ICANN84 Session Summary

Reducing Domain Name Abuse with AI

ccNSO & GAC Discussion | Dublin, Ireland

Executive Summary

This session explored how country code top-level domains (ccTLDs) can leverage artificial intelligence to combat domain name abuse, using the .nl registry's experience as a case study. The Netherlands' SIDN registry has successfully deployed machine learning systems to detect and reduce fake web shops, online impersonation, and malicious registrations. The discussion emphasized that AI is not just a technological solution but requires careful governance, human oversight, and continuous adaptation to remain effective.

Participants discussed the importance of keeping humans in the loop, with AI systems providing recommendations rather than automatic decisions. The session highlighted three key areas: technical implementation using open-source tools and continuous retraining, governance structures including privacy boards and compliance with the EU AI Act, and operational collaboration between registries and registrars. The .nl registry manages over 6 million domain names and has made substantial investments in data analytics expertise to strengthen their TLD against abuse.

Key outcomes included recognition that AI can serve both defensive and offensive purposes in the domain ecosystem. While registries use it to detect abuse, malicious actors also employ AI to create sophisticated attack networks. The discussion underscored the need for ccTLDs to develop principles aligned with RFC1591 responsibilities, establish expert oversight bodies, and share best practices internationally.

Key Discussion Outcomes

Technical Implementation

The .nl registry employs machine learning to detect patterns in large datasets, using both internal registration data and external sources. The system analyzes various parameters including domain characteristics, registrant information, DNS traffic patterns, and website content. SIDN has developed multiple specialized systems: RegCheck for suspicious registrations, fake web shop detection, and online impersonation identification through logo misuse detection.

Continuous retraining of AI models is essential because datasets become obsolete as abuse patterns evolve. The crawler creates snapshots of domain content, taking 1-2 weeks to complete a full scan of the zone. The system generates false positives, which is why human verification remains crucial. The support team handles cases flagged by AI, conducting manual reviews and reaching out to registrars, when necessary, before taking enforcement action.

Open-source software resources form the foundation of the technical implementation, with detailed information available through SIDN Labs blogs. The technology stack combines DNS traffic analysis, content checking through active crawling, and registration pattern analysis. External feeds like Netcraft complement internal detection systems, providing additional data points for comprehensive abuse identification across the domain lifecycle from registration through publication.

Questions were raised about scalability and costs when moving into larger gTLD zones. The consensus was that while costs increase with zone size, the strategic decision to invest in TLD strength through abuse reduction remains worthwhile. Market forces have also played a role, with commercial services like Netcraft now offering abuse detection feeds that complement registry efforts.

Governance and Compliance

The .nl registry operates under EU data protection regulations and has established a Privacy Board since 2014 to oversee AI considerations. This board's charter is being extended to include compliance with the EU AI Act, demonstrating proactive governance. The multi-faceted governance structure ensures that AI deployment aligns with legal requirements and ethical principles while maintaining operational effectiveness.

A critical principle emphasized throughout the discussion was keeping humans in the loop. AI systems produce recommendations with confidence levels, but final decisions always involve human adjudication. This approach balances automation efficiency with accuracy and fairness. The support team reviews flagged cases, validates AI outputs, and makes final enforcement decisions based on expert judgment combined with AI insights.

The governance framework requires board vision and strategic commitment from leadership. The registry board receives advice from experts who can evaluate AI system performance and ethical implications. Regular auditing of AI systems ensures they remain accurate and fair over time. This governance approach extends beyond technology to encompass stakeholder engagement, including discussions with government authorities, registrars, and other stakeholders about AI use cases and implications.

Use cases determine which data gets analyzed, with privacy considerations guiding data collection and processing decisions. The governance structure provides accountability while enabling innovation, ensuring AI deployments serve the public interest while respecting individual rights and maintaining the trust essential to domain registry operations.

