Constraint-Driven Intent-Based Networking for Interdomain Decision Support

Arthur V. C. Camargo*, Lisandro Z. Granville*, Leandro M. Bertholdo*, Renan P. Barreto†,
*Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil
†Universidade Federal do Rio Grande (FURG), Rio Grande, Brazil
Email: avccamargo@inf.ufrgs.br

***Abstract*—Interdomain routing is a complex and error-prone task, as Autonomous Systems must translate high-level intents into BGP configurations under partial visibility and distributed governance. While IBN is typically conceived for intradomain, automation-centric settings, its assumptions of authoritative control and closed-loop operation do not hold in environments such as Internet Exchange Points (IXPs). This paper argues that interdomain intent handling should be framed as a decision-support problem rather than full automation. We propose an LLM-assisted, human-in-the-loop architecture that supports intent interpretation, contextual reasoning, and pre-deployment impact assessment using external, non-authoritative routing evidence sources. A proof-of-concept prototype is evaluated in an operational IXP setting using reactive-intent scenarios, showing that LLM-assisted reasoning can improve operator awareness and reduce the risk of misconfigurations. The results highlight the feasibility and limitations of adapting IBN to multi-AS environments, rather than a production-ready solution.**

***Index Terms*—Intent-Based Networking, BGP, Large Language Models, Natural Language Processing, Network Automation**

I. Introduction

Interdomain routing remains a complex and error-prone task, rooted in the policy-driven nature of BGP, which allows each Autonomous System (AS) to define and enforce its own routing policies [1]. High-level operational goals—such as avoiding route leaks or steering traffic through preferred peers—must be manually translated into low-level device configurations, creating a gap between operator intent and actual network behavior. This misalignment increases the risk of misconfiguration, which can lead to severe incidents such as route leaks and large-scale outages [2].

The widespread adoption of Internet eXchange Points (IXPs) has intensified interdomain peering complexity by enabling dense connectivity among heterogeneous ASes. While IXPs improve reachability and performance, they also introduce operational challenges, including diverse route-server implementations, non-standardized BGP community semantics, and intricate peering policies [3]–[5]. As a result, operators increasingly need decision support to map high-level operational intents to safe, context-aware BGP actions—an area where recent LLM-based efforts focus on configuration synthesis but fall short of an intent-centric IBN lifecycle.

Although LLM-based approaches are being explored to automatically synthesize interdomain BGP policies into device configurations [6], intent-based networking (IBN) concepts remain largely unexplored in this context. This is partly due to the difficulty of establishing controlled environments for IBN in interdomain systems, given their distributed governance, heterogeneity, and limited observability [1]. Existing LLM-based solutions for intent translation and policy reasoning have shown promising results, but remain constrained to intradomain or service-specific scenarios [7], [8].

To address this gap, this paper presents a *constraint-driven reference architecture* for LLM-assisted, human-in-the-loop interdomain intent handling at IXPs as an initial step toward interdomain IBN environments, which provide a practical operational setting with observable interdomain evidence and shared metadata for decision support under distributed governance [4]. The framework targets small and medium-sized ASes, aiming to operationalize core IBN principles—intent translation, reasoning, and validation—while preserving human oversight and explainability. Rather than proposing full automation, this work is positioned as a *design study* that explores feasibility, architectural requirements, and validation strategies for trustworthy LLM-assisted interdomain operations. This design study makes three contributions: (i) we characterize the interdomain constraints that break core intradomain IBN assumptions in IXP operations; (ii) we derive a constraint-driven, human-in-the-loop architecture for evidence-grounded interdomain intent handling (decision support rather than closed-loop enforcement); and (iii) we validate feasibility through an operational prototype and a case-study evaluation with operator-derived reactive intent scenarios.

The remainder of this paper is organized as follows. Section II reviews related work; Section III presents the interdomain constraints and derived requirements; Sections IV and V describe the architecture and prototype; Section VI reports the case-study validation; and Section VII concludes the paper.

