# Cybersecurity and Artificial Intelligence Opportunities for MerTerre's Zéro Déchet Sauvage Platform

Ethan McDowell - *Brown University*

In collaboration with MakeSense in Paris, France and MerTerre in Marseille France

*This research was supported by the Laidlaw Foundation through the Laidlaw Scholars Leadership in Action Program*

August 12tt, 2025

## I. Abstract

This report examines the cybersecurity needs and artificial intelligence (AI) potentials for MerTerre's Zéro Déchet Sauvage participatory science platform, which aggregates citizen-generated waste data from coastal clean-ups across France. Drawing on academic frameworks and emerging technologies, it analyzes vulnerabilities in current data flows and proposes measures to ensure data integrity, confidentiality, and availability. Simultaneously, the report identifies opportunities to leverage AI in automating error detection, classifying waste, predicting pollution hotspots, and validating contributions. These dual enhancements aim to strengthen MerTerre's technical foundation, support scalable operations, and enhance policy advocacy.

## II. Introduction

In recent decades, the challenge of marine pollution has emerged as one of the most pressing environmental issues confronting coastal nations. The Mediterranean Sea, an enclosed basin bordered by over twenty countries, is especially vulnerable to waste accumulation due to its limited water exchange with the Atlantic and its high-density coastal populations. France, with its extensive Mediterranean coastline, plays a pivotal role in shaping conservation strategies that mitigate diffuse abandoned waste, that is, waste that escapes formal collection systems and accumulates in natural environments through a variety of pathways. Within this context, MerTerre, a Marseille-based non-governmental organization (NGO), has positioned itself at the forefront of participatory marine conservation efforts. Founded in 2000, MerTerre has dedicated itself to reducing marine litter through data-driven advocacy, volunteer mobilization, and collaboration with policy-makers at local, national, and international levels.

Central to MerTerre's approach is Zéro Déchet Sauvage, a participatory science platform designed to collect, store, and analyze data from waste clean-ups across France. The platform aggregates inputs from hundreds of volunteer groups, environmental associations, and institutional partners, creating a nationwide database that supports both grassroots activism and formal legislative action. This dual mission of bridging community engagement with policy influence has been recognized as a model of citizen science, yet it faces the technical challenges common to many environmental data platforms: inconsistent data quality, limited interoperability, and minimal integration of advanced computational tools.

My work with MerTerre during the summer of 2025 was carried out as part of the Leadership in Action program of the Laidlaw Scholars Network, which pairs undergraduate scholars with high-impact social projects around the globe. In this capacity, I brought a

background in computer science with particular interests in artificial intelligence (AI) and cybersecurity to assess the platform's current technical state, identify vulnerabilities, and explore pathways for AI-driven enhancement. Although the platform's primary goal is environmental protection, it operates within a data ecosystem that demands both technical robustness and future-proof scalability. As environmental data increasingly intersects with regulatory compliance, open-data movements, and cross-border information sharing, ensuring the integrity, confidentiality, and availability of such data has become essential not only for operational efficiency but also for public trust.

Globally, citizen science platforms are undergoing a technological transformation. Advances in AI have opened new possibilities for automating classification, detecting anomalies, predicting environmental trends, and validating submissions in near real-time. These capabilities, however, depend on the presence of clean, well-structured, and secure datasets—conditions that are often difficult to achieve in resource-constrained NGOs. At the same time, these platforms have become attractive targets for cyberattacks, ranging from opportunistic data breaches to more sophisticated attempts to corrupt environmental records. Such attacks not only jeopardize organizational credibility but can also have downstream effects on environmental policy-making, especially when data is used to justify legislative measures or international agreements.

In the case of Zéro Déchet Sauvage, these dynamics are particularly significant. The platform's dataset is not merely an academic resource; it is a living repository of evidence used to inform France's compliance with European Union directives, to support the negotiation of global agreements like the United Nations Global Plastics Treaty, and to evaluate the efficacy of national waste reduction strategies. A compromised dataset could weaken advocacy campaigns, undermine public trust, and limit the NGO's capacity to influence decision-making. Conversely,

a well-secured and AI-enabled platform could dramatically enhance MerTerre's ability to mobilize stakeholders, provide actionable intelligence to policy-makers, and anticipate emerging waste trends.

This report therefore examines the cybersecurity posture and AI-readiness of the Zéro Déchet Sauvage platform, situating these technical domains within the broader environmental and policy landscape. It begins by analyzing the platform's current architecture and identifying potential vulnerabilities in data collection, transmission, and storage. It then explores AI-driven opportunities for waste classification, geospatial hotspot prediction, and automated data validation. Throughout, the analysis is informed by best practices from comparable environmental data initiatives, frameworks from the National Institute of Standards and Technology (NIST), and emerging literature on the ethical deployment of AI in conservation. By addressing both defensive and innovative technical measures, this report seeks to demonstrate how a citizen science platform can evolve into a resilient, intelligent system that advances marine conservation while safeguarding the digital infrastructure upon which its mission depends.

