IEEE Open Journal of the Communications Society

Received 31 January 2025; revised 20 March 2025; accepted 20 April 2025. Date of publication 22 April 2025; date of current version 9 May 2025.

Digital Object Identifier 10.1109/OJCOMS.2025.3563415

Explainable Reinforcement and Causal Learning for Improving Trust to 6G Stakeholders

MIGUEL ARANA-CATANIA [1], AMIR SONEE [2] (Member, IEEE), ABDUL-MANAN KHAN [2], KAVAN FATEHI [3], YUN TANG [1] (Member, IEEE), BAILU JIN [1], ANNA SOLIGO [2], DAVID BOYLE [2] (Member, IEEE), RADU CALINESCU [3] (Senior Member, IEEE), POONAM YADAV [3] (Member, IEEE), HAMED AHMADI [3] (Senior Member, IEEE), ANTONIOS TSOURDOS [1] (Member, IEEE), WEISI GUO [1] (Senior Member, IEEE), AND ALESSANDRA RUSSO [2] (Member, IEEE)

[1] School of Aerospace, Transport and Manufacturing, Cranfield University, MK43 0AL Cranfield, U.K.
[2] Department of Computing, Imperial College London, SW7 2AZ London, U.K.
[3] Department of Computer Science, University of York, YO10 5DD York, U.K.

CORRESPONDING AUTHOR: W. GUO (e-mail: weisi.guo@cranfield.ac.uk)

This work was supported by the EPSRC CHEDDAR: Communications Hub for Empowering Distributed Cloud Computing Applications and Research under Grant EP/X040518/1 and Grant EP/Y037421/1.

(Miguel Arana-Catania, Amir Sonee, Abdul-Manan Khan, Kavan Fatehi, Yun Tang, and Bailu Jin contributed equally to this work.)

ABSTRACT Future telecommunications will increasingly integrate AI capabilities into network infrastructures to deliver seamless and harmonized services closer to end-users. However, this progress also raises significant trust and safety concerns. The machine learning systems orchestrating these advanced services will widely rely on deep reinforcement learning (DRL) to process multi-modal requirements datasets and make semantically modulated decisions, introducing three major challenges: (1) First, we acknowledge that most explainable AI research is stakeholder-agnostic while, in reality, the explanations must cater for diverse telecommunications stakeholders, including network service providers, legal authorities, and end users, each with unique goals and operational practices; (2) Second, DRL lacks prior models or established frameworks to guide the creation of meaningful long-term explanations of the agent's behaviour in a goal-oriented RL task, and we introduce state-of-the-art approaches such as reward machine and sub-goal automata that can be universally represented and easily manipulated by logic programs and verifiably learned by inductive logic programming of answer set programs; (3) Third, most explainability approaches focus on correlation rather than causation, and we emphasise that understanding causal learning can further enhance 6G network optimisation. Together, in our judgement they form crucial enabling technologies for trustworthy services in 6G. This review offers a timely resource for academic researchers and industry practitioners by highlighting the methodological advancements needed for explainable DRL (X-DRL) in 6G. It identifies key stakeholder groups, maps their needs to X-DRL solutions, and presents case studies showcasing practical applications. By identifying and analysing these challenges in the context of 6G case studies, this work aims to inform future research, transform industry practices, and highlight unresolved gaps in this rapidly evolving field.

INDEX TERMS 6G, explainable AI, reinforcement learning, trust, stakeholders, causal learning.

I. INTRODUCTION

A. CONTEXT AND MOTIVATION

NEVER before has telecommunication infrastructure been responsible for such a diverse range of services in human history, enabling transformative use cases that extend connectivity, intelligence, and immersive experiences across industries. To meet such dynamic service demands (bandwidth, coverage, latency, energy), the network needs to perform large-scale multi-objective optimisation over highly variable environments with partially observable dynamics that are constrained with limited resources. As such, 6G will require advanced artificial intelligence, especially

Reinforcement Learning (RL), to efficiently and dynamically manage heterogeneous infrastructure, optimize communication protocols, and orchestrate network resources in real-time [1].

However, the integration of ubiquitous AI in 6G introduces unprecedented challenges. For the first time on a large scale and in real-time, human well-being becomes deeply intertwined with 6G services. This extensive use of AI in 6G can potentially cause significant harm to human agency, safety, privacy, fairness, and social and environmental well-being [2], [3]. Ensuring the above principles requires the AI in 6G to be understandable for any stakeholders and in any context of the AI value chain, which can be addressed with proper XAI techniques.

B. REVIEW OF EXISTING RELATED SURVEYS

Extensive efforts have been put into proposing and applying novel XAI techniques for different AI-powered processes in 6G [4]. Review papers have also been published summarizing the benefits and challenges of XAI in 6G. For example, [5] covered a broad range of XAI application areas in 6G, including technical areas of 6G network (e.g., intelligent radio, trust and security, privacy, resource management, edge AI, Zero Touch Network and Service Management), AI-powered 6G use cases (e.g., intelligent health and wearable, industry 5.0, connected autonomous vehicles, smart grid 2.0, multi-sensory XR applications and smart governance) and XAI-related research projects and research challenges. Reference [6] reviews the XAI with a specific focus on O-RAN in 6G, including topics such as the deployment of XAI pipelines in O-RAN, potential applications of XAI in existing AI-driven O-RAN solutions, XAI for O-RAN use cases (Quality-of-Experience Optimization, traffic steering, user access management etc.), and research projects and standards on XAI for O-RAN. We identify the following gaps in the literature:

- *Gap 1*: Existing reviews primarily focus on stakeholder-agnostic XAI methodologies [6] or briefly address only the dimensions of "who" and "why" [5]. Limited discussion is present on "what" and "how," specifically concerning stakeholder-specific requirements and approaches.
- *Gap 2:* the reviewed XAI techniques often target a single ML/AI process that does not involve sequential learning as in reinforcement learning (RL), which is widely adopted to perform sequential decision-making for real-time 6G network optimization. This is basically due to the highly variable environments with partially observable dynamics that are constrained with limited resources to meet a large number of dynamic demands.
- *Gap 3:* the reviewed XAI techniques often focus on learning correlations between input data features and the outcomes of the AI models. This curbs the causal understanding of the network's decision-making processes.

C. MAIN CONTRIBUTIONS AND ORGANISATION OF PAPER

To address the limitations identified above, we identify the following research questions that we will try to answer in this review article:

1) *Stakeholders:* What stakeholders require what kinds of interpretability/explainability?
2) *Explainable Reinforcement Learning:* How can we explain the sequences of decisions made by an agent in RL-based approaches, as one of the promising solutions toward solving complicated 6G optimisation problems?
3) *Causal Understanding for Explainability and Performance:* How can the causal understanding of deep models help to improve AI/ML pipeline performance, from reducing variable search space to transfer learning and improving explainability?
4) *Current Challenges:* What are the remaining open challenges?

To demonstrate, we select 1) network slicing (NS) and 2) uncrewed aerial vehicles (UAVs) as two main mission-critical grounds in 6G where failing to provide explainable and safe decision-making can lead to catastrophic consequences. The reviewed XAI techniques will be contextualised in the two use cases from the stakeholder's perspective in practice:

- Network slicing is a key enabler for 6G networks, enabling virtualized and isolated slices managed by distinct resource policies. Each slice meets specific Quality-of-Service (QoS) requirements and Service-level Agreements (SLAs). Employing explainable decision-making for resource allocation—including resource block (RB) assignment, user admission control, and scheduling—offers transparency, aiding service providers in quickly identifying and mitigating service delivery issues for end users.
- Drones may be used as mobile base stations (BSs) and communication relays to enhance connectivity in complex and dynamic environments, underserved or disaster-hit areas. The UAV scenario is discussed here to provide insights into the challenges regarding the need to understand and control drone behaviour to ensure effective communication. Critical aspects involve making autonomous trajectory planning and service provisioning interpretable, explainable, and transparent for different stakeholders, thereby ensuring trust and legal compliance.

The paper organization is shown in Figure 2.

II. BACKGROUND

A key challenge in AI development is ensuring safe performance in unforeseen real-world situations not encountered during training or testing. According to EU AI ethics guidelines, AI systems require human oversight, accountability, and transparency, necessitating interpretability

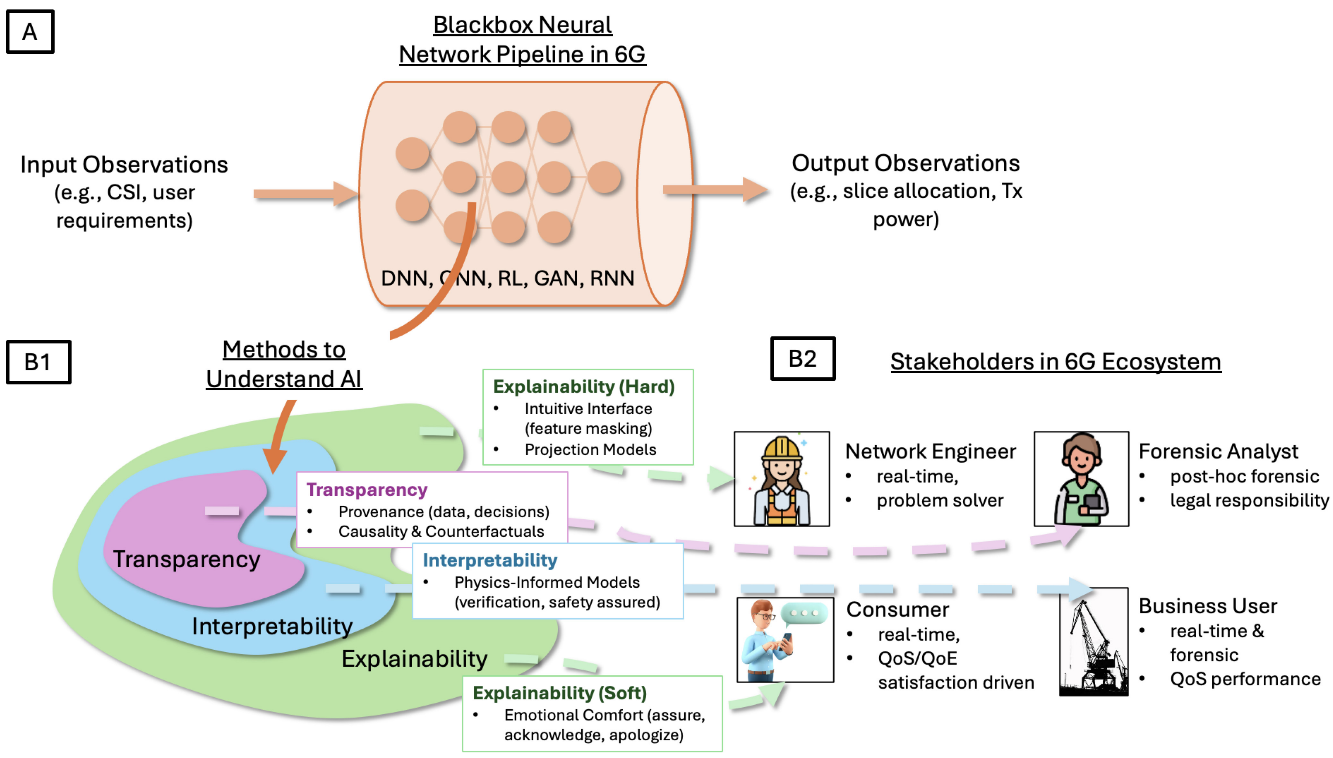

FIGURE 1. Forms of AI understanding methods and application to 6G stakeholder examples.

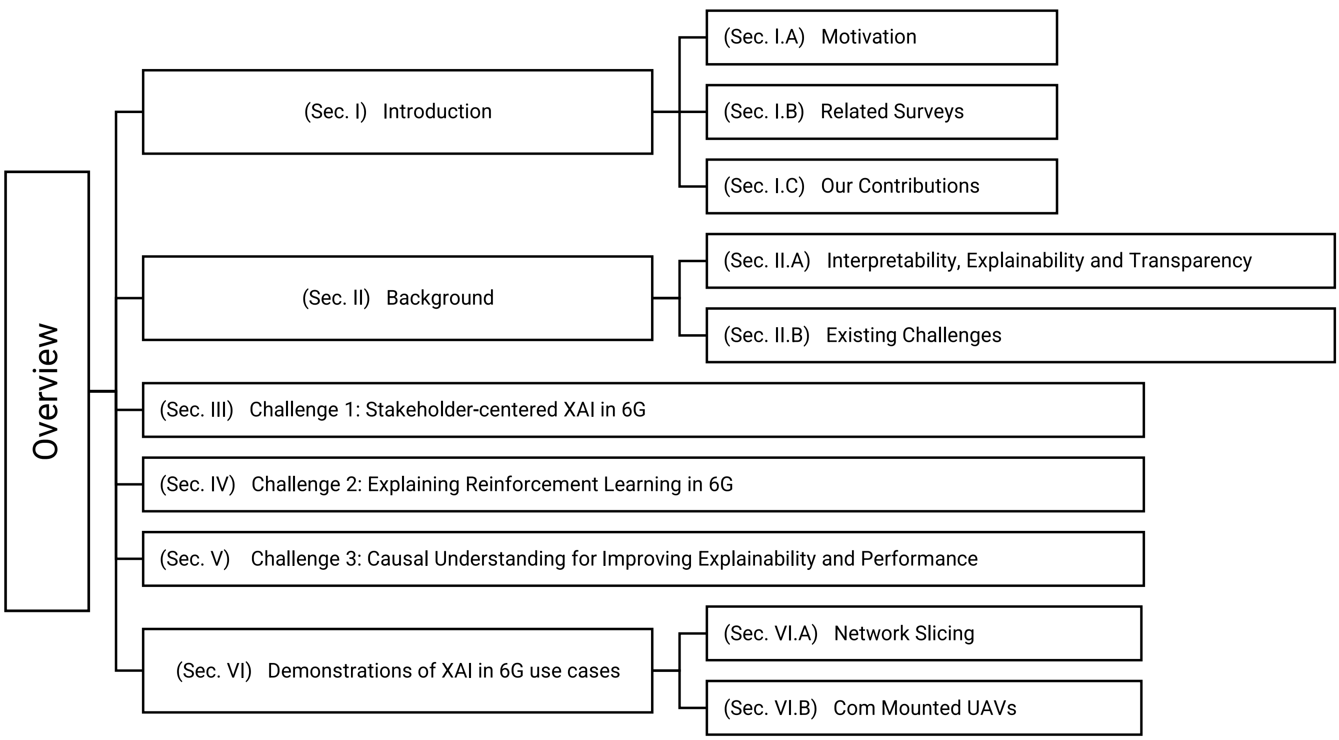

FIGURE 2. Paper organization.

and explainability for effective human engagement. Such engagement ranges from formal system verification during design to audits at runtime. However, we have yet to agree on clear definitions for related terms, often using interpretability, explainability, and transparency interchangeably. In this section, we focus primarily on *interpretability* and *explainability*, briefly addressing *transparency* within the context of XAI for 6G.

A. INTERPRETABILITY, EXPLAINABILITY, AND TRANSPARENCY OF AI

For clarity, it is worth distinguishing between the two most prevalent terms: *interpretability* and *explainability* [7]. Following Glanois et al. [8], interpretability may be defined as an inherent model quality describing the extent to which the inner workings of a model can be examined and understood. Interpretability is frequently enhanced through the use of "white-box" architectures, such as sets of first-order logic rules. Conversely, explainability allows us to describe an external understanding of model behaviour, based on active, *post hoc*, efforts to explain the decision-making process. Explainability techniques are typically applied to trained models to provide insights based on externally observed relations rather than internal mechanisms. An example is feature importance analysis, where input variables are systematically altered and the resulting impact on model outputs is quantified. Model interpretability and explainability share the common goal of increasing human understanding of the AI system and the insights gained through either approach should ultimately be communicated in a stakeholder-centered format, referred to in both cases as an "explanation".

Although interpretability offers insights which are better grounded in model internals, it is frequently associated with reduced model performance (e.g., higher Mean Square Error [9] in RAN slicing AI model), and methods are often architecture or application-specific, with limited scalability [8]. Explainability methods, conversely, can offer greater flexibility and broader applicability, making them valuable for gaining insights in complex scenarios where full internal interpretability may be neither necessary nor feasible. Importantly, interpretability and explainability approaches can be viewed as existing on a spectrum, and are not mutually exclusive: insightful results may be achieved through applying explainability methods to partially interpretable models, for instance [10].