II. Related Work

IBN has evolved from policy-based and model-driven management paradigms, aiming to bridge the gap between operator objectives and network configuration through automation and assurance loops [9]. Most IBN frameworks were conceived for *single-domain and fully observable environments*, typically data centers and enterprise networks [12], where intents can be expressed as formal policies and translated deterministically into device-level actions. These systems rely on continuous verification and an automonous closed-loop control (inner-loop) to maintain compliance with the desired state. The

TABLE I
COMPARATIVE SUMMARY OF INTENT-BASED NETWORKING REFERENCE MODELS AND SYSTEMS

Work	Domain Scope	Intent Processing	Assurance Mechanism	Human Interaction	Explainability	Distinctive Contribution
Conceptual Reference Model						
RFC 9315 [9]	Reference Model	Formalized intent lifecycle model	Conceptual intent assurance and control loops	Minimal (articulate the intent and assess outcome)	Traceability mentioned, not specified	Defines intent lifecycle and assurance concepts; does not specify an implementable architecture, operational mechanisms, or automation details.
NMRG Draft 2025 [10]	Intradomain (5G / Cloud)	Generative-AI-based intent reasoning and model specialization	Model-centric feedback loops (telemetry-driven adaptation)	Limited (automation-oriented)	Conceptual discussion only	Generative-AI-enabled IBN with model specialization and lifecycle management; assumes centralized governance and closed-loop operation, without addressing interdomain constraints.
Research Frameworks and Prototypes						
Tu et al., 2024 [11]	Intradomain (Enterprise)	LLM-based translation into configurations	Static validation before deployment	Optional manual review	Natural-language explanation of configs	Demonstrates LLM use for intent parsing and configuration generation in controlled networks.
Ficzere et al., 2025 [8]	Intradomain (Datacenter)	Ontology/NLP-based intent translation	Telemetry-driven feedback verification	Limited (semi-automated)	Partial explainability via semantic mapping	Analyzes intent translation research gaps; assumes homogeneous, centralized environments.
Mendoza et al., 2025 [6]	Interdomain (BGP / Multi-vendor)	LLM-based translation of high-level BGP policies into device configurations	Simulation-based verification (syntactic, semantic, reliability metrics)	Offline (no operator interaction)	Partial explainability through benchmark metrics, not human-readable explanations	Introduces PeeringLLM-Bench, a benchmark for evaluating LLMs in BGP configuration synthesis across vendors; focuses on performance assessment rather than intent reasoning or assurance.
This Work	**Interdomain (Multi-ASes / IXPs)**	**LLM-based, context-aware reasoning using external data (e.g., RPKI, Looking Glass)**	**Pre-deployment impact assessment and partial intent assurance (no closed loop)**	**Human-in-the-loop decision support**	**Explainable configuration output with audit logs and validation evidence**	**Operationalizes RFC 9315 concepts for interdomain by repositioning IBN as AI-assisted, human-in-the-loop decision support under distributed governance and limited visibility.**

human interaction is restricted to specify the intent and assess the outcome (outer-loop).

Building on this line of work, subsequent studies extended IBN principles to *virtualized and software-defined infrastructures*, enabling automated service orchestration, QoS optimization, and fault remediation under centralized control [13], [14]. Nonetheless, these approaches assume homogeneous data models, a single administrative authority, and mostly make use of specific description languages for intents (e.g., SDL).

Recent works have explored cognitive and AI-driven techniques as tools for intent translation, policy validation, or configuration assistance, primarily in controlled intradomain environments [6], [8], [10], [11], [15]. However, applying IBN to *distributed or interdomain environments* remains an open challenge, as intent translation and assurance must operate under limited network visibility, heterogeneous governance, and incomplete trust [16].

From the literature, we distill a set of interdomain-specific constraints—limited authority, partial visibility, heterogeneous semantics, and incomplete trust—that fundamentally challenge the assumptions underlying existing IBN approaches. As summarized in Table I, existing IBN systems—largely designed for intradomain environments—implicitly assume environmental homogeneity and centralized authority. These assumptions do not hold in interdomain environment, requiring IBN to be positioned as a decision-support paradigm under partial visibility and distributed governance.

