AI-Enhanced Biodiversity Risk Assessment for Hospitality Sector Compliance in New Jersey

Abstract

Biodiversity loss in New Jersey has become a growing concern, especially around commercial zones such as hotels. In response to newly introduced state regulations aimed at promoting environmental compliance, this project presents *BiodivProScope*, a location-based risk assessment and mitigation tool specifically designed for hotel use. The system enables hotel users to register, log in, and manage their locations by adding, viewing, or deleting addresses. It then performs geospatial analysis within a 5-mile radius using the Haversine formula, supported by a PostgreSQL database, to identify local ecological risks such as threatened and invasive species, and pollution indicators. The tool assigns risk scores and generates mitigation strategies using a hybrid approach—combining rule-based logic with AI-enhanced retrieval from a curated knowledge base via ChromaDB. Users can interact with a dynamic map to visualize risks, preview categorized mitigation actions, and generate downloadable reports in multiple formats. *BiodivProScope* ultimately serves as a practical solution to help hotel managers comply with biodiversity regulations while enhancing their sustainability efforts.

Keywords

Biodiversity, Web-Based Risk Assessment, New Jersey, Hotels, IUCN Red List, Invasive Species, Mitigation Strategy, PostgreSQL, Leaflet, Marine Water HCI, Freshwater HCI, React

---

1. Introduction

Biodiversity is facing an unexpected global decline driven by human-caused factors such as habitat destruction, water pollution, the spread of invasive species, and climate change. This loss has far-reaching consequences, not only for ecological stability but also for public health, economic systems, and resilience to climate-related impacts. Commercial sectors, particularly hotels and restaurants located near ecologically sensitive areas, may unknowingly contribute to this degradation through extensive land use, water consumption, and waste output [1][2].

In recognition of these environmental risks, the New Jersey Department of Environmental Protection (NJDEP) has implemented a set of biodiversity-conscious operational guidelines specifically aimed at the hospitality industry [3]. These frameworks include recommendations for identifying local ecological threats, minimizing operational impact, and adopting sustainable practices tailored to New Jersey’s diverse ecosystems, such as its coastal wetlands and urban parks. While these initiatives are timely and well-intentioned, they are often difficult for hotel operators to interpret or implement without access to ecological data or specialized knowledge.

To address this gap, this paper introduces **BiodivProScope**, a location-based biodiversity risk assessment and mitigation tool designed to support hotel compliance with New Jersey's environmental regulations. The system simplifies regulatory adherence by offering automated risk detection, spatial analysis, and evidence-based mitigation recommendations. It repositions environmental conservation as a strategic advantage—enhancing brand image, meeting sustainability expectations of modern travelers, and aligning operations with emerging regulations. The paper outlines the tool's architecture, integrated data sources, and hybrid risk assessment framework that combines species presence data with weighted pollution scoring. It also highlights the use of machine learning (ML) and natural language processing (NLP) to generate adaptive, regulation-informed mitigation strategies—making ecological responsibility more accessible for non-technical users.

Existing environmental assessment tools such as the International Union for Conservation of Nature (IUCN) STAR system and the Environmental Protection Agency's (EPA) Enviromapper offer valuable biodiversity data, but they are often broad in scope and not optimized for business-specific, localized decision-making. In contrast, BiodivProScope is tailored for New Jersey hotels, integrating IUCN threat data, NJDEP regulatory documents, pollution metrics, and Habitat Condition Index (HCI) scores into a unified, user-friendly platform.

BiodivProScope enables hotel managers to add and manage locations, perform biodiversity risk analyses within a 5-mile radius, and access tailored, AI-powered mitigation strategies. In doing so, it bridges the gap between ecological science and real-world business operations, empowering users to make informed, sustainable decisions. Ultimately, the system helps hotels demonstrate compliance with biodiversity regulations while contributing to long-term environmental stewardship.