Broadly, interpretability and explainability serve to address incompleteness in the formalisation of a problem, improving understanding of how a system will behave in new situations or concerning auxiliary criteria such as fairness [11]. In this sense, model explanations enable stakeholders to gain insights into abstract or complex attributes that are challenging to fully define or directly optimise for. These may include causality, reliability or scientific knowledge: facets of understanding which may be sufficient to alleviate ethical, legal and operational concerns about a system.

1) INTERPRETABILITY AND REINFORCEMENT LEARNING

For RL, both interpretability and explainability imply a level of understanding of the reasoning behind agent decisions. Despite a general consensus that this is desirable, there is, however, no precise definition of when it is necessary or what it entails, neither broadly nor for specific use cases. Consequently, there is a lack of standardised metrics or benchmarks for assessing the quality or utility of explanations. As Doshi-Velez and Kim [11] critically observe, the field often appears to default to a "you'll know it when you see it" approach.

Despite ambiguity regarding the definition, there exists a multitude of methods for increasing the interpretability of RL agents. We broadly categorise these as offering high or low-level understanding, with the latter further comprising direct and indirect approaches, and describe a non-exhaustive selection. Focus is placed on the primary challenge of understanding decision-making, but it is relevant to note that interpreting inputs and transition models are also important topics when considering complex architectures.

High-level interpretability approaches offer broad, generally top-down perspectives on agent-decision making. Hierarchical RL decomposes goals into sub-tasks (Feudal approaches) or sub-policies (policy tree approaches) [12]. This decomposition most effectively contributes to interpretability when discernible sub-behaviours, such as motor primitives [13], are explicitly learned. Alternatively, high-level interpretability can be obtained via direct incorporation of declarative knowledge into RL frameworks [14]. Examples include defining high-level rules to guide actions [15], and integrating knowledge-based reasoning paradigms with learning architectures [16]. However, the interpretability offered by these approaches is constrained: in HRL, interpretability is limited to the level of abstraction of the sub-task or policy, and in knowledge-based systems, it is broadly restricted to the scope of the incorporated knowledge. As evidenced by these approaches, imposing high-level interpretability frequently relies on prior knowledge of the desired RL solution and interpretation. This limits both applicability and performance: policy-tree HRL approaches, for instance, constrain the policy space and thus may not reach optimal policies.

Low-level interpretability can be directly achieved by learning more interpretable architectures for action selection. Decision trees are a prevalent example: acyclic graphs pass input variables through decision nodes, which select subsequent nodes based on feature values until a leaf node is reached. These trees can represent Q-values or policies, and recent work on 'soft decision trees' relaxes their classically discrete nature and enables efficient RL through gradient descent [17]. Alternatively, diverse methods exist to generate effective RL policies in the form of symbolic equations, e.g., using genetic programming to efficiently search a space of function trees [18] or training a recurrent neural network (RNN) to directly generate policy equations [19]. Beyond mathematical operators, policies have been learned as weighted combinations of first-order logic rules [20]. Recent efforts have increased the flexibility and scalability of this approach, for instance by weighting predicates, the building blocks of rules, rather than the rules themselves [21], or by using logical reasoning modules to induce separate policies [22].

Indirect approaches use similar underlying architectures but differ in the manner in which these are obtained. While direct approaches immediately search for interpretable policies, indirect approaches aim to replicate non-interpretable ones, using techniques analogous to policy distillation or imitation learning. VIPER, an imitation-learning-like algorithm which compresses DRL policies into decision trees, demonstrates how this indirect approach can improve scalability [23]. Similarly, RNN-based equation generation employs the indirect approach to scale to large action spaces by sequentially generating action-equations using a neural network "anchor" policy to select actions for which an equation is yet to be defined [19].

These low-level methods rely on using fundamentally different model architectures, which are generally deemed to be interpretable due to characteristics such as comprising fewer model components, containing transparent and tractable component interactions, and consisting of components which can be objectively translated to natural language. However, these architectural modifications can lead to issues regarding scalability and performance. Scaling the described methods rapidly becomes computationally prohibitive, even when an indirect approach is adopted [8]. Moreover, such scaling can compromise interpretability as a result of increasing the number of model components and their interactions. Notably, few existing low-level interpretability approaches apply their method to real-world RL problems, instead evaluating in simple control environments such as "Cart Pole" and the grid-based "Cliff Walk".

2) EXPLAINABILITY

Rather than directly deriving explanations from model internals, explainability approaches apply external methods to detect and describe input-output relations. This distinction can be illustrated using a third application of decision trees: as purely explanatory models [24]. Here, as for indirect interpretability, the decision tree is distilled from a deep neural network, with the distinction that it is used solely to explain action selection, which is still performed by the original neural network.

Explainability places a strong focus on the presentation of insights, where saliency maps or textual explanations are common approaches. Saliency maps highlight the image regions considered important for an agent's decision and can be constructed based on attention, gradients, feature perturbations, or object segmentation [25]. The result offers appealingly understandable visuals, but their utility and accuracy have been challenged. Atrey et al. [25] found their use in RL to be subjective and insufficiently falsifiable to be used as an explanatory tool, corroborating earlier works which raised their potentially misleading nature [26]. Alternative textual approaches frequently involve translating an agent's state-action space into human concepts and can enable interactive explainability. For example, by evaluating user queries with respect to a policy and transition model to generate natural language descriptions of expected action consequences [27].

The need to explicitly define concepts has recently been addressed with the integration of Large Language Models (LLMs), e.g., in autonomous driving research [28]. TransGPT [29] represents a state-of-the-art approach where a multimodal dataset is used to train a model capable of answering queries about driving actions. However, it lacks a direct evaluation of explanation accuracy, and the use of external models to retroactively justify input-output relations can be seen as a concerning example of how explainability may generate plausible-sounding explanations without genuinely improving understanding.

These post-hoc explainability methods offer high scalability and certain methods, such as perturbation-based feature analysis, allow for application to diverse model architectures. However, this is counterbalanced by a lack of grounding in model internals and the risk of offering misleading results. A challenge thus emerges regarding how the accuracy of post-hoc explanations can be effectively measured.

B. EXISTING CHALLENGES

As suggested by [5], there are multiple challenges and limitations in using XAI in 6G. There are no standard quantifiable metrics for XAI [30], where commonly the explanations consist of visual and textual inputs which can not be quantified. Efficient metrics seem to be stakeholder and domain-dependent, although some attempts have been made to propose general metrics for 6G XAI [3]. It is difficult to find the right trade-off between interpretability and performance [31], and between explainability and security and privacy [32]. Especially, a potential privacy leakage from XAI is an important challenge [33]. There are legal challenges related to the explainability of AI in 6G for regulatory compliance. In the European Union, the GDPR grants users the right to explanation in algorithmic decision-making, and failing to comply with the GDPR may result in fines up to 20 million euros or 4% of the company's global revenue. A serious engagement of legal experts is required to ensure legal compliance. Beyond that, a general engagement of all stakeholders involved is required to ensure that the explanations generated are appropriate for each of them.

> **Appropriate Explanations to Audience**: Interpreting and explaining AI models as well as improving transparency need to be audience and stakeholder-centric. It needs to reflect the usage scenario, the skill level, the operational environment, the legal requirements and the application context.

III. CHALLENGE 1: REQUIREMENT OF STAKEHOLDER-CENTERED XAI IN 6G

A. WHY STAKEHOLDER-CENTERED XAI

A stakeholder of an AI system in general can be any person, group or organization directly or indirectly involved in the

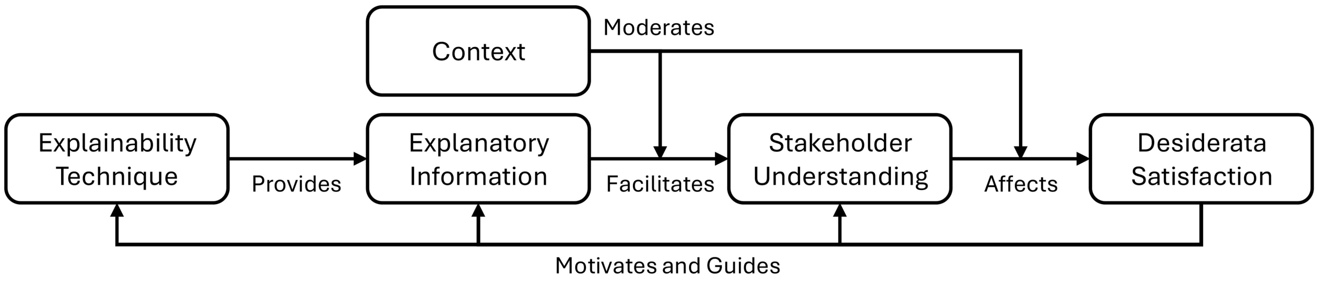

FIGURE 3. From XAI techniques to stakeholder desiderata satisfaction, a concept model adapted from [34].

AI system. Different stakeholders have different needs and expectations for the explanations of the same AI system. As a result, to provide useful and understandable explanations, the XAI approaches should adapt to the expertise and the social and cultural background of stakeholders since their cognition and perception are socio-culturally conditioned [35], [36], [37], [38].

The literature has proposed different ways of categorizing stakeholders, [39] identifies four main stakeholder communities with different motivations and requirements for XAI: *developers* who build AI applications, primarily concerned with quality assurance, system testing, and debugging, *theorists* who advance AI theory, with a focus on understanding fundamental properties and improving the state of the art, *ethicists* who are concerned with fairness, accountability, and transparency in AI systems such as policymakers, commentators, and critics, and *end users* who require explanations to understand and trust AI outputs, and to make decisions based on those outputs. Reference [34] identifies five stakeholder classes: *non-expert users* who interact with AI systems, *developers* who design, program, and build AI systems, *affected parties* who are impacted by AI decisions, often without their direct interaction, *deployers* who decide where and how to implement AI systems and *regulators* who are responsible for creating legal and ethical frameworks for AI usage.

From the legal or standardization literature, according to the Ethics Guidelines for Trustworthy AI [2], the stakeholders include *developers* who research, design or create AI systems, *deployers* who are public or private organizations that integrate AI systems into their business operations and use them to offer products or services, *end-users* who interact with the AI system, either directly or indirectly and *the broader society* including all others who are directly or indirectly impacted by AI systems. Based on this guideline, the EU AI Act [40] identifies additional stakeholder categories for prescribing legal obligations such as *importers* who bring AI systems from third countries into the EU market, *authorised representatives* who act on behalf of providers from third countries within the EU, *national competent authorities* who oversee the implementation and enforcement of the AI Act within member states and the *AI Office* which facilitate the development of codes of conduct, provide guidance, and ensure proper application of the AI Act.

To understand how XAI approaches can be designed and evaluated to satisfy such diverse stakeholder needs and expectations, [34] present a conceptual framework as shown in Fig. 3. Specifically, **explainability techniques** are applied to AI systems to generate explanatory information tailored to the needs of different stakeholders. The format, completeness, accuracy and currency [41] of **explanatory information** influence how well stakeholders understand the AI system. The degree of **understanding** achieved by stakeholders determines whether their specific needs and expectations (**desiderata**) are met. The **context** affects every stage, altering how explanatory information is interpreted and how understanding translates into satisfaction. This conceptual framework highlights the requirement of being stakeholder-centered for the XAI approaches.

B. STAKEHOLDERS OF AI IN 6G

AI in 6G concerns more than just network providers but also end users and legal authorities. Based on the literature [5], [40], [42], [43], we summarize the stakeholder specifications of XAI in Table 1 with three main stakeholder groups.

1) LEGAL AUTHORITIES

Legal authorities include legal regulators who need XAI to create, implement and enforce AI laws and regulations [44], and legal auditors (a third-party "notified body" or internal "authorised representative") who need XAI to audit the AI systems for legal compliance [40] (e.g., "Forensic Analyst" in Fig. 1). For example, when drafting AI laws and regulations, legal regulators need to specify what constitutes an adequate explanation or transparency for AI decisions. Understanding what XAI techniques could offer helps them set realistic and enforceable standards that AI service providers must meet. As a result, the legal auditors must examine the explanatory information by XAI techniques against those legal principles to report compliance and give recommendations. When AI-related mistakes happen, legal authorities need XAI to identify who is accountable for the erroneous decisions.

While the legal frameworks for 6G-specific XAI are still in a nascent stage, legal authorities can consider the existing general regulations to derive legal principles on agency, privacy, security, and safety for XAI in 6G. Examples may include AI-specific regulations such as the EU AI Act [40]

TABLE 1. Stakeholder-centered XAI specifications summarized based on literature (Who: relationship with AI, Why: the reason for XAI, What: explanatory information, How: XAI approaches to generate explanatory information).

	Legal Authorities	End Users	Service Providers
Who	set and enforce AI laws and regulations; audit AI for legal compliance.	directly or indirectly use AI	develop, deploy and maintain AI
Why	prescribe responsibility, audit conformity and enforce accountability	trust	**everything left plus** optimization, debugging, and maintenance etc.
What	written policies, procedures and instructions such as regulatory compliance strategies, design and development procedures, data management systems, risk management, post-market monitoring, incident reporting, communication protocols with authorities, and record-keeping	feature-based (feature attribute, shape and interaction), example-based (similar, typical and counterfactual examples), rule-based (decision rules and trees), supplementary (input, output, dataset, performance metrics)	**everything left plus** saliency maps, feature importance scores, counterfactual explanations, attention visualizations, surrogate models, etc.
How	quality management system, technical documentation	personalized interactive user interface	**everything left plus** (scope) global and local approaches, (stage) ante-hoc and post-hoc approaches, (example) counterfactual, influential instances, adversarial and prototype-criticism

and General Data Protection Regulation [45], telecommunications regulations such as ITU standards and FCC guidelines, ethical guidelines and soft law such as OECD AI Principles [46], IEEE Ethically Aligned Design [47], sector-specific regulations such as financial and healthcare regulations regarding the use of AI, national AI strategies, and international collaborations and frameworks such as Global Partnership on AI [48].

The requested information by legal auditors to understand AI in 6G can be in the form of a well-documented quality management system containing written policies, procedures and instructions such as regulatory compliance strategies, design and development procedures, data management systems, risk management, post-market monitoring, incident reporting, communication protocols with authorities, and record-keeping (Article 17, EU AI Act [40]).

2) END USERS

End users are the consumers of AI-powered services within the 6G network and require understandable explanations to build trust in AI decisions that directly affect them (e.g., "Consumer" and "Business User" in Fig. 1). Gaining a user's trust relies on three pillars [49], i.e., *ability*, *benevolence*, and *integrity*, in addition to the user's *propensity to trust*. *Ability* refers to the functional performance of AI systems, which can be reflected by explanatory information such as accuracy and confidence in AI predictions. *Benevolence* means the extent to which the AI systems are seen as genuinely concerning the end users' welfare, which can be influenced by explanatory information such as decision trees in safety, security or privacy-concerned scenarios [50], [51], [52]. And lastly, *integrity* refers to the user's belief that AI systems consistently follow a set of principles or values, which is closely related to legal conformity.

Considering the diverse backgrounds and preferences of end-users, [42] identified 12 end-user-friendly explanatory forms in four categories: *Feature-based explanations* (e.g., feature attribution, feature shape, feature interaction), *Example-based explanations* (e.g., similar examples, typical examples, counterfactual examples), *Rule-based explanations* (e.g., decision rules, decision trees) and *Supplementary information* (e.g., input, output, dataset information, performance metrics). To facilitate this, interactive user interfaces can be developed to enable users to explore and customize the variety of explanatory information forms at different depths [53], [54].

3) SERVICE PROVIDERS

Service providers are those who deploy, operate, and maintain AI services in 6G and need explainability to understand, predict, control, debug and improve the AI-enabled components of their systems (e.g., "Network Engineer" in Fig. 1). For example, network service providers may 1) work on the data mining and intelligent control layers using XAI to diagnose causes of incorrect decisions by AI models [55], 2) use XAI to enhance the performance of the network, manage operational risks, and understand the relationships between the input data or the training parameters and the learning efficiency, as well as signal the presence of biases, and 3) use XAI to better understand network maintenance and monitoring [5]. In addition, service providers also need explainability to facilitate end-user trust and legal compliance.