III. DESIGN OVERVIEW: INTERDOMAIN CONSTRAINTS THAT BREAK TRADITIONAL IBN ASSUMPTIONS

Unlike intradomain environments, interdomain intent handling violates several core assumptions implicitly adopted by most IBN systems designed for a single administrative domain. In particular, interdomain environments are characterized by conflicting and non-authoritative data sources (e.g., Looking Glass services), partial and fragmented views of the relevant network state (e.g., routing table vantage points), and the absence of a shared semantic authority (e.g., per-AS interpretation of BGP community attributes). As a result, autonomous closure of the IBN inner control loop becomes infeasible, and ambiguity resolution must rely on explicit safeguards such as deterministic validation, evidence cross-checking, and, when needed, human oversight.

Importantly, this does not invalidate the IBN intent lifecycle itself, but alters how intent assurance and enforcement can be realized in interdomain settings.

With a particular emphasis on interdomain intent handling, we use IXPs as a representative operational setting in which these constraints become observable through public control-plane evidence and operator workflows. To structure the discussion, we characterize these challenges as a set of interdomain-specific constraints with direct design consequences on the intent pipeline, and map them to a minimal set of functional requirements (FR1–FR4) capturing the capabilities needed under distributed governance and partial visibility.

FR1. Intent Interpretation without Semantic Authority

The system must interpret operators' objectives, rather than merely translate predefined intents. It should accept structured and unstructured inputs (e.g., natural language and visual artifacts), identify relevant entities (e.g., ASNs, prefixes, and BGP attributes), resolve semantic gaps, and confirm ambiguous interpretations with the operator.

Interdomain constraint: In the absence of a shared semantic authority or enforceable intent language, intent understanding

cannot be reduced by parsing only; it requires context-aware reasoning and explicit disambiguation with the operator.

FR2. Evidence-Grounded Intent Translation

Given an interpreted intent, the system must translate it into a concrete, feasible, and operationally valid set of actions for the current operational state. Translation must correctly scope the action (e.g., peer-, AS-, or IXP-specific) and check feasibility against applicable constraints.

To this end, the system must ground translation on external and near-real-time evidence—such as BGP collectors, peering databases, IXP metadata, session-state signals, and RPKI validators—to determine which actions are realizable in the current context and to parameterize the resulting configuration.

Interdomain constraint: Interdomain environments provide no authoritative or complete view of network state; therefore, intent translation must be grounded in correlated, non-authoritative evidence to derive feasible actions

FR3. Conflict Detection and Impact Assessment

Translating high-level intents into actions requires mapping abstract objectives to specific configuration elements, such as BGP community values, local preference rules, AS-path adjustments, or prefix filters. This process may generate internal and external contradictions; for instance, an intent to "prefer routes through AS64500" may conflict with a local policies "rejects all routes containing AS64500 in the path," or a publicized peer policy that prohibits such behavior. To prevent such inconsistencies, the system must perform *logical pre-deployment validation* to detect internal conflicts avoiding unintented behaviours before any configuration is suggested.

Rather than providing guarantees, pre-deployment impact assessment aims to reduce configuration risk by identifying consequences and inconsistencies, acknowledging that deterministic assurance is unattainable under partial visibility.

Interdomain constraint: Interdomain conflicts stem from the absence of centralized governance and are often expressed implicitly through BGP attributes or loose intent rules [17].

FR4. Human-Centered Validation and Impact Awareness

This requirement focuses on validating the interpreted intent with the operator, informed by the consequences of possible actions and explicit impact awareness, rather than on autonomous intent assurance. Given the impossibility of autonomous assurance at the interdomain scope, assessment results must be presented to the operator in a structured and explainable manner. To this end, the system must expose assessment outcomes as explicit information, including: (i) the data sources consulted, (ii) inferred assumptions, (iii) confidence or uncertainty indicators, and (iv) identified risk categories. This ensures that operators receive sufficient contextual information to review, accept, or reject recommendations affecting independent administrative domains.