---

2. Literature Review

A. Prior Work on Biodiversity Risk Platforms and Environmental Web Tools

Biodiversity loss has been widely recognized as an urgent global issue, with significant implications for both ecological stability and human well-being [1] [2]. In recent years, numerous digital platforms have emerged to support biodiversity tracking and visualization efforts. Tools such as the IUCN's Species Threat Abatement and Restoration (STAR) system [4], the U.S. EPA's Enviromapper [5], and the Global Biodiversity Information Facility (GBIF) [6] provide regional to global-scale insights on biodiversity conditions and threats. However, these platforms are primarily designed for scientific monitoring and policy-level assessment and fall short in delivering actionable guidance tailored to business users—particularly within the hospitality industry.

Other platforms like AquaMaps [7] focus on visualizing species richness and distribution across broad geographic ranges.. While informative, they do not support business-specific decision-making or offer concrete strategies for mitigating biodiversity risks at a localized level.

B. Use of Spatial Tools (Leaflet, GIS) in Environmental Science

Geographic Information Systems (GIS) and JavaScript libraries like Leaflet.js have become integral in developing web-based tools for spatial ecology and pollution monitoring [8] [9]. The integration of PostGIS, an extension of PostgreSQL, with Leaflet has enabled real-time geospatial applications that allow users to conduct spatial queries, analyze environmental patterns, and visualize risks dynamically [10].

Prior implementations in conservation planning, disaster management, and environmental monitoring demonstrate the utility of interactive web maps in local decision-making contexts [11] [12]. These tools lay the technological foundation for web-based risk assessment platforms like BiodivProScope, which depend on spatial interactivity and geographic filtering to deliver location-specific ecological insights.

C. Risk Frameworks Using Species Distribution and Pollution Metrics

Contemporary ecological risk frameworks often combine multiple data layers, including species presence, conservation status, and anthropogenic impact indicators [14]. One widely referenced dataset in this context is the Human-Caused Impact (HCI) index developed by the World Bank, which provides normalized scores representing pressure on marine, freshwater, and terrestrial ecosystems [13].

While IUCN-based risk scoring remains a cornerstone of conservation science, few existing tools integrate such scoring systems with pollution and invasive species datasets in a cohesive way [2] [15]. The fusion of these diverse risk indicators is particularly relevant for hotel sites, which are embedded within complex human-natural environments subject to overlapping ecological threats.

D. Hotel Industry Sustainability Tools and Regulation Compliance Systems

The hospitality sector has adopted several sustainability certification programs—including LEED, Green Key, and EarthCheck—which emphasize energy efficiency, water conservation, and waste reduction [16] [17]. However, these programs often neglect biodiversity-specific risk factors. The NJDEP's Hotel and Lodging Sustainability Guidance represents a significant departure from this trend by explicitly incorporating biodiversity awareness and local ecological considerations into hotel compliance strategies [3].

Despite this progress, tools that translate these regulatory recommendations into actionable, location-specific decisions for hotels remain scarce. This gap highlights the need for platforms that are both policy-aware and context-sensitive to hotel operations and surrounding ecosystems.

E. Summary of Gaps Addressed by BiodivProScope

Most of the platforms and frameworks discussed above share several key limitations. They typically do not address the unique biodiversity compliance needs of hotels in New Jersey. Few offer automated biodiversity risk scoring aligned with business user workflows or provide mitigation strategies grounded in real-world regulatory documents. Additionally, current tools rarely support multi-format report exports (PDF/CSV/Excel), which are crucial for sustainability audits and reporting.

BiodivProScope directly addresses these shortcomings by offering a hotel-specific compliance tool that enables interactive spatial querying (via Leaflet), integrates AI-driven mitigation recommendations using ChromaDB and a Retrieval-Augmented Generation (RAG) model, and draws from credible sources including the IUCN Red List, NJDEP guidance, and World Bank biodiversity datasets.

---

3. Data (III)

A. Core Datasets for Biodiversity Risk Assessment

Commented [1]: Regarding the use of database like IUCN, will there be automatic update in our website when there is any change in species in IUCN database?