As a result, service providers are responsible for not only generating explanatory information for their own (e.g., saliency maps, feature importance scores, counterfactual explanations, attention visualizations, and surrogate models [43]) but also presenting that information in ways understandable and legally compliant for end users and legal auditors. To generate low-level explanatory information (e.g., mainly visual, numerical or statistical features that require the expertise of the AI systems to understand) for improving the AI services, service providers may directly adopt the existing XAI approaches for general AI [43] or adapt them in the 6G application context. To generate high-level explanatory information (mainly natural language that requires minimal expertise of the AI systems), service providers may adopt legally compliant quality management systems and comprehensive documentation practices for legal auditors and design personalized interactive user interfaces for end users.

4) DISCUSSION

Table 1 presents significant responsibilities for the 6G service providers, including but not limited to, 1) communicating the XAI capabilities to others (legal authorities and end users), 2) collecting XAI needs or expectations of others, 3) generating the appropriate explanatory information for others. Moreover, the needs of XAI vary across network layers (i.e., edge, RAN, core, and cloud):

- At the edge, where resource-constrained AI models operate on end-user devices and base stations, explainability is crucial for end users to understand local AI-driven decisions (e.g., personalized AI services). Legal auditors may require localized compliance checks on AI privacy and security.
- Within the RAN, explainability is essential for service providers to diagnose performance bottlenecks and optimize AI-based network orchestration. Legal regulators may also need explanations for spectrum allocation fairness, ensuring non-discriminatory AI-driven network policies.
- In the core network, where AI-driven traffic management, authentication, and policy enforcement occur, interpretability needs are broader. Service providers may require detailed insights into AI decisions affecting routing, congestion control, and security monitoring, relying on surrogate models, counterfactual explanations, and dependency graphs. Legal auditors may focus on AI accountability in cybersecurity, necessitating forensic-level XAI capabilities to trace automated decisions affecting network access control, encryption policies, and service differentiation.
- At the cloud layer, where AI models are trained and orchestrated at scale, explainability is vital for both service providers (e.g., debugging global AI models, ensuring fairness, and improving network automation strategies) and legal authorities (e.g., ensuring compliance with regulatory frameworks). Explainability at this level typically involves comprehensive documentation, model interpretability reports, and traceability mechanisms.

As such, here we identify a lack of a common and formal specification language on explainability to capture the actual requirements across network layers and stakeholder communities for the verification of XAI techniques.

IV. CHALLENGE 2: EXPLAINING REINFORCEMENT LEARNING IN 6G

Various approaches and taxonomies are proposed in the literature discussing the explainability of RL as reviewed in [8], [24], [56], [57], [58]. A more RL-oriented review study is given by [24] where interpretability/explainability concerning different levels of RL is discussed: (1) feature importance in taking action for a given input state, (2) influential past experiences of MDP and reward (objective) components affecting the learning process for current behaviour, and (3) long-term policy over time capturing an abstraction or summarisation of the agents' behaviour about subtasks and planning during training to achieve the goal. A comprehensive study is provided by [8] on exploring/exploiting interpretable models for different components of RL including (1) inputs used by agents for learning and decision-making, (2) transition model of MDP, (3) preference models of reward function, and (4) value function and policy directly or indirectly. An overview of the state-of-the-art models with their interpretability purpose is listed in Table 2.

A. DIRECT FUNCTIONAL EXPLANATIONS

Direct model explanations are challenging for deep neural networks (DNN) due to the large functional space they operate in. Attempts to use generalised functional descriptors to collapse the dimensionality of the DNN into more tractable expressions suffer from several issues:

1) Mapping to a discontinuous functional space
2) Loss of information
3) Human bias in interpretability
4) Lack of prior target function space resulting in large function search space

For example, one might be motivated to use general functional approximators (e.g., Gaussian Processes [59], Hyper-geometric functions such as Meijer-G, Fox-H [60]) via Kolmogorov–Arnold Superposition Theorem (KST) to model the activation functions as they can (with the right parameters) be approximated to most functions. The combination of these approximators is a form that is the same as itself, reducing the function complexity of combining many activation functions. What we are effectively doing is asking for a DNN to be represented in some kind of parametric model space, where changing the parameters equates to searching on a discontinuous function landscape - see Fig. 4A. We are bound to lose information in the projection process. There may be many functional representations that satisfy some kind of information loss function, and hence we are likely to use cognitive bias to select functions that are

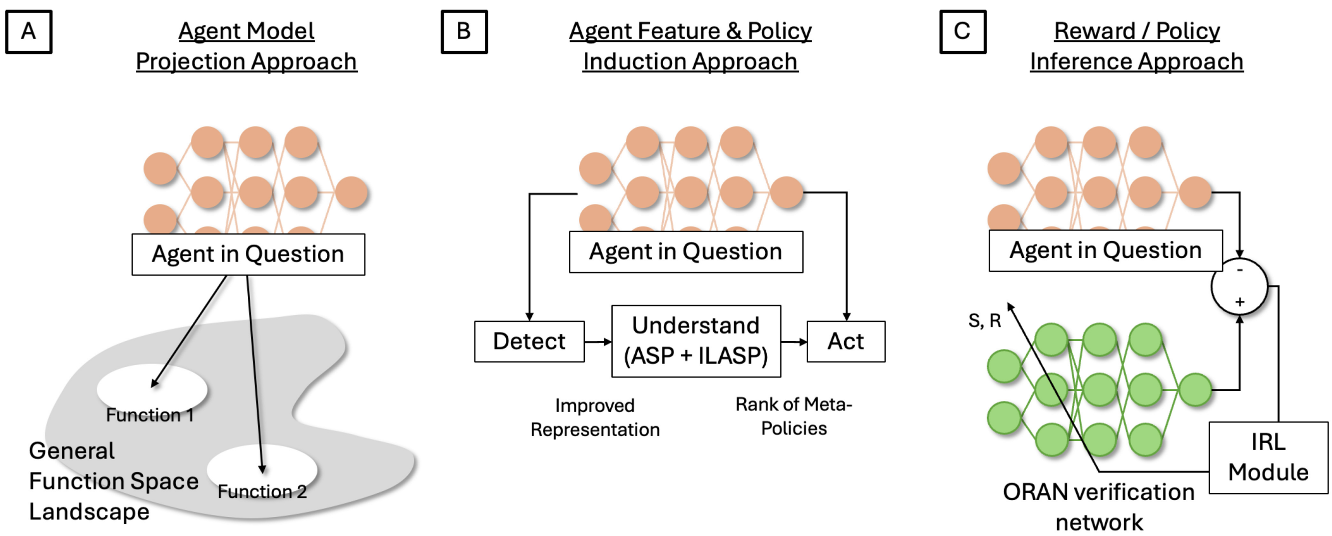

FIGURE 4. **Forms of XRL: (a) agent model projection to function space, (b) capture agent's behaviour to meet the objectives and sub-goals/subtasks with temporal abstraction, and (c) reward or policy inference approaches.**

more familiar to us or are easier to analyse/interpret. Finally, unless we know what the target function will look like (e.g., we know from prior convex optimisation solutions), the search space can be extremely large.

A further critical challenge is how such direct modelling approaches would work for more complex neural networks such as convolution, recurrent, transformer, and RL architectures (e.g., actor-critic, double duelling). It is worth noting there are more complementary data-driven models where a projection is made to a self-organising-map (SOM) with temporal-difference (TD) error used to control for the process [61], but the SOM is used more for improved performance than explainability. Nonetheless, these methods in general are limited to relatively small neural networks with prior explainable model beliefs.

B. SYMBOLIC EXPLANATIONS

Traditional RL methods are inherently limited by the assumption that tasks follow a Markov process, where the future state depends only on the current state and action, not on the sequence of events that preceded it. This assumption renders traditional RL methods ineffective for non-Markovian tasks, as they cannot naturally incorporate the necessary historical context without making policy learning practically intractable due to the exponential growth in state-space complexity. Reward Machines (RMs) present a compelling approach to address these challenges by explicitly modelling the structure of tasks. By breaking down tasks into a series of sub-goals and transitions, RMs facilitate the learning of policies that are more interpretable but also better aligned with human-understandable objectives. This task decomposition allows for the accommodation of non-Markovian tasks by integrating memory into the learning process. We summarise below the key aspects of this neuro-symbolic approach to RL.

1) REWARD MACHINE-BASED RL

Reward machines (RM) are finite-state *automata* enabling temporal high-level abstractions that represent the reward function and the structure of RL tasks in a compact form in terms of a finite number of states and transitions between them that are independent of the RL state and action spaces, rendering them more applicable towards autonomy of ultra-large networks. In the context of RL, automata have been generally used to represent hierarchies of decisions made by agents [62], [63], memory of the trajectories in partially observable environments [64], reward functions and the structure (RM) of an agent's task [65], [66], and encoded policies by a neural network [67]. This will induce more interpretability for the policy taken and the level of learning task performed by the agent. More specifically, RMs are exploited toward encoding sub-goals of a goal-oriented episodic RL task by two types of abstraction [68], [69]. The first one makes an abstraction of the original state space into the automaton set of states by recognising the level of task completion in terms of the initial launch (initial state), sub-goals achieved (accept state), those yet to be achieved (incomplete state) and not achieved by the end of the episode (reject state). The second one abstracts the actions into the set of sub-goals labelling the transition edges as local objectives toward the subsequent automaton state. Once the agent observes the next visible state of the environment, the automaton state transition occurs when the observables (if any) contributing to the sub-goals of the task are perceived by the agent. The abstractions made by this automata enable simpler learning of the whole task by learning easier subtasks, better exploration by quick transition to abstract states, generalisation to different similar or related tasks by sharing common subtasks, and deal with *partial observability* of environment and *non-Markovian* rewards by providing an external memory for the agent.

TABLE 2. Overview of the XRL methodologies.

XRL Purpose	XRL Method Ref	Impact on RL performance	Disadvantages	Relevant Stakeholder Group
Action	• Feature importance e.g. SHAPLY • DT using imitation learning e.g. VIPER, DAGGER • Augmented MDP	• speed up learning and better generalisation • simulatable, highly portable and deployable	• Unable to capture the memory and summarisation of the behaviour and subgoals or planning • Not easily robust to changes in the environment or task • Decision trees are not suitable for applications with approximate reasoning	• Service providers • End users
Reward/Value function	RDF, VFF, DVN, QMIX, QTRAN, RM, LTL, Boolean algebra, DT	• Convergence guarantees the correct decomposed values • Simulatable and achieve average reward comparable to the non-interpretable models	• Requires combination with influential experience (transition tuple) identifying approaches for a more insightful explanation of reward decompositions • Requires domain information about reward components prior to learning and is not suitable for post-hoc explainability	Service providers
Policy (long-term behaviour)	ISA, TLdR, Abstract policy graph	• Simulatable and decomposable • More intelligibility through temporal and hierarchical abstractions • Better exploration by a quicker transition between abstract states • Enables summarisation of policies to learn the shortest path toward the goal	• The policy summarisation may not be helpful for distribution shifts • learning the temporal abstraction structure is not scalable to the number of abstract states and their transition rules	• Legal authorities • End users
Transition model and MDP	DT, FOL, GP, VAE, object-oriented, Graph-based (state-space, attribute)	• larger generalisability and better transfer learning • More effective exploration and data-efficient RL • Rendering potential decomposability of the environment structure	• Require further verification by a human to identify mistaken influential experiences • learning the probabilistic, causal or influence model of the transition model is computationally expensive • Graphs are unable to capture continuous state spaces and stochasticity of transition model accurately	Service providers

To learn the RM described above, a *sub-goal automata learning task* is first formed by considering the automata model and a set of observation traces extracted from the traces of visible states and actions. Second, the automata learning task is represented by Answer Set Programs (ASP) [70] which is a declarative logical programming language to express knowledge and reasoning in the representational form of *logical programs* inherited from computational logic. ASP consists of a group of statements characterising the objects of a domain and their relations, and the set of possible outcomes of ASP, called answer sets, are the semantics of the program and reflect the agent's beliefs related to this program. Accordingly, various tasks can be reduced to finding the corresponding ASP program and their solutions as the answer sets of the respective ASP programs. ASP solutions can be learnt by the Inductive Logic Programming (ILP) system [71] and are so referred to as Inductive learning of ASP (ILASP) [72], [73]. ILP enables model developers to develop more verifiable machine learning algorithms by practising logical programs as a universal representation that are much easier to manipulate by pure clausal logic changes. The agent in question in Fig. 4B applies an ILASP system [72], [73] to perform learning by forming three components: (i) *background knowledge*, (ii) *hypotheses set*, and (iii) *examples* from the ASP representation of the automata model and observation traces, respectively. Third, the inductive solution to the ILASP learning task is released in the form of hypotheses or rules that comply with the background knowledge and cover the context-dependent examples. This solution is the minimal automaton, i.e., with a minimum number of states and a limited number of edges between states, capturing the sub-goal structure of tasks. The sub-goal automata learned in this way are referred to as Inductive Sub-goal Automata (ISA) and can be further exploited by RL algorithms to learn the policy conducting the agent to achieve the goal [69] using Q-learning. This can be performed in either a hierarchical manner (HRL) or a direct RL manner. The hierarchical way is carried out by first learning the policy over the set of available *options* in an automaton state, followed by learning the policy of an option (for an outgoing edge) that satisfies a sub-goal of the task. However, in the direct way, a single option policy is learned in a given automaton state that might not satisfy a sub-goal of the task but ultimately reach the goal of the task, and hence globally sounds to yield the best and fastest policy to achieve the goal.

In addition to exploiting RM for RL, it can also be *interleaved* with RL for iterative refinement of RM from the experience of an RL agent. This can be done by first checking if the current visible, terminal, and goal states of the environment observed by the agent are correctly validated by the current state of the automaton, and if not, then add it to the respective set of observation traces (goal, dead end, and incomplete) for subsequent relearn of the RM. If the accuracy of the relearned RM is still unsatisfactory, then another state is iteratively added to the automata set of states to enlarge the hypotheses space that captures the rule of automata. It should be mentioned that this way of learning RM is not scalable to the number of automata states and edges. The RM approach is also extended to the case where perceptions of high-level propositional events from the environment are noisy (probabilistic) [74]. Another extension is to the multi-agent scenario where RMs of subtasks are individually learned in a decentralised way to cooperatively guide the policy of agents toward a common goal [75].

Application in 6G Exploiting RMs toward symbolic RL for next-generation 6G applications is a promising way to address the high computational burden imposed by very large state spaces when performing joint optimisation over different layers of communication network as well as joint communication, control, computing, sensing and localisation (3CSL) required for broader connectivity over space-air-ground networking [76]. In this context, RMs can be leveraged within each layer toward layer-wise interpretability of the policies used to meet the respective goals of each layer and also in higher layers to hierarchically explain the orchestration of the functionalities performed interactively among layers. This can also be extended to capture the structure of the sub-goals when addressing RL-based optimisation for 3CSL. Additionally, RMs enable an interpretable framework to deal with the partial observability of the communication environment by capturing the memory of visible states of the network, as well as the structure of tasks while taking the reward function and policies into account for sequential decision-making. Further discussions on the way RM can be exploited toward this end for the *network slicing* use case are provided in Section VI-A.

2) NEUROSYMBOLIC APPROACH

Neurosymbolic (NeSy) AI tries to bridge the gap between the low-level connectionist approach of the "Neuro" component for statistical learning from raw data and the high-level cognitive and human-like approach of the "Symbolic" component for reasoning [77], [78], [79]. The integration of these two different areas of AI enables more efficient learning of data structure and knowledge representation from data as well as reasoning and explaining the learned experience in a form that is *understandable* and *interpretable* by humans. The various approaches with which these two components interact with or incorporate each other give rise to different categories of NeSy AI, as described in [80]. One categorisation proposed by [81], studies this integration based on learning for reasoning [82], [83], reasoning for learning [84], [85], or joint learning and reasoning. Further details and discussions on the current trends of such integration and their respective challenges and future opportunities can be found in [80], [86].