## III. Cybersecurity Analysis

The protection of environmental datasets is no longer a peripheral concern but a central operational necessity for organizations whose work depends on the integrity and credibility of the information they collect. For MerTerre's Zéro Déchet Sauvage participatory science platform, cybersecurity safeguards are vital to ensure that the marine waste data submitted by volunteers and partner organizations remains accurate, authentic, and secure. These datasets are not merely archival; they are actively used to inform conservation strategies, influence policy decisions, and underpin advocacy efforts at both the national and international levels. Any compromise to the platform's security, whether through malicious tampering, unauthorized access, or large-scale data loss, could undermine years of coordinated environmental monitoring and damage the trust of stakeholders who depend on its findings.

The threat landscape facing an environmental data platform such as Zéro Déchet Sauvage is distinct from that of financial institutions or healthcare providers, yet it is far from insignificant. While the monetary value of the data may be limited, its strategic and political value can be considerable. Industrial actors seeking to obscure pollution levels, political groups aiming to discredit environmental regulations, or even opportunistic hackers motivated by disruption could all find reasons to target such a platform. This is particularly relevant in contexts where pollution data could influence the enforcement of environmental regulations, the distribution of conservation funding, or the public perception of specific industries. In addition, the open, distributed nature of MerTerre's data collection model, relying on contributions from individuals, NGOs, and local authorities across multiple regions, creates potential points of vulnerability. Without robust contributor authentication, it remains theoretically possible for

actors to submit falsified data designed to skew statistics, introduce doubt, or reduce the platform's perceived credibility.

A preliminary review suggests that MerTerre's current cybersecurity approach covers essential practices such as access control and periodic data backup but does not appear to be guided by a formalized security framework, such as the National Institute of Standards and Technology (NIST) Cybersecurity Framework or the ISO/IEC 27000 series. While these frameworks are extensive and may not be fully practical for a small NGO to implement in their entirety, adopting scaled and prioritized elements would meaningfully enhance resilience against both targeted and opportunistic attacks. Current authentication practices appear to rely heavily on email-based verification, with little evidence of multi-factor authentication (MFA) for either administrative or contributor accounts. Similarly, while it is likely that data transmission is encrypted, there is no public documentation confirming compliance with modern protocols such as TLS 1.3. Backup procedures seem to exist but may be stored locally rather than geographically redundant, leaving the platform vulnerable to localized system failures or ransomware attacks. Administrative privileges also appear centralized, which can reduce complexity but could be further refined by implementing role-based access control (RBAC) to minimize unnecessary permissions.

The potential attack vectors facing the platform are diverse. Data poisoning, in which false or manipulated submissions are introduced to corrupt the dataset, remains one of the most serious threats. Denial-of-service attacks, either via mass submission of junk data or via direct network exploitation, could reduce platform availability during critical reporting periods. Man-in-the-middle attacks could intercept unencrypted submissions, while credential theft through phishing or brute-force attempts could compromise administrator accounts. Insider

threats, although less common, cannot be dismissed, particularly in small teams where individuals may have broad access rights. These risks are compounded by the platform's reliance on trust-based collaboration with external partners whose own systems and practices may not meet equivalent security standards.

Addressing these vulnerabilities requires a phased and realistic approach. Multi-factor authentication should be introduced for all administrative accounts and, where feasible, for contributor accounts as well. Role-based access control would ensure that users have access only to the data and functions necessary for their responsibilities, reducing the risk of both intentional and unintentional misuse. All data transmission should be encrypted using TLS 1.3, and HTTPS should be enforced site-wide. Data validation and sanitization protocols should be automated, allowing the system to flag submissions with improbable geospatial coordinates, extreme values, or inconsistencies with known waste categorization standards. Regular vulnerability scanning, using open-source or commercial tools, would help identify weaknesses before they can be exploited. Finally, backup systems should be restructured to ensure geographic redundancy and secure recovery capabilities in the event of data loss.

Investing in these cybersecurity measures is not merely defensive. In an era where environmental data is increasingly scrutinized by competing stakeholders, robust cybersecurity strengthens MerTerre's position as a trusted and authoritative voice in marine conservation. It also lays the groundwork for more advanced technical integrations, such as artificial intelligence-driven data analysis, which require secure and high-quality datasets to function effectively. By framing cybersecurity as a foundational enabler rather than an ancillary cost, MerTerre can ensure that its data not only remains safe but continues to serve as a reliable basis

for scientific research, policy advocacy, and public engagement in the fight against marine pollution.