Interdomain constraint: Autonomous intent assurance is infeasible in interdomain environments due to the lack of authoritative control, complete observability, and enforceable semantics. Consequently, intent handling must rely on human-centered validation supported by evidence-based assessments.

Table II makes explicit the causal relationship between interdomain constraints, the derived functional requirements, and their architectural realization. The inclusion of an interdomain rationale highlights why these design choices differ fundamentally from intradomain IBN assumptions and reinforces the decision-support positioning of the proposed design.

IV. System Architecture

This section instantiates the constraint-driven requirements into a modular architecture for interdomain intent handling. Figure 1 provides a high-level view of the main components and their interactions: intents are iteratively refined with evidence, candidate actions are checked before deployment, and recommendations are returned with an auditable trace for operator approval.

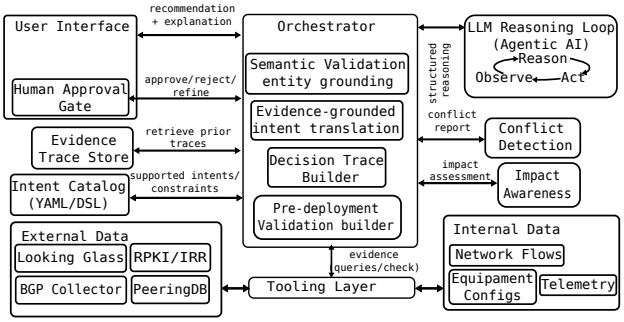

Fig. 1. High-level architecture for evidence-grounded interdomain intent handling (human approval gate, pre-deployment checks, and decision trace).

Architecturally, the Orchestrator coordinates the interaction between (i) the LLM Reasoning Loop, which refines intent hypotheses and proposes candidate actions, (ii) the Tool Interaction Module, which retrieves external evidence and executes pre-deployment checks, and (iii) the Evidence and Trace Store, which persists consulted sources, assumptions, and rationale for explanation and auditing. The User Interface exposes these artifacts to support clarification and approval. The architecture decomposes into five main components: User Interface, Orchestrator, LLM Reasoning Loop, Tool Interaction Module, and Evidence and Decision Trace Store. The following subsections describe the workflow and core components.

A. System Workflow

Given a user intent such as "`prefer routes via peer AS64500 at IXP` x", the system ① receives the intent through the UI. ② The Orchestrator performs semantic validation by forwarding the prompt to the LLM Reasoning Loop, which interprets entities and scope (AS64500, IXP x). ③ After reasoning, the agent requests evidence about sessions (Looking Glass tool), policies (via the store or by asking the user), and observed routing (Looking Glass tool). ④ The LLM agent then proposes candidate changes (e.g., localpref, communities, prepend) and return the result to the Orchestrator. ⑤ The Orchestrator triggers pre-deployment validation to flag

TABLE II
CONSTRAINT-DRIVEN DESIGN FOR INTERDOMAIN INTENT-BASED NETWORKING

Interdomain Constraint	Functional Requirement	Architectural Consequence	Interdomain Rationale
Heterogeneous operator intents, vocabularies, and semantics	FR1: Intent Interpretation without Semantic Authority	Conversational UI supporting natural language, structured (DLS) and unestructured data (pictures, snapshots, documents).	Operators express intents using heterogeneous, non-standardized vocabularies across AS boundaries, with no shared semantic authority or enforceable intent language.
Partial and non-authoritative visibility of routing state	FR2: Evidence-Grounded Intent Translation	Dedicated data collection module integrating external, non-authoritative data sources	No single entity has a complete or authoritative view of the interdomain routing state.
Independent administration and policy autonomy	FR3: Conflict Detection and Impact Assessment	Iterative orchestration with early detection of out-of-scope or inconsistent intents	Routing policies may conflict across independently managed administrative domains.
Lack of authoritative control, global observability, and enforceable assurance	FR4: Human-Centered Intent Validation and Impact Awareness	Pre-deployment impact assessment with evidence aggregation, explicit risk and uncertainty characterization, and human-in-the-loop decision support	Autonomous assurance cannot be achieved under partial visibility and distributed governance; safe interdomain operation requires evidence-based assessment and explicit human approval across administrative boundaries.

conflicts with the local baseline and identify plausible side effects. ⑥ Finally, it returns an explainable recommendation with evidence and uncertainty indicators for operator approval.