Operational Collaboration

The relationship between the .nl registry and its registrars proved essential to successful abuse mitigation. In the Netherlands, a strong association of registrars facilitates collaboration and information sharing. When RegCheck flags suspicious registrations, experts reach out to registrars who work cooperatively to address potential abuse. This collaborative approach works well within the Dutch market structure.

Joint development initiatives demonstrate the value of inter-registry collaboration. The RegCheck system is being developed jointly by .nl and .be registries, allowing

both to benefit from shared expertise and resources. This model of collaboration could extend to other ccTLDs, enabling smaller registries to access sophisticated abuse detection capabilities without bearing full development costs independently.

Providing information rather than just taking enforcement action characterizes the .nl approach. When abuse is detected, the registry works with registrars to resolve issues rather than immediately suspending domains. This educational and collaborative stance builds stronger relationships and more sustainable abuse reduction. The registry also shares threat intelligence and patterns with the broader community through blogs and reports.

The discussion acknowledged that the specific market conditions in each ccTLD jurisdiction affect implementation approaches. While the Netherlands benefits from established registrar relationships and regulatory frameworks, other ccTLDs may need to adapt governance and operational models to their local contexts while maintaining the core principles of human oversight, transparency, and collaborative enforcement.

Concrete Example: Fake Webshop Detection

The fake webshop problem in .nl peaked between 2016 and 2018, when approximately 12,000 fraudulent online stores were identified and taken down in 2018 alone. These fake shops created networks of websites designed to lure people into providing payment information or personal data without delivering products. The scale of the problem threatened consumer confidence in .nl domains and required urgent action.

SIDN developed machine learning tools specifically to detect fake webshops by analyzing website features and behavioral patterns. The system examines various parameters including payment methods offered, contact information validity, domain age, content similarity to known fake shops, and technical infrastructure characteristics. The ML model was trained on labeled examples of legitimate and fraudulent websites to recognize distinguishing features.

The registry partnered with registrars and the Internet Systems Consortium (ISC) to remove identified fake shops systematically. This collaborative enforcement approach proved highly effective. By 2019, takedowns had dropped to 4,340, then to 481 in 2020, and by 2023 only 241 fake shops required removal. The dramatic reduction from 12,000 to 241 represents a 98% decrease in just five years.

Today, fake shops are virtually gone from the .nl zone, and a dashboard system is in active use at SIDN's anti-abuse desk for continuous monitoring. The success of this initiative demonstrates how targeted AI application combined with strong partnerships can effectively address specific abuse types. The market has also responded, with services like Netcraft now providing commercial feeds that complement the registry's internal detection capabilities.

Summary of Findings

The following table summarizes the key findings and approaches discussed during the session:

Category	Key Points
Technology & Methods	• Machine learning to detect patterns in large datasets • Open-source software resources • Continuous retraining as datasets become obsolete • DNS traffic analysis and content crawling (1-2 weeks per scan) • Multiple detection systems: RegCheck, fake webshop detection, online impersonation
Governance & Oversight	• Privacy Board established 2014, charter extended for EU AI Act compliance • Multi-faceted governance structure with board vision • Human-in-the-loop principle: AI provides recommendations, humans make decisions • Confidence level adjudication and regular auditing • Stakeholder engagement with government, registrars, and community
Operational Approach	• Strong collaboration with registrar association in Netherlands • Support team handles flagged cases and reaches out to registrars • Joint development with .be registry (RegCheck system) • Focus on information sharing rather than immediate enforcement • Issue tracking for false positives and system improvement
Results & Impact	• Fake webshops reduced from ~12,000 (2018) to 241 (2023) - 98% decrease • Online impersonation detection operational with BrandGuard service • Dashboard systems in active use at anti-abuse desk • Market adoption with commercial services like Netcraft offering feeds • Strategic investment in TLD strength serving 6+ million domains
Challenges & Considerations	• False positives requiring human verification • Labor-intensive expert review process • Scalability concerns for larger zones • AI used by malicious actors creating attack networks