## IV. Artificial Intelligence Analysis

Artificial intelligence (AI) presents significant opportunities for enhancing the efficiency, accuracy, and impact of participatory science platforms like Zéro Déchet Sauvage. While MerTerre's current data management processes rely primarily on manual review and standard categorization systems, the integration of AI could automate key functions such as data cleaning, anomaly detection, and predictive analysis. These capabilities would not only reduce administrative workload but also improve the quality and usability of the dataset for both conservation practitioners and policymakers. Importantly, AI can be deployed in ways that directly align with MerTerre's mission: supporting a citizen-science-driven, decentralized model while ensuring that contributions from diverse actors are consistently standardized and scientifically valid.

One of the most immediate applications of AI within MerTerre's platform lies in automated data validation. Machine learning algorithms, trained on historical waste collection data, could be used to detect submissions that fall outside expected parameters—such as implausible geospatial coordinates, waste composition totals that are inconsistent with historical patterns, or categorization anomalies caused by user error. Such models could provide real-time feedback to contributors, prompting them to confirm or correct entries before final submission. By intercepting errors at the point of entry, the platform could drastically reduce the time and resources required for post-hoc data cleaning, freeing up staff and volunteers to focus on strategic analysis and outreach.

AI could also play a transformative role in image recognition and categorization for waste monitoring. While current reporting systems rely on manual categorization of debris, incorporating computer vision models could enable contributors to upload photographs of

collected waste that are automatically analyzed and classified. Such technology has already been successfully implemented in environmental monitoring initiatives by organizations like The Ocean Cleanup and the Marine Debris Monitoring and Assessment Project. For MerTerre, this would standardize categorization across regions and skill levels, reducing subjective variation between contributors and ensuring that waste type classifications remain consistent over time. Moreover, the visual database generated through this process could serve as an additional layer of verifiable evidence, strengthening the credibility of datasets used in advocacy and policymaking.

Another promising avenue is the use of predictive analytics to anticipate waste accumulation patterns based on historical data, weather events, tourist activity, and regional policy changes. Machine learning models could identify correlations between environmental conditions and pollution surges, enabling targeted clean-up campaigns in areas likely to experience high waste deposition. For example, predictive models could inform seasonal action plans, alerting partner organizations when certain coastal zones are at heightened risk. This would not only optimize volunteer mobilization but also support the development of policy interventions that address root causes of pollution at key moments.

AI can further enhance policy advocacy by generating advanced visualizations and summaries of large datasets. Natural language processing (NLP) tools could be used to convert complex waste data into plain-language summaries tailored for different audiences, from local community groups to legislative bodies. In parallel, AI-driven geospatial analysis could overlay waste data with other environmental and socioeconomic indicators—such as fishing zones, shipping routes, or demographic density—to strengthen evidence-based arguments for regulatory change. These capabilities could be instrumental in bridging the gap between raw scientific data

and the actionable insights needed to influence decision-making at municipal, national, and international levels.

However, the integration of AI into MerTerre's workflows must be approached cautiously to avoid potential pitfalls. AI systems are only as reliable as the data they are trained on, and biases or inaccuracies in the existing dataset could propagate through automated decision-making processes. This underscores the importance of combining AI with robust cybersecurity measures, as compromised or manipulated data could directly misinform models and lead to flawed outputs—a risk particularly acute in contexts where environmental data may be politically sensitive. Furthermore, AI tools require ongoing maintenance, computational resources, and, in some cases, specialized expertise that may be beyond the current scope of MerTerre's in-house capabilities.

To maximize the benefits of AI while mitigating risks, MerTerre should consider adopting a phased implementation strategy. Initial phases could focus on low-risk, high-impact tools such as automated anomaly detection for data entries and AI-assisted image classification for debris categorization. These tools could be deployed as opt-in features for contributors, allowing the organization to refine models and address usability issues before scaling them platform-wide. Later phases could integrate predictive analytics and policy-facing visualization tools, provided that the platform's security, governance, and data quality standards have matured sufficiently to support more complex applications. Collaboration with universities, research institutes, and technology partners could also provide access to AI expertise without requiring significant internal restructuring.

In sum, artificial intelligence has the potential to significantly amplify the reach and effectiveness of MerTerre's mission by transforming a diverse, volunteer-driven dataset into a

more consistent, actionable, and policy-relevant resource. When paired with the right safeguards and phased adoption, AI can help bridge the gap between grassroots environmental engagement and the large-scale systemic change necessary to protect the Mediterranean's marine biodiversity.