B. Core Components

User Interface (UI): Captures intents in natural language or a DSL and supports clarification and approval steps. It also accepts auxiliary artifacts (e.g., screenshots, diagrams, and plots) to better convey the operator's intent. The UI presents each recommendation with a natural-language explanation of the intended outcome.

Human Approval Gate: The system enforces an explicit approval step to prevent automatic changes. The operator can accept, reject, or edit the action before any deployment.

Orchestrator: Is the central control mechanism of the architecture and separates the LLM reasoning from the deterministic system. Although the LLM is responsible for triggering the steps, the Orchestrator constraints when and how each action will be executed, including access to the intent catalog, external tools and validation modules. By sequencing these interactions, the Orchestrator preserves reproducibility.

LLM Reasoning Loop: Based on the *reAct* synergy [18], the reasoning loop consists of three steps, observation, reasoning and action. Observation step gather the relevant information and input the LLM. Reasoning implements a chain-of-though prompting where the model reflect upon the received data and the requests being made in a LLM-to-LLM interaction. Lastly, Action prompts the LLM to interact with the system by replying to the user, asking clarification when uncertain and utilizing the internal and external tools.

Tool Interaction Module: The system relies on internal and external to have up to date data. For that reason, the tool interaction module queries external and internal sources (e.g., collectors, Looking Glass, RPKI/IRR, PeeringDB/IXP metadata) and runs pre-deployment checks. This module must interact with APIs and external software interfaces which requires error handling, rate limits and custom checks.

Pre-deployment Validation checks Two validation pre-deployment checks are required by FR3 and FR4: *(i) conflict detection* which tests candidate actions against the existing policy baseline and known constraints; and *(ii) impact/risk assessment*, which derives plausible outcomes and uncertainty indicators from non-authoritative evidence. Differently from the other sources those tools require a more active type of engagement. Both tools output structured artifacts attached to the decision trace.

Evidence and Trace Store: Persists consulted evidence, reasoning assumptions, and operator decisions to support explainability and auditability. Implemented as a database, it maintains a long-term record of key artifacts, including proposed/deployed configurations, operator approvals and edits, and external/internal tool queries.

Intent Catalog: Defines the supported intents and provides structured guidance for their interpretation. For each intent, it specifies the relevant entities, required evidence sources, applicable constraints, and possible actions. The catalog provides domain-specific knowledge and LLM guardrails, enabling greater control and reducing hallucinations.

V. PROTOTYPING

A prototype of the proposed framework was developed to explore the feasibility and design trade-offs of using LLM-assisted reasoning as a decision-support mechanism for interdomain IBN in IXP environments, rather than as a fully automated configuration system. It implements the main components described in Section IV, providing a concrete instantiation of the proposed architecture to examine how reasoning, orchestration, and contextual data interact under interdomain constraints. However, the pre-validation modules were abstracted as mock components, as implementing authoritative conflict detection and impact modeling requires domain-specific data and operational access that fall outside of the scope of this design study.

The interface is a web-based, chat-style UI where operators state intents in text and review system recommendations as part of a human-in-the-loop decision process.

The LLM serves as the reasoning component, responsible for interpreting inputs, generating candidate actions, and proposing the next steps in the execution flow, including requests for external and internal data sources. These requests are mediated by the Orchestrator, which validates them, enforces the prescribed workflow, and executes only authorized interactions with tools and system modules. The prototype uses *Gemini-Flash-2.5* [19]—chosen after exploratory qualitative

tests of responsiveness and reasoning stability—and integrates a set of external data sources utilized in interdomain operations, including Internet Routing Registry (IRR) databases, IXP Looking glasses, and Route Servers dumps. These sources are accessed through API wrappers and are treated as non-authoritative evidence to support reasoning All retrieved data are presented to the operator alongside contextual explanations of their meaning and relevance.