Category	Key Points
	• Need for continuous investment in expertise and infrastructure

Emerging Topics for Future Discussion

AI Principles and Transparency

The principle of transparency emerged as critical. ccTLDs should make their use of AI systems transparent to stakeholders, explaining what data is analyzed, how decisions are made, and what safeguards exist. Human-in-the-loop requirements ensure that automated recommendations don't replace human judgment on consequential decisions like domain suspensions. This transparency builds trust and accountability consistent with RFC1591's emphasis on responsible management.

Future work is suggested to focus on how AI deployment can support rather than undermine the public interest mandate of ccTLDs. This includes developing guidelines for when AI is appropriate, what oversight mechanisms are necessary, and how to balance efficiency gains with fairness considerations. The goal is to ensure AI serves the community interest while maintaining the trust relationships essential to Internet governance.

Operationalization Mechanisms Beyond Technology

The discussion revealed that technology alone is insufficient for responsible AI deployment in registry operations. Supporting mechanisms are required to operationalize ethical principles and governance frameworks. The suggestion of ccTLD-specific AI boards or expert panels represents one approach to ensuring ongoing oversight and adaptation as AI systems evolve and new use cases emerge.

Regular retraining of AI systems requires dedicated expertise and resources. These experts must understand both the technical aspects of machine learning and the specific context of domain name abuse. Building and maintaining this expertise internally or accessing it through partnerships becomes a key operational consideration. The expertise gap could be addressed through training programs, knowledge sharing networks, or shared service models among multiple registries.

Audit mechanisms ensure AI systems perform as intended and don't drift toward biased or inaccurate outputs over time. These audits might examine false positive rates, demographic impacts of enforcement actions, and whether the system's recommendations align with registry policies. Establishing audit frequencies, methodologies, and response protocols for concerning findings represents important future work for the ccTLD community.

Best Practices and Joint Development

The RegCheck collaboration between .nl and .be registries demonstrates the value of joint development. Rather than each registry building separate systems, pooling resources and expertise creates more robust solutions while reducing individual costs. This model could extend to other ccTLDs, particularly smaller registries that

lack resources for independent AI development but could benefit from shared systems adapted to their contexts.

Formulating best practices that work across multiple ccTLD contexts emerged as an important goal. While specific implementations will vary based on zone size, legal frameworks, and market structures, core principles around human oversight, transparency, and continuous improvement could apply broadly. Sharing lessons learned, both successes and failures, accelerates learning across the community and helps smaller registries avoid pitfalls.

The discussion suggested creating forums for ongoing knowledge exchange about AI applications in registry operations. These could include technical working groups, policy discussions about governance frameworks, and operational exchanges about collaboration models with registrars and other stakeholders. Building a community of practice around responsible AI use in ccTLDs would benefit the entire ecosystem.

Questions remain about what can be standardized versus what must be customized. Data privacy requirements vary by jurisdiction, affecting what information can be collected and analyzed. Enforcement mechanisms depend on local legal frameworks and contractual relationships with registrars. Best practices must be flexible enough to accommodate these variations while maintaining core principles of fairness, transparency, and effectiveness.

AI as Both Defense and Attack Vector

A sobering reality discussed was that AI serves as both a defensive tool for registries and an offensive weapon for malicious actors. While registries use machine learning to detect abuse patterns, criminals employ AI to create more sophisticated attack networks. These automated systems can generate convincing fake websites at scale, adapt to detection methods, and create interconnected networks designed to lure victims while evading identification.

This arms race dynamic means registries cannot deploy AI once and consider the problem solved. As defensive capabilities improve, attackers adapt their techniques, requiring continuous evolution of detection systems. The continuous retraining requirement emerges not just from dataset obsolescence but from adversarial adaptation. Future discussions should address how registries can stay ahead in this ongoing competition.