## V. Intersection of Cybersecurity and AI

The integration of artificial intelligence (AI) into MerTerre's digital infrastructure necessarily raises complex cybersecurity considerations. AI systems are only as reliable as the data they process, and in the case of the Zéro Déchet Sauvage platform, this data originates from a distributed network of citizen scientists, NGOs, and institutional partners. Ensuring the authenticity and integrity of these inputs is essential—not only for maintaining the credibility of the dataset but also for protecting against malicious interference that could skew environmental reporting or erode public trust. In practice, AI-driven analytics introduce new attack surfaces, such as adversarial input manipulation, model poisoning, and unauthorized extraction of proprietary algorithms, all of which must be mitigated through proactive cybersecurity design.

One of the most salient intersections between AI and cybersecurity in MerTerre's context is data validation at scale. Manual verification of waste categorization submissions becomes impractical as the platform expands, making AI-assisted anomaly detection an attractive option. For example, machine learning models could be trained to flag inconsistent or outlier entries, such as improbable waste counts, impossible geolocation coordinates, or sudden surges in a specific category, that may indicate either human error or coordinated disinformation. However, the reliance on such automated systems creates a dependency on the confidentiality, integrity, and availability (CIA) of both the underlying data and the AI models themselves. If these systems were compromised, attackers could alter detection thresholds, suppress alerts, or generate false positives, undermining both conservation outcomes and stakeholder confidence.

From a systems architecture perspective, secure model deployment becomes critical when AI is embedded into real-time decision-making processes. Techniques such as model watermarking, differential privacy, and federated learning could be adopted to protect intellectual

property while limiting the exposure of sensitive environmental data. For example, federated learning would allow MerTerre to train AI models across partner organizations' datasets without transferring raw data to a central repository, thereby reducing the risk of large-scale breaches. This approach not only safeguards against external intrusion but also mitigates insider threats, ensuring that access to data is strictly limited and traceable.

Another intersectional challenge lies in explainability and auditability. In a policy-driven context, MerTerre must be able to justify AI-derived insights to regulators, funders, and the public. This means implementing AI architectures that are not only accurate but also transparent enough to allow for third-party audits without exposing vulnerabilities that could be exploited by malicious actors. The convergence of cybersecurity and AI in this domain demands layered safeguards: robust authentication and authorization protocols, secure data pipelines, tamper-evident logging, and redundancy measures that maintain system functionality even during an active cyber incident.

Ultimately, the responsible integration of AI into MerTerre's operations requires an AI-cybersecurity symbiosis, a design philosophy where cybersecurity is not a separate, after-the-fact add-on, but an intrinsic element of AI development and deployment. By embedding secure-by-design principles into its AI initiatives, MerTerre can ensure that technological innovation does not outpace the organization's capacity to manage risk. In doing so, the platform can maintain the integrity of its environmental mission while setting a precedent for other NGOs seeking to merge participatory science, AI-driven analytics, and rigorous cybersecurity practices.

## VI. Conclusion

The integration of cybersecurity and artificial intelligence within environmental data platforms such as Zéro Déchet Sauvage presents both transformative opportunities and critical challenges. From a cybersecurity perspective, ensuring the confidentiality, integrity, and availability of environmental datasets is paramount, particularly when these datasets inform public policy and international advocacy. Unauthorized data manipulation, whether through malicious intrusion or unintentional mismanagement, could undermine the credibility of marine conservation initiatives and erode stakeholder trust. Proactive security measures, including encryption protocols, secure authentication systems, and regular vulnerability assessments, must therefore be embedded as foundational elements of any platform upgrade.

Artificial intelligence, in parallel, offers a powerful suite of analytical tools capable of accelerating data processing, detecting anomalies, and producing actionable insights for decision-makers. For MerTerre, AI-driven classification systems could streamline the categorization of waste types, while predictive modeling could help forecast pollution hotspots based on historical and environmental patterns. Yet the deployment of such systems necessitates a responsible framework, ensuring that AI outputs are transparent, interpretable, and free from biases that could distort conservation priorities.

The intersection of AI and cybersecurity is where the potential for both innovation and risk intensifies. While AI can bolster cybersecurity through intelligent threat detection, adversarial AI techniques could be used to exploit vulnerabilities in the very systems designed to safeguard data. This interplay demands a governance approach that blends technical safeguards with ethical oversight, ensuring that the evolution of the platform aligns with both technological best practices and MerTerre's mission-driven values.

Ultimately, the path forward for MerTerre lies in treating cybersecurity and AI not as ancillary considerations, but as integral pillars of platform design and organizational strategy. By investing in secure, intelligent, and transparent systems, MerTerre can enhance the resilience of its participatory science platform, strengthen its role as a trusted source for environmental data, and amplify its influence in shaping both national and international marine conservation policy. In doing so, the organization not only protects the integrity of its data, but also reaffirms its commitment to a future in which technology serves the preservation of the Mediterranean Sea and the communities that depend upon it.