VI. Design Validation Through Case Studies

To explore the proposed architecture under realistic operational conditions, we conduct a qualitative case study grounded in operator-reported IXP operational scenarios collected from an operator forum hosted on Telegram (2018 to mid-2025) [1] We distill these reports into six representative *reactive intent classes* capturing common interdomain challenges, and manually instantiate 24 concrete scenarios used in our evaluation. The intent classes are summarized in Table III.

TABLE III
Reactive intent classes used as IXP case studies.

Intent class (reactive)	**Outcome / evidence focus**
P1: Prepend anomalies	Consistency w/ policy; configured vs. observed paths; community-based actions.
P2: Prefix rejection	Root-cause of rejection; IRR/PeeringDB; RPKI.
P3: Propagation issues	Missing reachability; advertised vs. received/accepted/filtered.
P4: Community semantics	Meaning/effect of tags; community registry; received tags from IXPs.
P5: Session stability	Instability causes; session state; IXP IGP reachability.
P6: Routing data lookup	Routing context and historical data; AS paths; AS-Cone; AS-relationship signals.

A. Evaluation Methodology

We evaluate the proposed framework using the 24 manually instantiated *reactive-intent* scenarios spanning the six intent classes (P1–P6) in Table III. Two complementary evaluation approaches were employed: (*i*) an expert operator walkthrough; and (*ii*) an *LLM-as-Judge* analysis.

In the expert walkthrough, we conducted a qualitative assessment involving four network operators with operational experience in interdomain environments. Each operator interacted with the prototype across the full set of scenarios and subsequently evaluated the system's outputs using a 5-point Likert scale, reflecting their perceived correctness of the responses and clarity of explanations. The assessment aims to capture the subjective perceptions of the system's decision-support value from a practitioner perspective.

As stochastic models, LLMs can exhibit variability in open-ended reasoning tasks [20]. To complement the expert walkthrough, we employed an auxiliary *LLM-as-Judge* analysis to explore patterns in the system's responses across 5 generations for each scenario. In this analysis, the prototype's outputs are compared against expert solutions using another LLM (*gpt-4.1-mini*) as a semantic evaluator, focusing on high-level alignment of reasoning and explanatory content rather than exact correctness. Similar to the expert walkthrough, the LLM-as-Judge evaluates the system response using the 5-point Likert scale for comparability. Additionally, it receives a semantic guideline of what would be a conceptually appropriate response and judges the chat excerpt based on that.

[1]The anonymized operator reports, the distilled intent classes, and the 24 implemented scenarios and evaluation forms are available at: *https://github.com/ArthurCamargo/ibn-interdomain-decision-support*.

B. Results

As shown in Table IV, both human operators and the *LLM-as-Judge* consistently assigned high scores across all intent classes. While these values do not constitute performance measures, they indicate that the system's responses were generally perceived as coherent, understandable, and operationally adequate. The limited variability across scenarios suggests that the architecture behaves in a stable manner under diverse reactive-intent contexts. Importantly, the close alignment between human and automated assessments supports the use of the *LLM-as-Judge* as a complementary exploratory lens, enabling automated analysis across scenarios.

TABLE IV
Human vs. LLM-as-Judge evaluations on a 5-point Likert scale.

Intent class	Human Avg. Score	Human Std. Dev.	LLM-as-Judge Avg. Score	LLM-as-Judge Std. Deviation
P1	3.87	1.15	4.50	0.57
P2	4.50	0.89	4.25	1.50
P3	4.56	0.51	4.75	0.50
P4	4.50	0.82	5.00	0.00
P5	4.62	0.50	4.25	1.50
P6	4.50	0.63	4.25	1.50
Averages	4.42	—-	4.50	—-

C. Discussion and Insights

Our evaluation revealed several challenges in the system's final reasoning stage. These include configuration redundancy, where multiple mechanisms (e.g., AS-path and BGP communities) were simultaneously used to express equivalent policies; community misinterpretation, caused by wrongly reasoning over numeric IXP-specific community values; syntactic and semantic inconsistencies in generated configs (e.g., '`set path`' vs. '`set as-path`'); and parameter inference errors.

