability and Grad-CAM models. *\*Machine Learning and Knowledge Extraction\**, 7(1), 12. ([<https://doi.org/10.3390/make7010012>](<https://doi.org/10.3390/make7010012>))
- Fenstermacher, L., Uzcha, D., Larson, K., Vitiello, C., & Shellman, S. (2023, June). New perspectives on cognitive warfare. In *\*Signal Processing, Sensor/Information Fusion, and Target Recognition XXXII\** (Vol. 12547, pp. 172–187). SPIE. ([<https://doi.org/10.1117/12.2666777>](<https://doi.org/10.1117/12.2666777>))

- Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *\*AI Magazine\**, 40(2), 44–58. ([<https://doi.org/10.1609/aimag.v40i2.2850>])(<https://doi.org/10.1609/aimag.v40i2.2850>)
- Islam, S. R., Eberle, W., Ghafoor, S. K., & Ahmed, M. (2021). Explainable artificial intelligence approaches: A survey. arXiv preprint arXiv:2101.09429. ([<https://doi.org/10.48550/arXiv.2101.09429>])(<https://doi.org/10.48550/arXiv.2101.09429>)
- Machlev, R., et al. (2022). Explainable artificial intelligence (XAI) techniques for energy and power systems: Review, challenges and opportunities. *\*Energy and AI\**, 9, 100169. ([<http://dx.doi.org/10.1016/j.egyai.2022.100169>])(<http://dx.doi.org/10.1016/j.egyai.2022.100169>)
- Mustafa, A. M., et al. (2024). Interpreting Arabic transformer models: A study on XAI interpretability for Qur'anic semantic-search models. *\*Jordanian Journal of Computing and Information Technology\**, 10. ([<https://doi.org/10.5455/jjcit.71-1704878720>])(<https://doi.org/10.5455/jjcit.71-1704878720>)
- Shafie, R. S., et al. (2024). Psychosemantics of “Us” and “Them” in social media warfare. *\*SEARCH Journal of Media and Communication Research\**, 16(2), 65–81. ([<http://dx.doi.org/10.58946/search-16.2.P5>])(<http://dx.doi.org/10.58946/search-16.2.P5>)
- Stanek, D., Klaban, I., & Coufalíková, A. (2025, May). Explainable artificial intelligence: State of the art and beyond. In *\*2025 International Conference on Military Technologies (ICMT)\** (pp. 1–7). IEEE. ([<https://ieeexplore.ieee.org/document/11005246>])(<https://ieeexplore.ieee.org/document/11005246>)
- Zhang, H., & Ogasawara, K. (2023). Grad-CAM-based explainable artificial intelligence related to medical text processing. *\*Bioengineering\**, 10(9), 1070. ([<https://doi.org/10.3390/bioengineering10091070>])(<https://doi.org/10.3390/bioengineering10091070>).