In *Learning for reasoning*, the Neuro component guides the reasoning process by approximating the symbolic computation [87], assigning probability distribution on the base knowledge [88], learning relational reasoning with *Relational networks* (RNs) [89], learning first-order logical rules in the form of weighted non-linear logic operators called logical neural networks (LNN) [90], learning symbolic representation of unstructured data, structured knowledge graphs to further improve the reasoning process by shrinking the search space of the symbolic system. However, in *reasoning for learning*, the symbolic component guides the Neuro part in various settings by catering high-level symbolic knowledge to a neural network that is involved in learning and decision-making with the ultimate goal of enhancing interpretability and reasoning, especially when dealing with mislabelled, noisy, ambiguous data or complex applications of dynamic environments [91], [92], [93] or facilitating efficient learning for the Neuro component [94]. As this requires manual engineering of the symbolic knowledge for the downstream task, known as *symbol grounding* problem, there is a need for a *joint reasoning and learning* where a combination of neural network training and inference of symbolic knowledge from data is performed by intermittent interaction between the learning and reasoning [85], [95], [96].

In the context of RL, [97] provides a comprehensive survey of such categorisation for NeSy integration in different aspects of RL algorithms, including environment state space, agent policy, and reward function. The NeSy RL has been proposed to address the challenges of deep RL (DRL) related to data-inefficient learning, poor generalisation to similar tasks, lack of high-level processes for various types of reasoning, and low human-comprehensibility for the sequence of actions taken by the agent [82]. Accordingly, deep symbolic RL (DSRL) was introduced by [82] in an end-to-end architecture to learn low-dimensional high-level symbolic state representations from the high-dimensional raw data in the back-end for subsequent front-end symbolic reasoning. DSRL is further extended by [83] to take into account common sense priors for the assignment of rewards and the aggregation of Q values. This is shown to achieve faster learning and also higher accuracy than just Q-learning and DSRL especially when trained on a simple environment and tested on more complicated environments. This extension can also offer a better balance between generalisation and specialisation. Reference [98] makes use of LNNs to train an RL policy that can directly render interpretability by neural training of logical functions. In addition to the representation perspective, NeSy RL is applied for safe exploration of the state and action space to allow efficient verification of the learned policies [99].

Reference [100] proposes a relational deep RL framework that leverages the relational reasoning of RNs and key-value attention mechanisms to further build and aid autonomous

RL agents that can learn out-of-distribution tasks expressed by temporal logic instructions. A neurosymbolic relational RL approach, called deep explainable relational RL (DERRL), is proposed by [101] where deep neural networks learn symbolic relational representations (in the form of logic programs) of policies to extract interpretable policies while enabling scalability under structural changes in the environment. A similar approach is used by [102] for a hierarchical RL (HRL) where high-level symbolic relational representations (in the form of ASP) of meta-policies over options are learned by an ILASP system and further used to guide a pre-trained DRL agent. Reference [22] also considers learning a task-level relational reasoning module for HRL. This can be considered as the hierarchical extension of [20] where neural logic machines (NLM) [103] are exploited to reason about policies in DRL by combining differentiable ILP (DILP) and policy gradient methods. Another direction of NeSy RL research is to extract and exploit a finite state machine (automata) structure for reward shaping, especially for sparse and non-Markovian reward functions, toward a more effective DRL algorithm by changing the reward with respect to different learning stages over time [104], [105], [106]. More details about NeSy RL can be found in surveys [8], [86], [97].

Application in 6G NeSy AI has been exploited in the context of 6G for the following use cases:

1) *zero-touch network and service management* (ZSM) to capture the dynamics of wireless Internet of Everything (IoE) environment for autonomous management of communication and computational service decisions using a directed acyclic graph (DAG)-based Bayesian networks as an explicit explainer for the neural network-based multivariate regression [107],
2) intent-based semantic communication to consider *semantic* and *effectiveness* of transmitted messages, without affecting their reliability, for integration of time-sensitive autonomous systems in future generation of communication networks. This will enable developing intelligent end-nodes that can efficiently and reliably communicate through a combination of knowledge representation and reasoning with machine learning. A Generative Flow network approach with DNN encoder and decoder structure is used by [108] to learn the probabilistic structure of the observed data at the receiver emanating from an optimal transmit message in form of a compact objective function.
3) provisioning-aware radio resource allocation by a gNB to meet QoS requirements [109]. More specifically, a *Bayesian Graph Neural Network* explainability approach is used to address an RL-based minimisation of the physical radio resource over-provisioning and under-provisioning while meeting the amount of requested downlink traffic.

Explaining Reinforcement Learning is Challenging: RL prospects into the future and explaining the reasoning is often done through symbolic mapping to known symbolic knowledge or belief heatmaps of associated observation-action pairs.

C. REWARD/VALUE, STATE TRANSITION, AND POLICY SPACE VERIFICATION & INFERENCE

In many legal and regulatory cases where motivation, opportunity and capability to act are sufficient to prescribe responsibility. As such for RL, we may only be interested in limited explanations of what its motivations (rewards) are, and/or what state-action mappings (policies) it is set to do. As a primary party (e.g., network service provider), we would have direct access to the reward or value function and policy, and we can in real-time display these as a function of time and events to assess performance.

For the reward/value function, the explanations can be provided by decomposing the reward function, also resulting in the decomposition of the value function, into several components that can provide intuitions about the contribution of various objectives to the final composite reward [110]. A similar idea was also considered for the decomposition of the value function in the context of cooperative multi-agent RL to account for the individual contributions of each agent in the joint value function, assuming it can be approximated by factorization of less complex functions, referred to as Value function factorisation (VFF) [111]. Various algorithms based on VFF include the value decomposition network (VDN) [111], the Mixture of Q-values (QMIX) [112], and the Transformation of Q-values (QTRAN) [113]. The decomposed variants of these algorithms, referred to as D-VDN, D-QMIX, and D-QTRAN, are further proposed in [114] by incorporating RDF into the VFF to capture the contribution of distinct reward components in the approximated individual value function of each agent. This adaptation was additionally improved in the complexity of the reward components per agent by a multi-headed architecture that performs multi-task learning of all reward components [114]. A use case of these algorithms was also studied to improve the key performance indicators (KPIs) of codec adaptation in XR traffic.

This approach avoids trying to explain what the agent is doing precisely but rather explains what motivates (rewards) the agent to behave in a certain way (policy). In many cases, we may wish to understand state transitions to better understand the agent-environment interactions rather than just the agent's internal design (e.g., reward and policy). This extrinsic explainability is essential to check the agent is not only designed well but also interacts well. This has been designed as explainable Q-learning for linear controller systems so far and can be expanded to data-driven problems [115]. Expanding some of these techniques to high dimensional states for 6G could be a challenge at the resource optimization level.

What becomes more challenging is if we are a secondary party (e.g., ORAN orchestrator) and we must verify what the policy and reward space is for a rApp/xApp. For example, an innovator inserts a suspicious ORAN app and we must verify whether the declared certificate is true. Here, we can use a variety of inference or inverse learning methods to verify. For example, we may set up a verification RL model that uses the declared policy or reward to check the error between them [116] - see Fig. 4C. If the agent in question declares what policy and/or reward functions it uses, the ORAN would verify the declared model in question is performing in accordance with expectations by minimizing the error feedback. In an undeclared case, or when the above is not true, the ORAN would use general functions (e.g., polynomials) to try to infer what potential functions are being used via an inverse reinforcement (IRL) policy or reward learner module. Either way, the difference in output can be feedback to the verification network as state (S) and rewards (R).

Another approach in XRL is through *feature importance* where the most influential *state features* affecting the action taken by the RL agent at each time instant are recognised. This explanation can provide insight to service providers and end users about the critical input contexts for making decisions. A use case for this type of XRL is in radio resource management (RRM) for V2X communication in the context of autonomous vehicular networks [117] toward automotive transportation. In this setting, vehicles communicate with the help of a roadside unit (RSU) and optimise real-time power control using a multi-agent DRL. The space for each agent comprises the transmit power in the previous time slots, direct channel gain to other vehicles, interference gains of the current and previous time slots, and the signal to interference in the previous time slot, which constitute a state space of overall $2K + 2$ dimension for a total number of K vehicles. The explainability approach in this study is based on SHAP (SHapley Additive exPlanations) values for computing and ranking the feature importance scores to reduce the size of state space and retraining the model using the most significant state features that globally contribute to the agent's decision, and hence explaining the allocated power at each time slot. The experimental results in [117] show that under the high mobility regime – which results in low correlation between fading coefficients of successive time slots– features of previous time slots have negligible influence in the agent's output and hence reducing the size of state space by almost 70%. However, this might not always be the case, and hence the characteristics of the communication environment might allow for more important state features to appear as the number of vehicles increases, which renders the feature importance method computationally infeasible.

Most of the discussions provided on the explainability techniques for RL are model-based such as decision-tree (DT), reward machine, structured causal models, graph-based abstractions for the policy and transition model of MDP, and reward decomposition. It should be noted that these models still need to be learned with interpretable methods in most cases to capture the trajectory of experiences in an RL task. The most famous model-free techniques which are not specific to (but can be also used in) RL are the feature importance-based techniques such as SHAPLY that computes the highly influential state feature for a given action taken at each time step. Other model-free techniques include post-hoc explainability techniques such as linear interpretable model-agnostic explanations (LIME), saliency maps, and conversion of policy into interpretable formats. The model-free approaches that can be considered more specific to RL are the "interpretable off-policy evaluation" and "influence functions" to identify the most influential experiences of an agent with respect to the estimates of value function [24].

V. CHALLENGE 3: CAUSAL UNDERSTANDING FOR IMPROVING EXPLAINABILITY AND PERFORMANCE

The explainability approaches discussed in the previous sections can be inadequate as they only expose variable correlations and thus misinterpret confounding factors and complex causal chains present in those dynamic environments. Causal analysis shows a promising pathway addressing this. Causal analysis can not only help to optimise the AI models but also produce a deeper understanding of both the AI models and the dynamics of the networks.

A. CAUSAL ANALYSIS

Causal analysis is a statistical framework which allows system analysis from three different perspectives: observational, counterfactual, and interventional [118], [119], [120], [121]. Traditional probabilistic approaches are mostly focused on the first observational level; observations of the system variables are used to infer a probabilistic model of the system, providing some explainability of it. The counterfactual and interventional levels opened by causal analysis allow the understanding of how such probabilistic models would be modified under hypothetical changes in the variables. The counterfactual level analyses how a past sample observation would have changed if some system variables had a different value, while the interventional level considers how the model would change if some modifications were forced on it, such as fixing a variable to a certain value. In that sense, causal analysis does not produce a single probabilistic model but a continuous stream of probabilistic models (the observed one as well as any possible variation of it). Common approaches to hypothetical changes such as sensitivity analyses offer a correlational understanding of the changes in variables but are not able to distinguish confounding factors and since they do not provide complete probabilistic models they are not able to produce analysis on specific past samples; i.e., they answer how in average

TABLE 3. Examples of causal representations for 6G networks.

Focus of the Causal Analysis	Nodes of the Causal Diagram	Edges of the Causal Diagram
Network	Network variables (e.g. downlink traffic, user numbers, QoE metrics)	Network causality (e.g. performance metrics vs. quality of service causality)
Network from AI model perspective	Latent variables, observations, interventions	Causality between network and AI model
AI model	RL states, actions, observations	Intra-model causality

some variables seem to affect others, while causal analysis can answer how, for a particular observation, changes in variables would have affected others. In other words, a sensitivity analysis offers a static probabilistic model, while a causal analysis offers probabilistic models for each past or hypothetical scenario.

Causal analysis uses structural causal models as a main tool to represent such scenarios. A structural causal model is defined by a series of structural equations which define how the different variables of the system relate to each other and any additional sources of variation such as exogenous noise. The structural causal model induces a graphical representation which is visualised using Directed Acyclic Graphs (DAGs) [118], [119], [120], [121]. These diagrams offer, among other benefits, a clear visualisation of the analysed system's causal structure, showing how the variables relate to each other. They also make it easy to understand for example which variables would need to be intervened to produce changes in specific variables of interest.

Table 3 shows examples of different causal representations that can be used when modelling the causality of a 6G network, the nodes of a graph may represent for example performance and quality indicators of the 6G network, allowing the understanding of how the performance indicators cause changes in the quality (see [122, Fig. 6]). Another possibility is that the causal model represents the network from the AI model perspective. In this case, the graph nodes may categorise the information into latent variables, observations, and interventions (see [123, Fig. 2]). A different possibility is to perform a causal analysis focused directly on the AI model. In this case, for RL the graph nodes may represent the states, actions, and observations (see [124, Fig. 1]).

Causal discovery techniques [125], [126], [127], [128], [129] produce such diagrammatic causal representation from observational and interventional data. Estimation techniques allow the quantification of the effect of the causal relationships to obtain the complete structural causal model, which enables the calculation of specific probabilities using causal inference methods [130], [131].

Recently, causal analysis is increasingly used in connection to machine learning, to improve the explainability of such techniques but also to make them more efficient. The interplay between the two approaches is what is called causal machine learning [132], [133], [134], [135].

Traditional supervised machine learning works on the observational or associational level, using the data to learn from its correlations to solve the supervised tasks. Regarding its efficiency, this leads to strong limitations such as the difficulty for such techniques to distinguish between causal and confounding information [133], [136], [137], [138], [139]. To overcome this problem, different causal machine learning approaches can be applied, going from a causal feature selection that preselects the causally relevant information for the machine learning models to work on to causal architectures of machine learning models which already have as a learning objective the distinction between causal and non-causal information [140], [141], [142], [143].

On the other side, beyond the efficiency argument, causal machine learning models are more explainable. The improvements may come, for example, from using more explainable features or feature selection [144], or from using architectures and representations which are explainable per se [145], [146].

B. CAUSAL ANALYSIS FOR 6G OPTIMISATION AND EXPLAINABILITY

As was presented in the previous section, causal analysis allows the creation of models of complex 6G networks that help understand how their variables relate to each other causally. Next, we review the key causal analysis approaches, then detail some use cases and their impact on stakeholder groups, and finalise by reviewing recent work in causal AI in communication networks.

1) NATIVE AI NETWORK OPTIMISATION

Prior or real-time causal understanding can be used to simplify the use of native AI operating in 6G networks and make it explainable by construction. There are many examples which we give insight into below.

Transfer Learning in Parallel Channel/Traffic: In channel estimation and traffic prediction amongst nearby Base Stations or Radio Units (RUs), transfer learning can be applied between edge prediction models such as RNNs and GPs [147]. While traditionally it would be required to train multiple parallel models naively assuming the independence of the data sources, causal understanding between channels or coverage areas allows for reducing their multiplicity. For example, if the data from a BS/RU is shown to be the cause of the data for others (e.g., due to users moving from one spatial location to another), causal directional

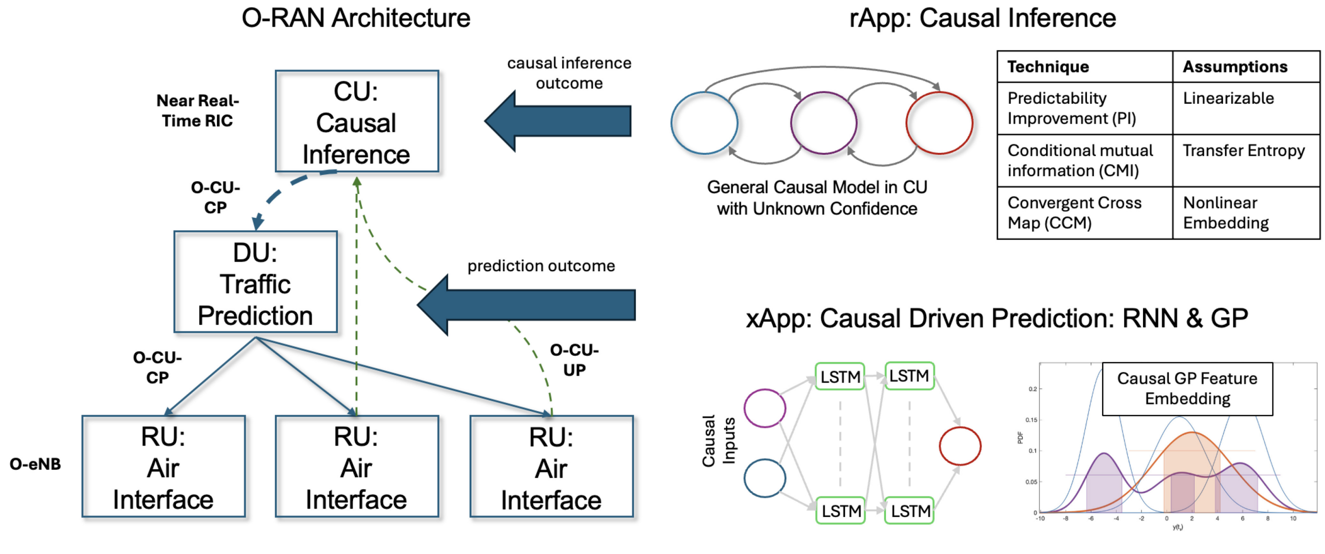

FIGURE 5. Causal learning rApp drives LSTM and probabilistic GP prediction tasks as xApps.

transfer learning could be applied between such sources of data instead of training multiple independent models or guessing which direction the causality lies using correlation approaches. In Fig. 5, we can see that future ORAN xApps may hold these parallel prediction models in Distributed Units (DUs), and the causal inference as a non-real-time rApp may be held in the Centralised Unit (CU) to indicate which direction the causal transfer learning should take place.