At the same time, the system exhibited robust behavior across the majority of intent classes, including P2 through P6, which encompass filtered prefix detection, propagation analysis, community interpretation, session stability, and data lookup tasks. These scenarios allow the LLM to correlate routing data, registry records, and community semantics to derive plausible interpretations. In contrast, P1 scenarios introduced higher levels of ambiguity and relied more heavily on implicit operator assumptions and undocumented configuration practices, which reduced the availability of reliable contextual grounding. This suggests that differences in system behavior across categories are primarily driven by the structure

and observability of the underlying evidence, rather than by isolated limitations of the reasoning process.

Taken together, these observations indicate that the effectiveness of LLM-based interdomain reasoning is coupled to the quality and structure of the available contextual evidence. The system performs best when operating over explicit, externally verifiable information, but remains vulnerable to inference errors in cases requiring implicit knowledge or undocumented operator conventions. This is an important nuance in interdomain routing scenarios where ASes have their own BGP community conventions and utilized an heterogeneous set of systems (e.g. Cisco, Huawei, Mikrotik ...) that can have distinct BGP implementations. This reinforces the importance of augmenting LLM reasoning with domain-grounded retrieval mechanisms, such as structured community databases and IXP-specific metadata, as well as integration between orchestration and validation components. More broadly, the results suggest that LLMs are well-suited for decision-support roles in interdomain IBN, provided that human oversight and architectural safeguards remain central.

VII. Conclusion and Future Work

This paper proposed a constraint-driven reference architecture to augment IBN systems with LLM-driven reasoning, instantiated as a proof-of-concept prototype that enables interdomain intent handling and decision support for ASes peering at IXPs. Unlike intradomain IBN approaches that assume authoritative control and closed-loop enforcement, the proposed system uses conversational reasoning to interpret high-level intents in natural language and translate them into evidence-grounded, operator-approved routing recommendations. By combining LLM-based reasoning with structured IXP evidence— such as AS-path observations, peer/session availability, prefix visibility, and BGP community semantics—the framework supports explainable and context-aware intent handling across multi-AS environments.

Future work will focus on advancing the prototype from exploratory reasoning to more robust *assisted* intent translation under explicit human approval, as well as exploring privacy-preserving techniques and data anonymization strategies. Toward operational trust and adoption, we envision two complementary safeguards: (i) evidence-grounded generation, where routing semantics are retrieved from authoritative sources rather than inferred from free text; (ii) deterministic guardrails enforcing vendor syntax, conflict rules, and safety constraints.

References

[1] N. Feamster, H. Balakrishnan, and J. Rexford, "Some foundational problems in interdomain routing," in *Proceedings of Third Workshop on Hot Topics in Networks (HotNets-III)*, 2004, pp. 41–46.

[2] H. Gao, N. Li, and Y. Xie, "Leak-detector: An improved route leak detection method," in *2024 IEEE International Conference on Dependability in Sensor, Cloud & Big Data Systems & Applications (DependSys)*. IEEE, 2024, pp. 1–6.

[3] V. Giotsas, G. Nomikos, V. Kotronis, P. Sermpezis, P. Gigis, L. Manassakis, C. Dietzel, S. Konstantaras, and X. Dimitropoulos, "O peer, where art thou? uncovering remote peering interconnections at ixps," *IEEE/ACM Transactions on Networking*, vol. 29, no. 1, pp. 1–16, 2021.

[4] Y. Dabone, T. F. Ouedraogo, and O. Sie, "Measures to enhance the characterization of internet exchange points," in *2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, 2022, pp. 169–173.