The dual-use nature of AI technology raises questions about information sharing. While registries want to share best practices and detection methods with each other, publicizing too much detail about how systems work could provide attackers with roadmaps for evasion. Balancing transparency with operational security becomes an important consideration for the community moving forward.

Investment Decisions and Market Dynamics

SIDN's experience demonstrates that effective AI deployment requires substantial strategic investment. Beyond technology costs, registries must invest in expertise, infrastructure, and organizational processes to support AI systems. For the .nl registry managing over 6 million domains, this investment was justified as necessary to maintain TLD strength and user trust. Smaller ccTLDs face questions about whether similar investments are viable given their scale.

Market dynamics complicate investment decisions. The emergence of commercial services like Netcraft offering abuse detection feeds represents both an opportunity and a challenge. Registries can now purchase some capabilities rather than building everything internally, potentially reducing costs. However, relying on external vendors raises questions about data sharing, customization, and long-term sustainability.

The strategic positioning of being ahead of the market, as .nl achieved with fake webshop detection, provides competitive advantage for ccTLDs. Domains in zones known for strong abuse controls become more valuable and trustworthy. This market differentiation can justify investment costs by enhancing the TLD's reputation and attractiveness to legitimate registrants.

Future discussions should explore various investment models: fully internal development, purchasing commercial solutions, participating in shared development initiatives, or hybrid approaches. Each model has implications for cost, control, customization, and sustainability. Understanding which approaches work best under different circumstances would help ccTLDs make informed strategic decisions about their AI investments.

Data Sources and Privacy Considerations

Effective AI systems require diverse data inputs. The .nl registry combines internal registration data, DNS traffic patterns, crawled website content, and external threat intelligence feeds. Each data source provides different insights, and the combination enables more accurate abuse detection than any single source alone. However, collecting and analyzing these data sources raises important privacy considerations.

EU data protection regulations impose requirements on what data can be collected, how it can be processed, and how long it can be retained. The .nl registry's Privacy Board oversees these considerations, ensuring compliance while enabling effective abuse detection. As the EU AI Act comes into force, additional requirements around algorithmic transparency and impact assessment will apply. Other jurisdictions have different regulatory frameworks that ccTLDs must navigate.

Future work should address how registries can build privacy-preserving AI systems that achieve abuse reduction goals while minimizing data collection and protecting individual rights. Techniques like differential privacy, federated learning, or anonymization might enable effective analysis with reduced privacy risks. Sharing approaches to privacy-respecting AI development would benefit the entire ccTLD community.

Conclusion

The ICANN84 discussion demonstrated that AI presents significant opportunities for ccTLDs to reduce domain name abuse while raising important questions about governance, collaboration, and responsible deployment. The .nl registry's success in reducing fake webshops by 98% over five years provides a compelling case study, showing what is achievable with strategic investment and careful implementation.

Several key themes emerged: the necessity of keeping humans in the loop to ensure accountability and accuracy, the importance of governance structures including expert oversight and regular auditing, and the value of collaboration both between registries and with other stakeholders. Technology alone does not solve abuse

problems; it must be embedded within frameworks that ensure ethical, effective, and sustainable deployment.

The emerging topics identified during the session provide a roadmap for future ccTLD work on AI. Developing principles aligned with RFC1591 responsibilities, establishing operational mechanisms beyond technology, sharing best practices and potentially building joint systems, addressing the arms race with malicious actors, making informed investment decisions, and navigating privacy requirements all represent important areas for continued discussion and collaboration.

As AI capabilities continue to evolve rapidly, the ccTLD community must maintain ongoing dialogue about responsible deployment. The .nl experience demonstrates both the promise and the challenges of using AI to strengthen domain name systems. By sharing knowledge, developing common principles, and collaborating on solutions, ccTLDs can harness AI's potential while upholding their responsibilities to serve the public interest and maintain the trust essential to Internet governance.