Global Cross-Layer Optimisation: The newly integrated services (that are previously siloed) such as integrated sensing and communications [148], [149], [150], shared native AI-as-a-Service [151], using mobile drone relays to expand and improve coverage [152], [153], [154], etc. require cross-layer and cross-module global optimisation, leading to a dimensionality explosion challenge, especially for RL. Causal AI can improve efficiency by reducing the search space in the training of network-wide cross-layer global models. This can reduce the dimensionality explosion problem in 6G [155]; if some of these data sources are proven to be causally redundant or even confounding information, they can be excluded from the training processes.

> **Causal Learning is Crucial**: causal learning allows us to understand why something has happened—leading to more reasoned and efficient downstream AI/ML implementation.

Causally Informed AI/ML Models: Causality may also be used in the design and architectures of the AI models themselves to improve their efficiency and explainability, as it was mentioned in the previous section. Some examples in other domains are the following: [145] propose an autoencoder with an internal causal layer that disentangles the latent variables. While traditional autoencoders produce latent variables that may have no semantic or causal meaning (since they use a commonly incorrect assumption of uncorrelated latent variables), here the latent variables are the causal variables of the data (which are learnt together with the causal model of it). Reference [143] propose a model with a causal training objective whose aim is to distinguish in the inputs the causally relevant information to produce the correct outputs. Similarly is done in the work by [141] in an RL case. Reference [140] propose models with specific causal architectures and objectives.

2) CAUSAL AI USE CASES AND STAKEHOLDERS IN 6G AND XAI TECHNIQUES COMPARISON

In the following, we motivate the use of causal approaches in 6G networks by presenting specific use cases and comparing them to other explainability methods. We illustrate the advantages of the causal framework through scenarios involving various stakeholders—legal auditors, service providers, and end-users—and highlight technical aspects relevant to 6G networks. The stakeholder groups and related applications, which we detail below, are as follows:

1) Human forensic post-hoc analysis by service providers and legal auditors
2) Human real-time analysis for resource management by service providers and end-users
3) Machine native AI network real-time optimisation

Next, we present example use cases related to them. Table 4 summarises the following content.

Bayesian Approaches in Spectrum Management: Reference [156] analysed the use of deep learning in channel management where the management optimisation was performed through a Bayesian technique which provides a certain degree of explainability, since it creates a statistical model of the system based on the knowledge about it. While Bayesian approaches focus on the relationship between the prior knowledge and the posterior estimated from it, causal models are focused on the causal relationship between the system variables, independently of our knowledge about

TABLE 4. Examples of causal AI use cases and stakeholders.

Application	Use Case Example	Stakeholder Example	Causal Approach
Monitoring, access, and management of the spectrum	Spectrum management to ensure high-quality connection for a vehicle-to-everything intelligent traffic safety application	Post-hoc analysis by legal auditor	Causal discovery
Resource management (massive channel access, interference management, hand-off management, IoT coverage, etc.)	Smartphone app resources management for connection diagnostics	Real-time analysis for bandwidth management by end-user	Counterfactual analysis
Edge computing	Edge app sharing vehicles real-time video streams to assist in traffic manoeuvres	Machine native AI network real-time optimisation by service provider	Causal inference for intervention estimation analysis

it. Causal analysis does not exhibit a sensitivity to the priors [157]. This may represent an advantage for certain explainability aspects. For example, consider a scenario of a critical service requiring a high-quality connection [158] (i.e., ultra-reliable low latency communication, high spectrum efficiency, etc), such as a vehicle-to-everything intelligent traffic safety application. In this scenario, a legal auditor stakeholder may be investigating an accident produced by the failure of the application due to a failed modelling of the channel requirements. The auditor should not consider explanations focused on -and thus highly dependent on- the prior knowledge observed up to that moment, but used instead explanations with a focus on the true causal mechanism of the system and how it affected the outcome. Thus, a causal approach may fit better the legal auditor explainability requirements.

In [156], the Bayesian approach is used as a way to optimise over unknown and latent parameters given some observed data in a channel estimation use case. In that sense is a more efficient approach than naive optimisation which disregards using the modelled priors. However, it is still less efficient than using in addition the knowledge of the causal relationship between the variables [159], which would be disregarded in a standard Bayesian optimisation. This latter work shows how a causal approach can be in some cases complementary to other approaches to explainability.

Counterfactual Explanations in Resource Management: Another relevant application of AI in 6G is resource management [160]. This includes aspects such as massive channel access, interference management, hand-off management, IoT coverage extension, etc. [161] consider multi-modal traffic classification using deep learning and apply SHAP techniques [162] for its explainability. These allow them to explain which set of input features contributes the most to the confidence probability value associated with the traffic of a given mobile app. It works by comparing the outcome of the model with the outcome of a model where a particular feature is withheld, analysing the average of this procedure for all the features.

While these explanations rely on some kind of counter-factuality, their use is very limited. The technique only analyses the local counter-factuality of the presence or absence of each feature concerning the particular outcome observed. In contrast, complete causal models can obtain the result of any counterfactual question over any outcome. That is, the counter-factuality is not just a binary question over the features' presence but allows the learning of how any particular value of each feature would have affected the results. Additionally, in a causal approach, this is done by relying on a global causal model and not just a local approximation of it. For example, in the case analysed in the previous work, we can consider the perspective of an end-user stakeholder wanting to know why their phone connection is experiencing problems. Using the techniques proposed in the article may produce a diagnostic pointing to the traffic of a specific app, which may lead the user to deactivate completely the app driving this process.

A complete causal model may produce instead a more complex counterfactual explanation, which may alternatively lead to a rebalance of the traffic of various apps, which may allow keeping all of them active, instead of producing a single binary change, which may be a more desirable scenario for the end-user stakeholder.

Requirement Forecasting in Edge Computing: The last AI application in 6G that will be considered is the edge computing case. Reference [163] study a scenario where real-time video streams are shared by vehicles to assist in traffic manoeuvres. In this scenario, the video streams are shared by an application running at the edge of the network. The AI application forecasts the Quality-of-Experience (QoE) perceived by the users based on QoE and QoS metrics as well as other information such as cell usage. The authors show how federated learning outperforms local or centralised learning. In this work the explainability of the AI models is produced using fuzzy rules [164]. Given this use case, we can consider a scenario where a service provider stakeholder uses the edge-AI models to optimise the QoE based on the previous data. Common induction algorithms for fuzzy rules are based on correlations, and thus this scenario may lead to misidentifying confounders as causes of the QoE. A causal analysis approach would on one side eliminate such errors, and on the other side would offer the service provider the specific interventions that can be executed to maximise the QoE when it is forecasted to be below a certain threshold. This possibility of operating both at the observational and interventional levels is unique to the causal framework. Authors such as [165] have shown how to combine causality with fuzzy rules in such cases.

In summary, the presented cases show how a causal analysis framework may benefit different stakeholders of the 6G network, from legal auditors to service providers through end-users. Some of the advantages include the possibility of asking counterfactual questions over the system analysed, being able to design specific interventions to increase the efficiency of the network, and a more robust and complete global explainability. As was shown, a causal framework can be applied on certain occasions in combination with other explainable techniques.

3) RECENT WORK IN CAUSAL AI IN COMMUNICATION NETWORKS

Reference [166] showed the possibility of increasing the efficiency of the energy allocation of the communication channels. They verified the causality between channels and used this common causal connection to better forecast their energy requirements. Reference [122] demonstrated the possibility of improving the user experience of the users of the network through the causal modelling of the relationship between performance and quality indicators. Causal discovery techniques were used to build the DAG of such variables showing their causal relationships, which were embedded in a graph attention network. As we mentioned above, such causal knowledge can also be used to infer which variables of the network have to be intervened to maximise the effect on the user experience. Reference [167] analysed how to use causality for a better operation and maintenance of the network by better predicting its conditions of use. Causality was applied to communication data analysis to select the features to use in an LSTM model to predict future call volume.

The exponential expansion in the amount of data expected in 6G networks together with the embedding of machine learning models at all network levels is leading to a recent increase in studies specifically targeted to causality and machine learning for 6G. Reference [168] applied causal discovery approaches to identify causal factors determining network performance patterns in mobile wireless networks. As an example of their results, in their datasets they found the uplink throughput to be the most relevant causal factor for the performance and a causal relationship between the number of reserved signalling resources in the physical uplink control channel and the uplink throughput. In a very recent work, [169] developed a very comprehensive vision of the advantages and principles for a causality-driven AI-native wireless network, including RL models. They identified challenges in the current use of AI in wireless systems, showed the advantages of the use of causality illustrated by use cases such as dynamic channel tracking, digital twin modelling, and ISAC, and proposed a causal inference-based framework for wireless control problems. In [170], some of the previous authors used causal representation learning concepts to design reasoning-driven semantic communication networks. This work also included proposals for reasoning capacity measures for computing and communication resources. The authors proposed definitions, visions, and building blocks of an end-to-end semantic communication network.

These previous studies justify the need for and advantages of further work on causality in 6G networks.

Examples of causal analysis applications for AI-driven 6G decisions include:

- increased performance in channel estimation and traffic prediction using causality for transfer learning between the AI estimation/prediction models;
- improved performance in training global cross-layer AI models for integrated sensing and communications by using causality to reduce the search space of redundant/confounding data sources;
- causal post-hoc analysis of AI-driven 6G decisions of spectrum management in ultra-reliable low latency communications;
- better performance of real-time bandwidth management by estimating counterfactual AI-driven decisions of channel access;
- better resource forecasting in edge computing using causal AI for improved estimation of QoE.

VI. XAI IN 6G USE CASES

In this section, we focus on two key 6G use cases—network slicing and autonomous robotic systems—where XRL and causal analysis play an essential role in addressing real-world deployment challenges. These use cases are selected based on their criticality to 6G infrastructure, their reliance on AI-driven decision-making, and the pressing need for interpretability in their operations. Network slicing is key to 6G's efficient resource management, but opaque AI policies can undermine compliance, SLA assurance, and trust. Similarly, UAV-based robotic systems enhance adaptive connectivity but require transparent decision-making for safety and reliability. While many 6G applications benefit from XAI, these two domains serve as representative examples where the integration of XRL and causal reasoning is not only beneficial but essential for real-world viability.

A. NETWORK SLICING

Network slicing, a key enabler for tailored service delivery in 6G networks, represents a complex resource allocation challenge that has been significantly enhanced through AI-driven approaches. As a combinatorial optimization problem [171], it has evolved significantly through DRL applications, demonstrating substantial improvements in various aspects of network management. Pioneering studies showcase DRL's effectiveness: [172] developed an intelligent slice admission control framework, while [173] and [174] achieved revenue improvements of up to 54.5% and 17% respectively through efficient slice orchestration. Recent advancements have further enhanced this domain, with [175] proposing a dynamic RAN slicing model that categorizes services into throughput-oriented, delay-sensitive, and delay-throughput-tolerant types while addressing system stability across heterogeneous traffic demands. Complementing this work, [176], [177] introduced a two-phase approach combining optimization theory and DRL to balance eMBB data rates with URLLC reliability constraints. This framework effectively manages the challenging trade-off between service types, demonstrating improved eMBB reliability while maintaining URLLC performance through dynamic resource allocation. Additional contributions include the EXPLORA framework [178], which enhances explainability in Open RAN systems, and autonomous slicing refinement algorithm [179], achieving up to 100% user satisfaction and 80% resource utilization.

While RL and DRL have demonstrated significant success in addressing sequential decision-making challenges in network resource management, their inherent opacity presents substantial implementation barriers. The emerging XRL approaches provide insights into environmental perception, motivational factors, and Q-value computations [180], making them particularly crucial for practical deployment in network resource management scenarios.

XRL has been pivotal in addressing transparency challenges in resource allocation. For example, [181] proposes an XRL framework for 6G networks that enhances both performance and interpretability. Their framework introduces an intrinsic interpretability approach that combines SHAP values with an entropy mapper mechanism. The AI model encompasses average SNR values, traffic volume, and remaining capacity metrics in its state space, while determining discrete allocation of Physical Resource Blocks (PRBs) through its action space. A distinctive feature of their approach is the composite reward mechanism that integrates traditional SLA-based rewards with an XAI reward derived from SHAP importance values and entropy calculations, guiding DRL agents toward more interpretable resource allocation decisions. The framework employs multiple XRL agents to allocate physical resource blocks across different network slices (URLLC, eMBB, and mMTC) while meeting specific SLA requirements. The SHAP values generate probability distributions over state-action pairs, while the entropy mapper calculates uncertainty metrics for selected actions, using the inverse maximum entropy as an XAI reward component. Their experimental results demonstrate the XRL approach's superiority over traditional RL baselines, achieving improved latency performance (1.9 ms versus 3.5 ms at the 50th percentile for URLLC) and reduced dropped traffic rates (5.2% versus 7.9% for mMTC). This research represents one of the first initiatives to incorporate explainability directly into the DRL training process rather than using it as a post-hoc analysis tool.

The evaluation of XRL techniques in network slicing requires specific metrics that address both the performance and the explainability aspects. Current research highlights four essential metrics for comprehensive assessment: (1) *SLA Violation Rate*, which measures how well XRL systems maintain promised service quality across different slice types (eMBB, URLLC, mMTC), providing insight into fault detection and service reliability [5]; (2) *Resource Utilization Efficiency*, assessing how optimally network resources are allocated, with evidence suggesting that better explanations lead to more efficient resource management; (3) *Adaptation Responsiveness*, evaluating how quickly and effectively XRL systems can reconfigure slices when facing changing network conditions and traffic patterns [182]; and (4) *Explanation-Action Alignment*, measured through SHAP-alignment scores and the Slice-Trust Index, which quantifies how well the system's explanations reflect its actual resource allocation decisions. The Explainability-SLA Balance (ESB) metric further enhances the evaluation by measuring the trade-off between improved transparency and potential performance impacts [183]. These metrics offer a structured approach to evaluate XRL techniques in network slicing, addressing both technical requirements and the growing demand for trustworthy AI in telecommunication systems.

RM (discussed in Section IV-B) is another approach to iently incorporate interpretability into RL tasks, which is independent of the RL state and action spaces, making them suitable for guiding RL algorithms in ultra-large networks. RMs can be exploited for the problem of network slicing in two (or multiple hierarchical) levels. The higher level RM in xApp near-real-time RAN intelligent controller (NR-RIC) explains and guides the policies made for allocation of PRBs across various slices (eMBB, uRLLC, and mMTC) over the longer slice-time (of multiple time-slots), and the lower level RM in real-time RIC explains the decision made by admission control within a slice to accept or refuse a service request upon arrival during the shorter time-slot. This can subsequently guide the PRB scheduling policy over the admitted requests of the slice to meet the QoS requirements. The RM approach can provide modular explainability and also at a long-term policy level as well as reward level to explain the optimal PRB allocation policy over slices that can guarantee successful service admissions and minimise the denial of services in the network and failure due to the limited resources.