[5] F. Mazzola, P. Marcos, and M. Barcellos, "Light, camera, actions: characterizing the usage of ixps' action bgp communities," in *Proceedings of the 18th International Conference on Emerging Networking EXperiments and Technologies*, ser. CoNEXT '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 196–203. [Online]. Available: https://doi.org/10.1145/3555050.3569143

[6] J. R. Mendoza and R. Ocampo, "Peeringllm-bench: Evaluating llms for bgp configuration tasks," in *Proceedings of the 20th Asian Internet Engineering Conference*, ser. AINTEC '25. New York, NY, USA: Association for Computing Machinery, 2025, p. 78–86. [Online]. Available: https://doi.org/10.1145/3763400.3763451

[7] E.-D. Jeong, H.-G. Kim, S. Nam, J.-H. Yoo, and J. W.-K. Hong, "S-witch: Switch configuration assistant with llm and prompt engineering," in *NOMS 2024-2024 IEEE Network Operations and Management Symposium*, 2024, pp. 1–7.

[8] D. Ficzere, G. Hollósi, and P. Varga, "Beyond Intent Translation: Research Gaps in the Application of Generative AI for Intent-Based Networking," in *NOMS 2025-2025 IEEE Network Operations and Management Symposium*, May 2025, pp. 1–7.

[9] A. Clemm, L. Ciavaglia, L. Z. Granville, and J. Tantsura, "Intent-Based Networking - Concepts and Definitions," RFC 9315, Oct 2022. [Online]. Available: https://www.rfc-editor.org/info/rfc9315

[10] P. Cassara', A. Gotta, G. Fioccola, A. Artigiani, R. Burrai, E. Kolaj, M. Martalo', and V. Pilloni, "Generative AI for Intent-Based Networking," Internet Engineering Task Force, Internet-Draft draft-cgfabk-nmrg-ibn-generative-ai-01, Oct. 2025, work in Progress. [Online]. Available: https://datatracker.ietf.org/doc/draft-cgfabk-nmrg-ibn-generative-ai/01/

[11] N. Tu, S. Nam, and J. W.-K. Hong, "Intent-based network configuration using large language models," *International Journal of Network Management*, vol. 35, no. 1, p. e2313, 2025.

[12] D. M. Manias, A. Chouman, and A. Shami, "Towards intent-based network management: Large language models for intent extraction in 5g core networks," in *2024 20th International Conference on the Design of Reliable Communication Networks (DRCN)*. IEEE, 2024, pp. 1–6.

[13] W. Cerroni, C. Buratti, S. Cerboni, G. Davoli, C. Contoli, F. Foresta, F. Callegati, and R. Verdone, "Intent-based management and orchestration of heterogeneous openflow/iot sdn domains," in *2017 IEEE Conference on Network Softwarization (NetSoft)*, 2017, pp. 1–9.

[14] A. S. Jacobs *et al.*, "Refining network intents for self-driving networks," in *Proceedings of the ACM SIGCOMM Afternoon Workshop on Self-Driving Networks*. ACM, 2018, pp. 15–21.

[15] Y. Zhou, K. Hsieh, S. K. Mani, S. Kandula, and Z. Liu, "Meshagent: Enabling reliable network management with large language models," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 9, no. 3, Dec. 2025.

[16] K. Yao, D. Chen, J. P. Jeong, Q. Wu, C. Yang, L. M. Contreras, and G. Fioccola, "Use Cases and Practices for Intent-Based Networking," Internet Engineering Task Force, Internet-Draft draft-irtf-nmrg-ibn-usecases-02, Nov. 2025, work in Progress. [Online]. Available: https://datatracker.ietf.org/doc/draft-irtf-nmrg-ibn-usecases/02/

[17] M. N. Inc., "Radb: Routing assets database," https://www.radb.net/, 2025, accessed: 2025-11-10.

[18] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. R. Narasimhan, and Y. Cao, "React: Synergizing reasoning and acting in language models," in *The eleventh international conference on learning representations*, 2022.

[19] G. Team, "Gemini: A Family of Highly Capable Multimodal Models," May 2025. [Online]. Available: https://arxiv.org/abs/2312.11805

[20] L. Li, L. Sleem, G. Nichil, R. State *et al.*, "Exploring the impact of temperature on large language models: Hot or cold?" *Procedia Computer Science*, vol. 264, pp. 242–251, 2025.