The BiodivProScope platform relies on a combination of reputable ecological datasets to assess biodiversity risks around hotel locations. The IUCN Red List of Threatened Species serves as the primary reference for species conservation status. Species occurrences within the geographical areas relevant to hotel sites are filtered based on their threat categories, such as Critically Endangered and Endangered classifications [4]. To supplement this, regional invasive species datasets are incorporated to identify and assess species that disrupt local biodiversity and hotel-adjacent ecosystems [18].

In addition to species-based data, aquatic ecosystem risks are evaluated using marine and freshwater pollution datasets [14] obtained from state and federal environmental agencies. These datasets include measurements of nutrient runoff, contaminant concentrations, and the health of nearby freshwater bodies, providing a comprehensive view of aquatic biodiversity pressures [19]. Terrestrial risks are assessed through the HCI , which quantifies human activity pressures such as land development, traffic density, and agricultural land use [20]. This multidimensional dataset selection ensures that BiodivProScope offers a balanced perspective on species conservation, ecological health, and anthropogenic impacts.

B. Relevance of Data to Hotel-Adjacent Biodiversity Risk

To ensure that biodiversity risk assessments are contextually meaningful for hotel operations, each dataset used in BiodivProScope is spatially filtered based on its proximity to user-specified hotel locations. The objective is to identify ecological threats that are both immediate and

actionable within the surrounding environment of hotel properties. This localized approach enhances the platform's effectiveness in supporting guest safety, promoting regulatory compliance, and reinforcing environmental responsibility.
The selection of datasets was guided by their ability to represent three critical and complementary dimensions of biodiversity risk: species preservation, referring to the presence of threatened or endangered flora and fauna; ecological balance, involving the spread of invasive species that disrupt native ecosystems; and human-induced pressures, such as pollution and habitat fragmentation, that accelerate biodiversity degradation. These dimensions were chosen based on their alignment with ecological risk frameworks widely recognized in conservation science [15].

C. Justification for the 5-Mile Radius Metric

To define the area of influence around each hotel, a 5-mile (approximately 8 kilometers) radius was selected. This metric balances the need for ecological relevance with computational efficiency, ensuring a manageable scope of data retrieval and analysis. A 5-mile radius typically encompasses critical ecological features such as nearby water bodies, parks, nature preserves, and residential developments that can influence the local biodiversity profile of a hotel [21]. Furthermore, the selected radius aligns with recognized environmental assessment guidelines and planning frameworks used in community-scale ecological evaluations.

D. Data Management Tools: Nominatim, PostgreSQL, and ChromaDB

To support efficient geospatial data processing, the Nominatim API from OpenStreetMap is used for geocoding zip codes and addresses submitted by hotel users. This service converts hotel input locations into precise latitude and longitude coordinates [22]. The backend relies on PostgreSQL, a robust relational database system, extended with the PostGIS module to enable spatial indexing and radius-based querying [9]. This architecture ensures rapid and accurate retrieval of environmental risk data surrounding hotel properties.

For mitigation action generation, BiodivProScope utilizes ChromaDB, an open-source vector database. ChromaDB is employed to store and retrieve vectorized embeddings of real-world mitigation strategies, enabling the use of a Retrieval-Augmented Generation (RAG) model. This approach supports dynamic, context-aware mitigation recommendations that are grounded in regulatory and field-validated conservation practices [23].

---

4. Methodology

A. System Architecture and Setup

The BiodivProScope platform is constructed using a three-tier architecture, comprising a frontend, backend, and database layer. This modular architecture ensures the system remains scalable, maintainable, and extensible to accommodate future enhancements.

The overall system architecture is illustrated in Figure 1.