The above XRL application examples deliver substantial benefits across multiple stakeholder groups. For network service providers, these methods significantly enhance operational efficiency by providing transparent, interpretable insights into resource allocation and admission control decisions. The intrinsic explainability embedded in these frameworks allows providers to proactively manage network

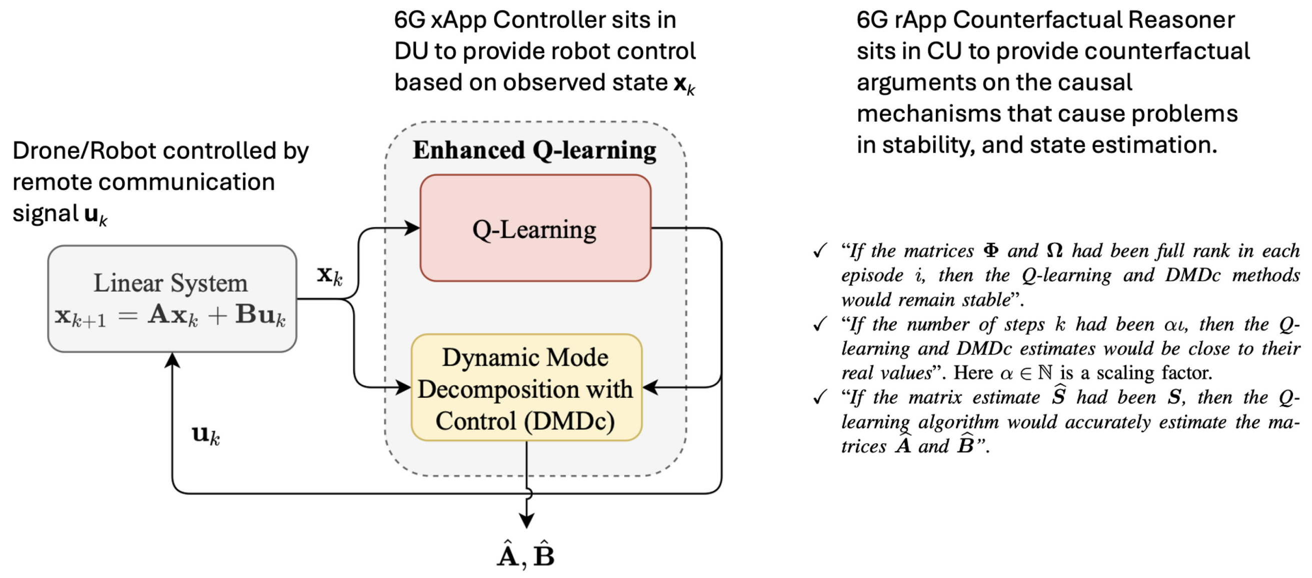

FIGURE 6. Remote wireless control of safety-critical robot/drone using Q-learning in the DU of 6G ORAN, where XRL is used to perform counterfactual reasoning in CU to ensure stability and performance accuracy - details in [184].

resources, mitigate potential bottlenecks, and ensure compliance with SLAs, ultimately improving resource utilization and reducing operational costs. Legal auditors gain from the modular explainability offered by RMs at multiple hierarchical levels, enabling precise monitoring and auditing of network decisions, thus facilitating regulatory compliance and strengthening governance in network management practices. End-users directly benefit through enhanced QoE, as demonstrated by reduced latency and lower service denial rates, fostering greater trust in service reliability and performance.

B. XRL FOR 6G WITH AUTONOMOUS ROBOTIC SYSTEMS

6G networks are poised to integrate autonomous robotic systems, such as UAVs, to enhance connectivity in underserved regions. These UAVs can function as mobile base stations or communication relays, dynamically extending network coverage and improving QoS [185], [186]. However, deploying autonomous UAVs at scale requires efficient decision-making frameworks that are both optimized for performance and safety, in addition to being transparent in their operations. XRL addresses these concerns by ensuring that UAVs make decisions that are interpretable and trustworthy (with provable sub-linear convergence properties and control stability guarantees) for various stakeholders, including service providers, regulators, and end-users [187], [188]. Such approaches, however, will need to be adapted to specific 6G contexts to ensure practical applicability and efficiency.

To ensure seamless service provision, RL-driven UAVs can dynamically adjust their flight paths for optimal resource allocation and service efficiency [189], [190]. RL techniques also facilitate intelligent task offloading and resource allocation, ensuring timely data transmission without overloading communication channels [191]. By leveraging DRL and multi-agent RL (MARL), UAVs can autonomously optimize their decision-making processes to enhance network reliability [189], [190], [192], [193], [194], [195].

Despite significant advancements in RL-driven UAV deployment, the incorporation of explainability remains relatively limited. Several recent studies have explored different XRL approaches to improve transparency and trust: [196] employs Evolving Behavior Trees (EBTs) via genetic programming to enhance system interpretability by integrating explicit safety behaviours; [197] utilizes SHAP values to provide post-hoc explanations of decision-making processes; [198] applies Saliency Maps to identify critical visual input regions influencing a Deep Q-Network (DQN) agent's actions; [199] also uses SHAP-based feature attribution to enhance transparency; [200] introduces human-UAV teamwork with interpretability and transparency through human-agent interaction; [201] explores constrained RL through probabilistic inference for intrinsic interpretability in high-stakes deployment scenarios.

An emerging perspective in XRL involves counterfactual reasoning, which enhances interpretability by identifying causal relationships in decision-making. This approach enables forensic diagnostics by answering "what if" questions, allowing stakeholders to understand the rationale behind an agent's choices. Counterfactual reasoning has been applied in Q-learning for networked control, providing powerful tools for fault diagnosis and prevention [202]. Recent studies, such as [184], integrate counterfactual reasoning with dynamic mode decomposition and control algorithms for state-transition function estimation, improving the robustness of explanations without requiring hyperparameter tuning.

Integrating XRL and counterfactual reasoning in 6G UAV networks benefits key stakeholders in several ways:

- *Service Providers* XRL enables flexible and optimized resource management, ensuring high QoS while maintaining network efficiency. Counterfactual explanations improve system debugging, facilitating better decision-making.
- *Regulators* Explainable UAV trajectory modifications ensure compliance with network and safety standards, making regulatory audits more transparent and effective.
- *End-Users* Enhanced transparency fosters trust in UAV-assisted communication services, ensuring reliable and interpretable connectivity solutions.

The synergy between XRL and causal AI within 6G infrastructure offers promising avenues for developing intelligent RAN controllers that can integrate, coordinate, and manage AI-enabled subsystems like UAVs [203]. By ensuring explainability and accountability, these technologies pave the way for secure, efficient, and stakeholder-friendly UAV operations for future 6G networks.

VII. CONCLUSION & FUTURE RESEARCH AREAS

Future telecommunications are set to increasingly integrate critical services into their network infrastructures, raising significant trust and safety concerns. The AI/ML modules orchestrating these critical services will inevitably rely on DRL to process multi-modal requirements datasets and make semantically modulated decisions. Despite its potential, DRL presents a critical challenge: its lack of explainability, which remains a key area of concern for the research community.

First, we reviewed how the explanations must cater for diverse telecommunications stakeholders, including network operators, service providers, and end-users, each with unique goals and operational practices. Second, when DRL lacks prior models or established frameworks to guide the creation of meaningful explanations, we reviewed key emerging research approaches to help tackle this problem for different parts of the RL. Finally, we demonstrated how causal explanations can further enhance the framework to improve 6G services.

As such, we advocate for a stakeholder-centric approach to XAI, which is especially challenging for XRL where we have to grapple with concepts such as policy and reward/value functions. Furthermore, as 6G increasingly deals with safety-critical areas such as autonomy (robots, drones, resource controller) and healthcare, explaining causal and not correlated reasons why something happened becomes more critical for insurance and liability. Counterfactual arguments need to be made where possible, and this can be challenging in RL. We have shown how one design can remote control drones and use xApp Q-learning at the DU to operate and create rApp counterfactual arguments at the CU of 6G ORAN. However, challenges such as trade-off between performance and explainability, lack of ground-truth for explanation, integration of XAI systems into the existing network infrastructure may still affect the real-world implementation. To this end, companies such as ERICSSON have tried to come up with solutions for XRL in adjusting the antenna for better KPIs and coverage region by USs.

In future work, we would like to think that all of these approaches may be too overwhelming for innovators and end-users outside the telecommunication and AI industry. Thankfully the recent advances in LLMs may help them specify requirements in communications, edge AI, privacy, and security. Our recent advances in 6G LLMs have embedded standards, research papers, and best practices in the LLM as an RAG database and can provide the knowledge we shared today as part of the 6G design specification for diverse but important end-user requirements [204].

Conclusion: Emerging techniques in explaining machine learning, especially RL, need to be tailored to the specific application, the usage context, and the human stakeholder involved. Future safety-critical 6G applications require more than correlated explanations: they need causal arguments and counterfactual reasoning for different stakeholders.

REFERENCES

[1] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[2] European Commission. "Ethics guidelines for trustworthy AI." Accessed: Aug. 20, 2024. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[3] C. Li, W. Guo, S. C. Sun, S. Al-Rubaye, and A. Tsourdos, "Trustworthy deep learning in 6G-enabled mass autonomy: From concept to quality-of-trust key performance indicators," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 112–121, Dec. 2020.

[4] W. Guo, "Explainable artificial intelligence for 6G: Improving trust between human and machine," *IEEE Commun. Mag.*, vol. 58, no. 6, pp. 39–45, Jun. 2020.

[5] S. Wang et al., "Explainable AI for 6G use cases: Technical aspects and research challenges," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 2490–2540, 2024.

[6] B. Brik et al., "A survey on explainable AI for 6G O-RAN: Architecture, use cases, challenges and research directions," 2023, *arXiv:2307.00319*.

[7] A. B. Arrieta et al., "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020.

[8] C. Glanois et al., "A survey on interpretable reinforcement learning," *Mach. Learn.*, vol. 113, no. 8, pp. 5847–5890, Aug. 2024. [Online]. Available: https://doi.org/10.1007/s10994-024-06543-w

[9] S. Roy, H. Chergui, and C. Verikoukis, "Toward bridging the FL performance-explainability tradeoff: A trustworthy 6G RAN slicing use-case," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 10529–10538, Jul. 2024.

[10] M. Nauta et al. "From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable AI." 2022. [Online]. Available: https://arxiv.org/abs/2201.08164

[11] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," 2017, *arXiv:1702.08608*.

[12] S. Pateria, B. Subagdja, A.-H. Tan, and C. Quek, "Hierarchical reinforcement learning: A comprehensive survey," *ACM Comput. Surveys*, vol. 54, no. 5, pp. 1–35, Jun. 2021. [Online]. Available: https://doi.org/10.1145/3453160

[13] J. Merel et al. "Neural probabilistic motor primitives for humanoid control." 2018. [Online]. Available: http://arxiv.org/abs/1811.11711

[14] A. Mohan, A. Zhang, and M. Lindauer, "Structure in deep reinforcement learning: A survey and open problems," *J. Artif. Intell. Res.*, vol. 79, pp. 1167–1236, Apr. 2024. [Online]. Available: http://dx.doi.org/10.1613/jair.1.15703

[15] H. Zhang, Z. Gao, Y. Zhou, H. Zhang, K. Wu, and F. Lin. "Faster and safer training by embedding high-level knowledge into deep reinforcement learning." 2019. [Online]. Available: http://arxiv.org/abs/1910.09986

[16] K. Lu, S. Zhang, P. Stone, and X. Chen. "Robot representing and reasoning with knowledge from reinforcement learning." 2018. [Online]. Available: http://arxiv.org/abs/1809.11074

[17] A. Silva, T. Killian, I. D. J. Rodriguez, S.-H. Son, and M. Gombolay. "Optimization methods for interpretable differentiable decision trees in reinforcement learning." 2020. [Online]. Available: https://arxiv.org/abs/1903.09338

[18] D. Hein, S. Udluft, and T. A. Runkler. "Interpretable policies for reinforcement learning by genetic programming." 2017. [Online]. Available: http://arxiv.org/abs/1712.04170

[19] M. Landajuela et al., "Discovering symbolic policies with deep reinforcement learning," in *Proc. 38th Int. Conf. Mach. Learn.*, Jul. 2021, pp. 5979–5989. [Online]. Available: https://proceedings.mlr.press/v139/landajuela21a.html

[20] Z. Jiang and S. Luo, "Neural logic reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 3110–3119.

[21] M. Zimmer et al. "Differentiable logic machines." 2021. [Online]. Available: https://arxiv.org/abs/2102.11529

[22] Z. Ma et al. "Learning symbolic rules for interpretable deep reinforcement learning." 2021. [Online]. Available: https://arxiv.org/abs/2103.08228

[23] O. Bastani, Y. Pu, and A. Solar-Lezama. "Verifiable reinforcement learning via policy extraction." 2018. [Online]. Available: http://arxiv.org/abs/1805.08328

[24] S. Milani, N. Topin, M. Veloso, and F. Fang, "Explainable reinforcement learning: A survey and comparative review," *ACM Comput. Surveys*, vol. 56, no. 7, p. 168, Apr. 2024. [Online]. Available: https://doi.org/10.1145/3616864

[25] A. Atrey, K. Clary, and D. D. Jensen. "Exploratory not explanatory: Counterfactual analysis of saliency maps for deep RL." 2019. [Online]. Available: http://arxiv.org/abs/1912.05743

[26] J. Adebayo, J. Gilmer, M. Muelly, I. J. Goodfellow, M. Hardt, and B. Kim. "Sanity checks for saliency maps." 2018. [Online]. Available: http://arxiv.org/abs/1810.03292

[27] J. van der Waa, J. van Diggelen, K. van den Bosch, and M. A. Neerincx. "Contrastive explanations for reinforcement learning in terms of expected consequences." 2018. [Online]. Available: http://arxiv.org/abs/1807.08706

[28] Z. Yang, X. Jia, H. Li, and J. Yan. "LLM4Drive: A survey of large language models for autonomous driving." 2024. [Online]. Available: https://arxiv.org/abs/2311.01043

[29] P. Wang, X. Wei, F. Hu, and W. Han. "TransGPT: Multi-modal generative pre-trained transformer for transportation." 2024. [Online]. Available: https://arxiv.org/abs/2402.07233

[30] D. Gunning and D. Aha, "DARPA's explainable artificial intelligence (XAI) program," *AI Mag.*, vol. 40, no. 2, pp. 44–58, 2019.

[31] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, "XAI—explainable artificial intelligence," *Sci. Robot.*, vol. 4, no. 37, 2019, Art. no. eaay7120.

[32] T. Senevirathna, V. H. La, S. Marchal, B. Siniarski, M. Liyanage, and S. Wang, "A survey on XAI for beyond 5G security: Technical aspects, use cases, challenges and research directions," 2022, *arXiv:2204.12822*.

[33] T. D. Grant and D. J. Wischik, "Show us the data: Privacy, explainability, and why the law can't have both," *George Washington Law. Rev.*, vol. 88, no. 6, p. 1350, 2020.

[34] M. Langer et al., "What do we want from explainable artificial intelligence (XAI)? A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research," *Artif. Intell.*, vol. 296, Jul. 2021, Art. no. 103473.

[35] H. Kopecka and J. Such, "Explainable AI for cultural minds," in *Proc. Workshop Dial. Explanation Argumentation Human-Agent Interact.*, 2020, p. 45.

[36] H. Kopecka, J. Such, and M. Luck, "Preferences for AI explanations based on cognitive style and socio-cultural factors," in *Proc. ACM Human–Comput. Interact.*, vol. 8, 2024, pp. 1–32.

[37] M. Ribera and A. Lapedriza, "Can we do better explanations? A proposal of user-centered explainable AI," in *Proc. ACM IUI Workshops*, 2019, p. 89.

[38] S. Mohseni, N. Zarei, and E. D. Ragan, "A multidisciplinary survey and framework for design and evaluation of explainable AI systems," *ACM Trans. Interact. Intell. Syst.*, vol. 11, nos. 3–4, pp. 1–45, 2021.

[39] A. Preece, D. Harborne, D. Braines, R. Tomsett, and S. Chakraborty, "Stakeholders in explainable AI," 2018, *arXiv:1810.00184*.