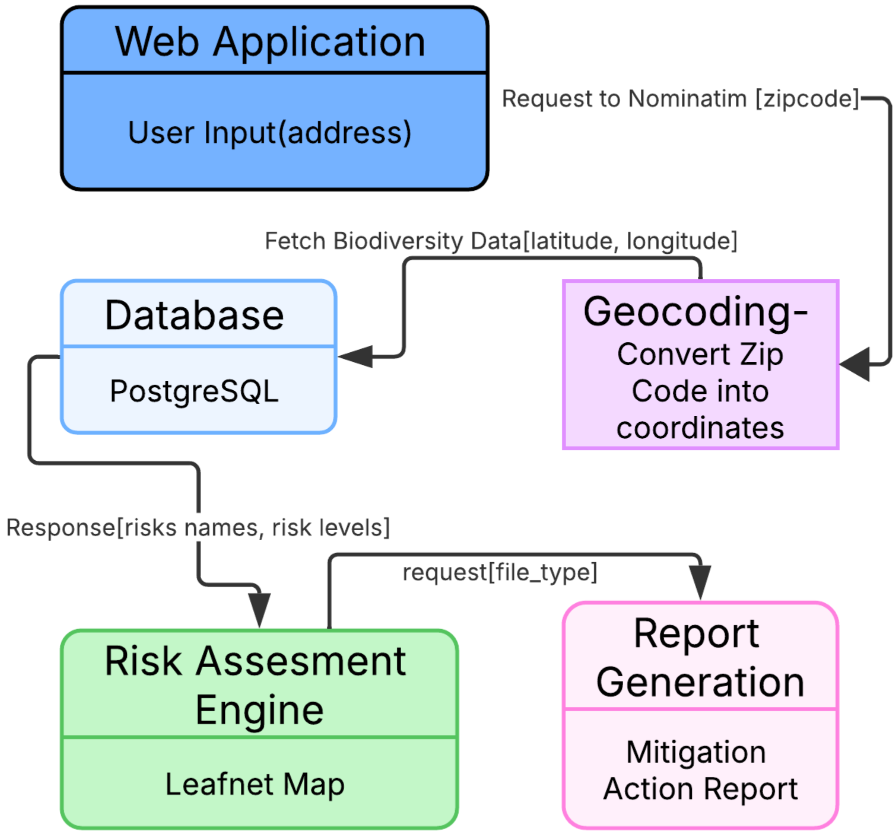


Figure 1:System architecture of BiodivProScope platform illustrating the user input process, geocoding, biodiversity data retrieval, risk assessment, and report generation modules.

The **frontend** is developed using **React.js**, a widely adopted JavaScript library maintained by Facebook. React's component-based structure and virtual DOM enable efficient dynamic rendering and responsive user interfaces. These capabilities are crucial for managing complex ecological datasets, supporting real-time interactive maps, and rendering downloadable mitigation reports [24].

The **backend** is implemented using the **Python Flask** microframework, chosen for its lightweight, modular design. Flask governs the server-side logic, handles HTTP requests, and manages communication between the frontend and the database. Additionally, it supports the integration of machine learning models and geospatial analytics, enabling the platform to deliver AI-driven mitigation strategies through RESTful API endpoints [25].

The **database** layer is powered by PostgreSQL, an open-source relational database management system recognized for its reliability and extensibility. PostgreSQL's integration with the PostGIS extension allows efficient handling of spatial data, enabling complex radius-based queries essential for calculating biodiversity risks within defined geographic boundaries [26].

This tri-layered system design provides the foundation for seamless data ingestion, geographic analysis, and AI-driven mitigation recommendations.

B. Technologies Utilized

BiodivProScope incorporates a range of modern technologies to support interactive visualization, AI-powered analysis, and collaborative development workflows.

Leaflet.js is utilized as the core geospatial visualization library. Its lightweight and mobile-friendly design enables dynamic map interactions such as marker clustering and multi-layer overlays, which are essential for conveying localized biodiversity risks in an accessible and visually engaging format [27].

For version control and team collaboration, the project uses **GitHub**, allowing structured source code management, branching, pull request workflows, and transparent issue tracking. This fosters efficient coordination across development cycles and ensures reproducibility [28].

To support AI-driven mitigation recommendation, **ChromaDB** is integrated as the vector database layer. It enables fast retrieval of semantically embedded data, allowing the system to match identified risks to real-world mitigation strategies using high-dimensional similarity search [23].

The **machine learning stack** includes **scikit-learn** for traditional models such as Random Forest classifiers, used to predict context-specific mitigation actions based on input features like risk type and severity [29]. For embedding unstructured textual data, the platform employs **SentenceTransformers** from Hugging Face, specifically leveraging pre-trained models like *all-MiniLM-L6-v2* for efficient, context-aware semantic matching [30].