[40] European Commission. "The EU AI act." Accessed: Aug. 20, 2024. [Online]. Available: https://artificialintelligenceact.eu/the-act/

[41] A. B. Haque, A. N. Islam, and P. Mikalef, "Explainable artificial intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research," *Technol. Forecast. Soc. Change*, vol. 186, Jan. 2023, Art. no. 122120.

[42] W. Jin, J. Fan, D. Gromala, P. Pasquier, and G. Hamarneh, "EUCA: The end-user-centered explainable AI framework," 2021, *arXiv:2102.02437*.

[43] V. Hassija et al., "Interpreting black-box models: A review on explainable artificial intelligence," *Cogn. Comput.*, vol. 16, no. 1, pp. 45–74, 2024.

[44] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data Soc.*, vol. 3, no. 2, p. 20, 2016.

[45] European Commission. "Regulation (EU) 2016/679 (general data protection regulation)." Accessed: Sep. 2, 2024. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679

[46] OECD. "AI principles overview—OECD.AI." Accessed: Sep. 2, 2024. [Online]. Available: https://oecd.ai/en/ai-principles

[47] I. S. Association. "Ethically aligned design." Accessed: Sep. 2, 2024. [Online]. Available: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

[48] GPAI. "Global partnership on artificial intelligence." Accessed: Sep. 2, 2024. [Online]. Available: https://gpai.ai/

[49] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Acad. Manag. Rev.*, vol. 20, no. 3, pp. 709–734, 1995.

[50] Z. Abou El Houda, B. Brik, and S.-M. Senouci, "A novel IoT-based explainable deep learning framework for intrusion detection systems," *IEEE Internet Things Mag.*, vol. 5, no. 2, pp. 20–23, Jun. 2022.

[51] Z. Abou El Houda, B. Brik, and L. Khoukhi, ""why should i trust your IDs? An explainable deep learning framework for intrusion detection systems in Internet of Things networks," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1164–1176, 2022.

[52] Z. Abou El Houda, H. Moudoud, B. Brik, and L. Khoukhi, "Securing federated learning through blockchain and explainable AI for robust intrusion detection in IoT networks," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2023, pp. 1–6.

[53] V. Arya et al., "One explanation does not fit all: A toolkit and taxonomy of ai explainability techniques," 2019, *arXiv:1909.03012*.

[54] K. Sokol and P. Flach, "One explanation does not fit all: The promise of interactive explanations for machine learning transparency," *Künstliche Intelligenz*, vol. 34, no. 2, pp. 235–250, 2020.

[55] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial-intelligence-enabled intelligent 5G networks," *IEEE Netw.*, vol. 34, no. 6, pp. 272–280, Nov./Dec. 2020.

[56] A. Heuillet, F. Couthouis, and N. Díaz-Rodríguez, "Explainability in deep reinforcement learning," *Knowl. Based Syst.*, vol. 214, Feb. 2021, Art. no. 106685.

[57] Y. Qing, S. Liu, J. Song, H. Wang, and M. Song, "A survey on explainable reinforcement learning: Concepts, algorithms, challenges," 2022, *arXiv:2211.06665*.

[58] Y. Bekkemoen, "Explainable reinforcement learning (XRL): A systematic literature review and taxonomy," *Mach. Learn.*, vol. 113, no. 7, pp. 355–441, 2023.

[59] S. C. Sun, C. Li, Z. Wei, A. Tsourdos, and W. Guo, "Scalable partial explainability in neural networks via flexible activation functions (student abstract)," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, May 2021, pp. 15899–15900. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/17946

[60] S. C. Sun and W. Guo, "Approximate symbolic explanation for neural network enabled water-filling power allocation," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, 2020, pp. 1–4.
[61] S. Blakeman and D. Mareschal, "A complementary learning systems approach to temporal difference learning," *Neural Netw.*, vol. 122, pp. 218–230, May 2020.
[62] R. Parr and S. Russell, "Reinforcement learning with hierarchies of machines," in *Proc. Adv. Neural Inf. Process. Syst.*, 1997, pp. 1043–1049.
[63] M. Leonetti, L. Iocchi, and F. Patrizi, "Automatic generation and learning of finite-state controllers," in *Proc. Artif. Intell. Methodol. Syst. Appl.*, 2012, pp. 135–144.
[64] R. Toro Icarte, E. Waldie, T. Klassen, R. Valenzano, M. Castro, and S. McIlraith, "Learning reward machines for partially observable reinforcement learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 15497–15508.
[65] R. T. Icarte, T. Klassen, R. Valenzano, and S. McIlraith, "Using reward machines for high-level task specification and decomposition in reinforcement learning," in *Proc. 35th Int. Conf. Mach. Learn.*, 2018, pp. 2107–2116.
[66] Z. Xu et al., "Joint inference of reward machines and policies for reinforcement learning," in *Proc. Int. Conf. Automated Plan. Schedul.*, 2020, pp. 590–598.
[67] A. Koul, A. Fern, and S. Greydanus, "Learning finite state representations of recurrent policy networks," in *Proc. 7th Int. Conf. Learn. Rep. (ICLR)*, 2019, p. 12. [Online]. Available: https://openreview.net/forum?id=S1gOpsCctm
[68] D. Furelos-Blanco, M. Law, A. Russo, K. Broda, and A. Jonsson, "Induction of subgoal automata for reinforcement learning," in *Proc. 34th AAAI Conf. Artif. Intell. (AAAI)*, 2020, pp. 3890–3897.
[69] D. Furelos-Blanco, M. Law, A. Jonsson, K. Broda, and A. Russo, "Induction and exploitation of subgoal automata for reinforcement learning," *J. Artif. Intell. Res.*, vol. 70, no. 1, pp. 1031–1116, 2021.
[70] M. Gelfond and Y. Kahl, *Knowledge Representation, Reasoning, and the Design of Intelligent Agents: The Answer-Set Programming Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[71] S. Muggleton and L. de Raedt, "Inductive logic programming: Theory and methods," *J. Logic Program.*, vols. 19–20, no. S1, pp. 629–679, 1994.
[72] D. Furelos-Blanco, M. Law, A. Jonsson, K. Broda, and A. Russo, "Induction and exploitation of subgoal automata for reinforcement learning," *J. Artif. Intell. Res.*, vol. 70, pp. 1031–1116, May 2021, doi: 10.1613/jair.1.12372.
[73] M. Law, A. Russo, and K. Broda. "The ILASP system for inductive learning of answer set programs." 2020. [Online]. Available: https://arxiv.org/abs/2005.00904
[74] R. Parać, L. Nodari, L. Ardon, D. Furelos-Blanco, F. Cerutti, and A. Russo, "Learning robust reward machines from noisy labels," in *Proc. 21st Int. Conf. Principles Knowl. Rep. Reason.*, 2024, p. 1445.
[75] M. Law, A. Russo, and K. Broda, "Inductive learning of answer set programs," in *Logic-Based Artificial Intelligence*, E. Fermé and J. Leite, Eds., Cham, Switzerland: Springer Int. Publ., 2014, pp. 311–325. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-11558-0_22
[76] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020.
[77] A. D. Garcez et al., "Neural-symbolic learning and reasoning: Contributions and challenges," in *Proc. AAAI Spring Symposia*, 2015, p. 5.
[78] T. R. Besold et al., *Neuro-Symbolic Artificial Intelligence: The State-of-the-Art: Chapter 1—Neural-Symbolic Learning and Reasoning: A Survey and Interpretation*. New York, NY, USA: IOS Press, 2021, pp. 1–51.
[79] A. D. Garcez and L. C. Lamb, "Neurosymbolic AI: The 3rd wave," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 12387–12406, 2023.
[80] M. Kamruzzaman Sarker, L. Zhou, A. Eberhart, and P. Hitzler. "Neuro-symbolic artificial intelligence: Current trends." 2021. [Online]. Available: https://arxiv.org/abs/2105.05330
[81] D. Yu, B. Yang, D. Liu, H. Wang, and S. Pan, "A survey on neural-symbolic learning systems," *Neural Netw.*, vol. 166, pp. 105–126, Sep. 2023.
[82] M. Garnelo, K. Arulkumaran, and M. Shanahan. "Towards deep symbolic reinforcement learning." 2016. [Online]. Available: https://arxiv.org/abs/1609.05518
[83] A. D. Garcez, A. R. R. Dutra, and E. Alonso, "Towards symbolic reinforcement learning with common sense," 2018, *arxiv.abs/1804.08597*.
[84] P. Barbiero et al., "Interpretable neural-symbolic concept reasoning," in *Proc. 40th Int. Conf. Mach. Learn. (ICML)*, 2023, pp. 1801–1825.
[85] D. Cunnington, M. Law, J. Lobo, and A. Russo, "Neuro-symbolic learning of answer set programs from raw data," in *Proc. 32nd Int. Joint Conf. Artif. Intell. (IJCAI)*, 2023, pp. 3586–3596.
[86] A. d'Avila Garcez, A. R. R. Dutra, and E. Alonso, "Towards symbolic reinforcement learning with common sense," 2018, *arXiv:1804.08597*.
[87] R. Manhaeve, S. Dumančić, A. Kimmig, T. Demeester, and L. De Raedt, "Neural probabilistic logic programming in DeepProbLog," *Artif. Intell.*, vol. 298, Sep. 2021, Art. no. 103504.
[88] Z. Yang, A. Ishay, and J. Lee, "NeurASP: Embracing neural networks into answer set programming," in *Proc. 29th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2021, pp. 1–7.
[89] A. Santoro et al., "A simple neural network module for relational reasoning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 4967–4976.
[90] R. Riegel et al., "Logical neural networks," 2020, *arXiv:2006.13155*.
[91] J. Xu, Z. Zhang, T. Friedman, Y. Liang, and G. Broeck, "A semantic loss function for deep learning with symbolic knowledge," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 5502–5511.
[92] R. Evans and E. Grefenstette, "Learning explanatory rules from noisy data," *J. Artif. Intell. Res.*, vol. 61, no. 1, pp. 1–64, Jan. 2018.
[93] P. Sen, B. W. de Carvalho, R. Riegel, and A. Gray, "Neuro-symbolic inductive logic programming with logical neural networks," in *Proc. AAAI Conf. Artif. Intell.*, vol. 36, 2022, pp. 8212–8219.
[94] E. Tsamoura, T. Hospedales, and L. Michael, "Neural-symbolic integration: A compositional perspective," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, 2021, pp. 5051–5060.
[95] W.-Z. Dai, Q. Xu, Y. Yu, and Z.-H. Zhou, "Bridging machine learning and logical reasoning by Abductive learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 2811–2822.
[96] W.-Z. Dai and S. Muggleton, "Abductive knowledge induction from raw data," in *Proc. 13th Int. Joint Conf. Artif. Intell. (IJCAI)*, Aug. 2021, pp. 1845–1851.
[97] K. Acharya, W. Raza, C. Dourado, A. Velasquez, and H. H. Song, "Neurosymbolic reinforcement learning and planning: A survey," *IEEE Trans. Artif. Intell.*, vol. 5, no. 5, pp. 1939–1953, May 2024.
[98] D. Kimura et al., "Neuro-symbolic reinforcement learning with first-order logic," in *Proc. Conf. Empirical Methods Nat. Lang. Process.*, 2021, pp. 3505–3511.
[99] G. Anderson, A. Verma, I. Dillig, and S. Chaudhuri, "Neurosymbolic reinforcement learning with formally verified exploration," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 6172–6183.
[100] B. G. León, M. Shanahan, and F. Belardinelli, "Relational deep reinforcement learning and latent goals for following instructions in temporal logic," in *Proc. 38th Int. Conf. Mach. Learn. Workshop Uncertainty Robustness Deep Learn.*, 2021, pp. 1–8.
[101] R. Hazra and L. De Raedt, "Deep explainable relational reinforcement learning: A Neuro-symbolic approach," in *Proc. Mach. Learn. Knowl. Disc. Databases Res. Track*, 2023, pp. 213–229.
[102] L. Mitchener, D. Tuckey, M. Crosby, and A. Russo, "Detect, understand, act: A Neuro-symbolic hierarchical reinforcement learning framework," *Mach. Learn.*, vol. 111, no. 4, pp. 1523–1549, 2022.
[103] H. Dong, J. Mao, T. Lin, C. Wang, L. Li, and D. Zhou, "Neural logic machines," in *Proc. Int. Conf. Learn. Rep.*, 2019, p. 9.
[104] M. Hasanbeig, N. Y. Jeppu, A. Abate, T. Melham, and D. Kroening, "DeepSynth: Automata synthesis for automatic task segmentation in deep reinforcement learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, 2021, pp. 7647–7656.

[105] A. Velasquez et al., "Dynamic automaton-guided reward shaping for monte carlo tree search," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, 2021, pp. 12015–12023.
[106] A. Velasquez et al., "Multi-agent tree search with dynamic reward shaping," in *Proc. Int. Conf. Autom. Plan. Schedul.*, vol. 32, 2022, pp. 652–661.
[107] M. S. Munir et al., "Neuro-symbolic explainable artificial intelligence twin for zero-touch IoE in wireless network," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 22451–22468, Dec. 2023.
[108] C. K. Thomas and W. Saad, "Neuro-symbolic artificial intelligence (AI) for intent based semantic communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2022, pp. 2698–2703.
[109] F. Rezazadeh et al., "Toward explainable reasoning in 6G: A proof of concept study on radio resource allocation," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 6239–6260, 2024.
[110] Z. Juozapaitis, A. Koul, A. Fern, M. Erwig, and F. Doshi-Velez, "Explainable reinforcement learning via reward decomposition," in *Proc. Int. Joint Conf. Artif. Intell. Workshop Explain. Artif. Intell.*, 2019, pp. 1–8.
[111] P. Sunehag et al., "Value-decomposition networks for cooperative multi-agent learning based on team reward," in *Proc. 17th Int. Conf. Auton. Agents MultiAgent Syst. (AAMAS)*, 2018, pp. 2085–2087.
[112] T. Rashid, M. Samvelyan, C. S. De Witt, G. Farquhar, J. Foerster, and S. Whiteson, "Monotonic value function factorisation for deep multi-agent reinforcement learning," *J. Mach. Learn. Res.*, vol. 21, no. 178, pp. 1–51, 2020.
[113] K. Son, D. Kim, W. J. Kang, D. E. Hostallero, and Y. Yi, "QTRAN: Learning to factorize with transformation for cooperative multi-agent reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 5887–5896.
[114] P. e. a. Iturria-Rivera, "Explainable multi-agent reinforcement learning for extended reality codec adaptation," 2024, *arXiv:2411.14264*.
[115] A. Perrusquía, M. Zou, and W. Guo, "Explainable data-driven Q-learning control for a class of discrete-time linear autonomous systems," *Inf. Sci.*, vol. 682, Nov. 2024, Art. no. 121283.
[116] A. Perrusquía, W. Guo, B. Fraser, and Z. Wei, "Uncovering drone intentions using control physics informed machine learning," *Nat. Commun. Eng.*, vol. 3, p. 36, Feb. 2024.
[117] N. Khan, S. Coleri, A. Abdallah, A. Celik, and A. M. Eltawil, "Explainable and robust artificial intelligence for trustworthy resource management in 6G networks," *IEEE Commun. Mag.*, vol. 62, no. 4, pp. 50–56, Apr. 2024.
[118] J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[119] M. Glymour, J. Pearl, and N. P. Jewell, *Causal Inference in Statistics: A Primer*. Hoboken, NJ, USA: Wiley, 2016.
[120] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*. Cambridge, MA, USA: MIT Press, 2017.
[121] P. Spirtes, C. N. Glymour, R. Scheines, and D. Heckerman, *Causation, Prediction, and Search*. Cambridge, MA, USA: MIT Press, 2000.
[122] K. Zhang, G. Chuai, W. Gao, and Q. Liu, "Cellular QoE prediction for video service based on causal structure learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 7, pp. 7541–7551, Jul. 2023.
[123] P. Lippe, S. Magliacane, S. Löwe, Y. M. Asano, T. Cohen, and S. Gavves, "Citris: Causal identifiability from temporal intervened sequences," in *Proc. Int. Conf. Mach. Learn.*, 2022, pp. 13557–13603.
[124] L. Buesing et al., "Woulda, Coulda, Shoulda: Counterfactually-guided policy search," in *Proc. Int. Conf. Learn. Rep.*, 2018, pp. 1–6.
[125] R. Guo, L. Cheng, J. Li, P. R. Hahn, and H. Liu, "A survey of learning causality with data: Problems and methods," *ACM Comput. Surveys*, vol. 53, no. 4, pp. 1–37, 2020.
[126] C. Glymour, K. Zhang, and P. Spirtes, "Review of causal discovery methods based on graphical models," *Front. Genet.*, vol. 10, p. 524, Jun. 2019.
[127] D. Malinsky and D. Danks, "Causal discovery algorithms: A practical guide," *Philosophy Compass*, vol. 13, no. 1, 2018, Art. no. e12470.
[128] M. J. Vowels, N. C. Camgoz, and R. Bowden, "D'ya like dags? A survey on structure learning and causal discovery," *ACM Comput. Surveys*, vol. 55, no. 4, pp. 1–36, 2022.
[129] A. Zanga, E. Ozkirimli, and F. Stella, "A survey on causal discovery: Theory and practice," *Int. J. Approx. Reason.*, vol. 151, pp. 101–129, Dec. 2022.
[130] D. Jacob, "CATE meets ML: Conditional average treatment effect and machine learning," *Digit. Finance*, vol. 3, no. 2, pp. 99–148, 2021.
[131] C. Glymour, K. Zhang, P. Spirtes, "Review of causal discovery methods based on graphical models," *Front. Genet.*, vol. 10, p. 524, Jun. 2019, doi: 10.3389/fgene.2019.00524.
[132] B. Schölkopf, *Causality for Machine Learning*, 1st ed. New York, NY, USA: ACM Press, 2022, pp. 765–804.
[133] B. Schölkopf et al., "Toward causal representation learning," *Proc. IEEE*, vol. 109, no. 5, pp. 612–634, May 2021.
[134] J. Kaddour, A. Lynch, Q. Liu, M. J. Kusner, and R. Silva, "Causal machine learning: A survey and open problems," 2022, *arXiv:2206.15475*.
[135] Y. Zeng, R. Cai, F. Sun, L. Huang, and Z. Hao, "A survey on causal reinforcement learning," 2023, *arXiv:2302.05209*.
[136] F. Locatello et al., "Challenging common assumptions in the unsupervised learning of disentangled representations," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 4114–4124.
[137] C. Lu, B. Schölkopf, and J. M. Hernández-Lobato, "Deconfounding reinforcement learning in observational settings," 2018, *arXiv:1812.10576*.
[138] S. Singla and S. Feizi, "Salient ImageNet: How to discover spurious features in deep learning?" 2021, *arXiv:2110.04301*.
[139] S. Beery, G. Van Horn, and P. Perona, "Recognition in terra incognita," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 456–473.
[140] Y. Liu, R. Cadei, J. Schweizer, S. Bahmani, and A. Alahi, "Towards robust and adaptive motion forecasting: A causal representation perspective," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 17081–17092.
[141] Z. Wang, X. Xiao, Z. Xu, Y. Zhu, and P. Stone, "Causal dynamics learning for task-independent state abstraction," 2022, *arXiv:2206.13452*.
[142] P. A. Ortega et al., "Shaking the foundations: Delusions in sequence models for interaction and control," 2021, *arXiv:2110.10819*.
[143] Y. Chen et al., "Learning causally invariant representations for out-of-distribution generalization on graphs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 22131–22148.
[144] J. Kim and J. Canny, "Interpretable learning for self-driving cars by visualizing causal attention," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2942–2950.
[145] M. Yang, F. Liu, Z. Chen, X. Shen, J. Hao, and J. Wang, "CausalVAE: Structured causal disentanglement in variational autoencoder," 2020, *arXiv:2004.08697*.
[146] J. Brehmer, P. De Haan, P. Lippe, and T. S. Cohen, "Weakly supervised causal representation learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 38319–38331.
[147] S. C. Sun and W. Guo, "Forecasting wireless demand with extreme values using feature embedding in Gaussian processes," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, 2021, pp. 1–6.
[148] F. Liu et al., "Integrated sensing and communications: Toward dual-functional wireless networks for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.
[149] Z. Wei et al., "Integrated sensing and communication signals toward 5G-a and 6G: A survey," *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11068–11092, Jul. 2023.
[150] D. K. P. Tan et al., "Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions," in *Proc. 1st IEEE Int. Online Symp. Joint Commun. Sens. (JC S)*, 2021, pp. 1–6.
[151] N. Janbi, I. Katib, A. Albeshri, and R. Mehmood, "Distributed artificial intelligence-as-a-service (DAIaaS) for smarter IoE and 5G environments," *Sensors*, vol. 20, no. 20, p. 5796, 2020.
[152] G. Amponis et al., "Drones in B5G/5G networks as flying base stations," *Drones*, vol. 6, no. 2, p. 39, 2022.
[153] R. Bajracharya, R. Shrestha, S. Kim, and H. Jung, "6G NR-U based wireless infrastructure UAV: Standardization, opportunities, challenges and future scopes," *IEEE Access*, vol. 10, pp. 30536–30555, 2022.
[154] D. Mishra, A. M. Vegni, V. Loscrí, and E. Natalizio, "Drone networking in the 6G era: A technology overview," *IEEE Commun. Stand. Mag.*, vol. 5, no. 4, pp. 88–95, Dec. 2021.

[155] L. Jiao et al., "Advanced deep learning models for 5G: Overview, opportunities and challenges," *IEEE Access*, vol. 12, pp. 133245–133314, 2024.

[156] H. Guo and V. K. Lau, "Robust deep learning for uplink channel estimation in cellular network under inter-cell interference," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1873–1887, Jun. 2023.

[157] G. Amponis et al., "Drones in B5G/6G networks as flying base stations," *Drones*, vol. 6, no. 2, p. 39, 2022.

[158] R. Kaliski, S.-M. Cheng, and C.-F. Hung, "Supporting 6G mission-critical services on O-RAN," *IEEE Internet Things Mag.*, vol. 6, no. 3, pp. 32–37, Sep. 2023.

[159] V. Aglietti, X. Lu, A. Paleyes, and J. González, "Causal Bayesian optimization," in *Proc. Int. Conf. Artif. Intell. Stat.*, 2020, pp. 3155–3164.

[160] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine learning for resource management in cellular and IoT networks: Potentials, current solutions, and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1251–1275, 2nd Quart., 2020.

[161] A. Nascita, A. Montieri, G. Aceto, D. Ciuonzo, V. Persico, and A. Pescapé, "XAI meets mobile traffic classification: Understanding and improving multimodal deep learning architectures," *IEEE Trans. Netw. Service Manag.*, vol. 18, no. 4, pp. 4225–4246, Dec. 2021.

[162] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 4765–4774.

[163] J. L. C. Bárcena et al., "Enabling federated learning of explainable AI models within beyond-5G/6G networks," *Comput. Commun.*, vol. 210, pp. 356–375, Oct. 2023.

[164] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 1, pp. 116–132, Jan./Feb. 1985.

[165] L. Kunitomo-Jacquin, A. Lomet, and J.-P. Poli, "Causal discovery for fuzzy rule learning," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, 2022, pp. 1–8.

[166] Z. Chen and Q. Liang, "Efficient energy power allocation for forecasted channel based on transfer entropy," in *Proc. 8th Int. Conf. Commun. Signal Process. Syst.*, 2020, pp. 1758–1765.

[167] K. Zhang, G. Chuai, W. Gao, X. Liu, S. Maimaiti, and Z. Si, "A new method for traffic forecasting in urban wireless communication network," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, p. 66, 2019.

[168] M. Sousa, P. Vieira, M. P. Queluz, and A. Rodrigues, "Towards the use of unsupervised causal learning in wireless networks operation," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 9, 2023, Art. no. 101764.

[169] C. K. Thomas, C. Chaccour, W. Saad, M. Debbah, and C. S. Hong, "Causal reasoning: Charting a revolutionary course for next-generation ai-native wireless networks," *IEEE Veh. Technol. Mag.*, vol. 19, no. 1, pp. 16–31, Mar. 2024.

[170] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, "Less data, more knowledge: Building next generation semantic communication networks," *IEEE Commun. Surveys Tuts.*, vol. 27, no. 1, pp. 37–76, 2nd Quart., 2025.

[171] D. Harutyunyan, R. Behravesh, and N. Slamnik-Kriještorac, "Cost-efficient placement and scaling of 5G core network and MEC-enabled application VNFs," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manag. (IM)*, 2021, pp. 241–249.

[172] B. Han et al., "Admission and congestion control for 5G network slicing," in *Proc. IEEE Conf. Stand. Commun. Netw. (CSCN)*, 2018, pp. 1–6.

[173] M. R. Raza, C. Natalino, P. Öhlen, L. Wosinska, and P. Monti, "A slice admission policy based on reinforcement learning for a 5G flexible RAN," in *Proc. IEEE Eur. Conf. Opt. Commun. (ECOC)*, 2018, pp. 1–3.

[174] C. Natalino, M. R. Raza, A. Rostami, P. Öhlen, L. Wosinska, and P. Monti, "Machine learning aided orchestration in multi-tenant networks," in *Proc. IEEE Photon. Soc. Summer Topical Meeting Ser. (SUM)*, 2018, pp. 125–126.

[175] Y. Cai, P. Cheng, Z. Chen, M. Ding, B. Vucetic, and Y. Li, "Deep reinforcement learning for online resource allocation in network slicing," *IEEE Trans. Mobile Comput.*, vol. 23, no. 6, pp. 7099–7116, Jun. 2024.

[176] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for eMBB and URLLC coexistence in 5G and beyond: A deep reinforcement learning based approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4585–4600, Jul. 2021.

[177] R. M. Sohaib, S. Tariq Shah, and P. Yadav, "Towards resilient 6G O-RAN: An energy-efficient URLLC resource allocation framework," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 7701–7714, 2024.

[178] C. Fiandrino, L. Bonati, S. D'Oro, M. Polese, T. Melodia, and J. Widmer, "EXPLORA: AI/ML EXPLainability for the open RAN," *Proc. ACM Netw.*, 2023, pp. 1–26.

[179] G. Sun, "Autonomous resource provisioning and resource customization for mixed traffics in virtualized radio access network," *IEEE Syst. J.*, vol. 13, no. 3, pp. 2454–2465, Sep. 2019.

[180] J. A. Hurtado Sánchez, K. Casilimas, and O. M. Caicedo Rendon, "Deep reinforcement learning for resource management on network slicing: A survey," *Sensors*, vol. 22, no. 8, p. 3031, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/8/3031

[181] F. Rezazadeh, H. Chergui, and J. Mangues-Bafalluy, "Explanation-guided deep reinforcement learning for trustworthy 6G RAN slicing," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2023, pp. 1026–1031.

[182] T. Senevirathna, V. H. La, S. Marchal, B. Siniarski, M. Liyanage, and S. Wang, "A survey on XAI for 5G and beyond security: Technical aspects, challenges and research directions," *IEEE Commun. Surveys Tuts.*, early access, Aug. 2, 2024, doi: 10.1109/COMST.2024.3437248.

[183] H. Sun et al., "Advancing 6G: Survey for explainable AI on communications and network slicing," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 1372–1412, 2025.

[184] M. Zou, A. Perrusquía, and W. Guo, "Explaining data-driven control in autonomous systems: A reinforcement learning case study," in *Proc. 10th Int. Conf. Control Decis. Inf. Technol. (CoDIT)*, 2024, pp. 73–78.

[185] J. Angjo, I. Shayea, M. Ergen, H. Mohamad, A. Alhammadi, and Y. I. Daradkeh, "Handover management of drones in future mobile networks: 5G technologies," *IEEE Access*, vol. 9, pp. 12803–12823, 2021.

[186] S. Luo et al., "Blockchain-based task offloading in drone-aided mobile edge computing," *IEEE Netw.*, vol. 35, no. 1, pp. 124–129, Jan./Feb. 2021.

[187] Y. Wang and D. Boyle, "Constrained reinforcement learning using distributional representation for trustworthy quadrotor UAV tracking control," *IEEE Trans. Autom. Sci. Eng.*, vol. 27, pp. 5877–5894, 2024.

[188] T. Xu, Y. Liang, and G. Lan, "CRPO: A new approach for safe reinforcement learning with convergence guarantee," in *Proc. 38th Int. Conf. Mach. Learn.*, Jul. 2021, pp. 11480–11491.

[189] J. Xiao et al., "A deep reinforcement learning based distributed multi-UAV dynamic area coverage algorithm for complex environment," *Neurocomputing*, vol. 595, Aug. 2024, Art. no. 127904.

[190] Z. Fan, Y. Lin, Y. Ai, and H. Xu, "Adaptive task migration strategy with delay risk control and reinforcement learning for emergency monitoring," *Sci. Rep.*, vol. 14, no. 1, 2024, Art. no. 17606.

[191] Z. Ma, Z. Zhang, Z. Hao, and Q. Guo, "Energy-efficient mechanism of task offloading and resource allocation for hierarchical MEC in UAV-assisted mmWave IABN," *Expert Syst. Appl.*, vol. 267, Apr. 2025, Art. no. 126145. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417424030124

[192] M. Hosseinzadeh et al., "A novel *Q*-learning-based secure routing scheme with a robust defensive system against wormhole attacks in flying ad hoc networks," *Veh. Commun.*, vol. 49, Oct. 2024, Art. no. 100826.

[193] Y. Wang and D. Boyle, "Constrained reinforcement learning using distributional representation for trustworthy quadrotor UAV tracking control," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 5877–5894, 2025.

[194] D. Rizvi and D. Boyle, "Multi-agent reinforcement learning with action masking for UAV-enabled mobile communications," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 3, pp. 117–132, 2025.

[195] Z. Huang et al., “A soft actor–critic reinforcement learning approach for over the air active beamforming with reconfigurable intelligent surface,” *Phys. Commun.*, vol. 66, Oct. 2024, Art. no. 102474.
[196] N. Potteiger and X. Koutsoukos, “Safeguarding autonomous UAV navigation: Agent design using evolving Behavior trees,” in *Proc. IEEE Int. Syst. Conf. (SysCon)*, 2024, pp. 1–8.
[197] T. Hickling, N. Aouf, and P. Spencer, “Robust adversarial attacks detection based on explainable deep reinforcement learning for UAV guidance and planning,” *IEEE Trans. Intell. Veh.*, vol. 8, no. 4, pp. 380–383, Oct. 2023.
[198] E. Çetin, C. Barrado, and E. Pastor, “Explainability of deep reinforcement learning method with drones,” in *Proc. IEEE/AIAA 42nd Digit. Avionics Syst. Conf. (DASC)*, 2023, pp. 1–9.
[199] L. He, N. Aouf, and B. Song, “Explainable deep reinforcement learning for UAV autonomous path planning,” *Aerosp. Sci. Technol.*, vol. 118, Nov. 2021, Art. no. 107052.
[200] C. Wang, L. Wu, C. Yan, Z. Wang, H. Long, and C. Yu, “Coactive design of explainable agent-based task planning and deep reinforcement learning for human-UAVs teamwork,” *Chin. J. Aeronaut.*, vol. 33, no. 11, pp. 2930–2945, 2020.
[201] Y. Wang, Q. Qian, and D. Boyle, “Probabilistic constrained reinforcement learning with formal interpretability,” in *Proc. 41st Int. Conf. Mach. Learn.*, Jul. 2024, pp. 51303–51327.
[202] S. Wachter, B. Mittelstadt, and C. Russell, “Counterfactual explanations without opening the black box: Automated decisions and the GDPR,” *Havard J. Law Technol.*, vol. 31, no. 2, pp. 1–47, 2018.
[203] S. Mahboob, J. Dai, A. Soysal, and L. Liu, “Transforming future 6G networks via O-RAN-empowered NTNs,” *IEEE Commun. Mag.*, vol. 63, no. 3, pp. 76–82, Mar. 2025.
[204] Y. Tang and W. Guo, “Automatic retrieval-augmented generation of 6G network specifications for use cases,” *IEEE Commun. Mag.*, vol. 63, no. 4, pp. 95–102, Apr. 2025.