These tools collectively empower BiodivProScope to deliver intelligent, responsive, and user-aligned biodiversity risk insights tailored to hotel operations.

5. Model Development

A. Logic for Biodiversity Risk Scoring

The biodiversity risk scoring framework implemented within BiodivProScope combines geospatial species presence data with environmental stress indicators to evaluate ecological threats surrounding hotel locations. The system adopts a **hybrid risk assessment approach**: (1) **presence-based identification** for species-related risks, and (2) **score-based evaluation** for ecosystem-level pollution threats [4], [31].

For **IUCN Red List species** and **invasive species**, the system identifies risk based on the **presence of such species within a 5-mile radius** of the hotel location, using data aligned with IUCN threat categories and regional biodiversity lists.

In contrast, for **pollution-related environmental stressors**—including **terrestrial**, **freshwater**, and **marine HCI (Habitat Condition Index)** datasets—the system employs a **numeric risk scoring formula** [13] to calculate composite threat levels:

$$\text{Risk Score} = \sum_{i=1}^{n} (w_i \times r_i) \tag{1}$$

where:

- w_i = weight assigned to risk factor i,
- r_i = normalized risk value of risk factor i (scaled between 0 and 1),
- n = total number of relevant risk factors.

Risks are categorized into three threat levels—High, Moderate, and Low—based on the computed composite risk score. The classification thresholds are defined as follows:

$$\text{Threat Level} = \begin{cases} \text{High}, & \text{if Risk Score} \geq 0.7 \\ \text{Moderate}, & \text{if } 0.4 \leq \text{Risk Score} < 0.7 \\ \text{Low}, & \text{if Risk Score} < 0.4 \end{cases} \tag{2}$$

The formula integrates normalized environmental risk scores weighted using the Analytic Hierarchy Process (AHP), based on enforcement violation data from NJDEP.

This approach reflects principles of Multi-Criteria Decision Analysis (MCDA) and Environmental Risk Assessment (ERA) [36] [37].

B. Use of Haversine Formula for Radius Search

Spatial filtering is a core component of BiodivProScope's risk assessment pipeline. Geospatial proximity is computed using the Haversine formula, which measures the shortest path over the Earth's surface between two coordinates. Only those species or ecosystem stressors located within a 5-mile radius of the hotel are retrieved for analysis, ensuring that risk assessments remain localized, relevant, and actionable [32]. This technique leverages latitude-longitude indexing and spatial querying strategies, optimizing database performance and minimizing unnecessary data retrieval.

C. Machine Learning Component for Mitigation Recommendation

This model is trained on a curated dataset comprising historical mitigation records, threat types, regional characteristics, and prior conservation actions. Features such as risk type, threat level, geographical region, and mitigation history are used to predict the most suitable mitigation strategy for new cases. Initial model evaluations demonstrate promising accuracy in strategy prediction, with plans to benchmark its performance against alternative classifiers, including decision trees and support vector machines (SVMs), to ensure robustness and generalizability [33]. The inclusion of a machine learning model enhances the system's ability to offer data-driven, context-specific recommendations for biodiversity risk management.

D. Hybrid Retrieval-Augmented Generation (RAG) Model

In addition to traditional machine learning approaches, BiodivProScope implements a hybrid Retrieval-Augmented Generation (RAG) model to further refine mitigation suggestions. The system first performs a similarity search over a ChromaDB vector store containing embedded conservation guidelines and mitigation reports, such as those published by NJDEP and IUCN [34]. Retrieved mitigation excerpts are merged with machine learning predictions to generate recommendations that are both grounded in documentation and tailored to user context. This hybrid approach leverages the factual reliability of retrieval-based methods alongside the flexibility of generative language models, producing high-quality mitigation recommendations even in cases where historical data is sparse or rule-based logic would otherwise fail.

6. Data Collection

A. Spatial Data Preprocessing and Standardization

For consistent geospatial analysis, all location-based data utilized within BiodivProScope was standardized to the WGS 84 coordinate system. This choice ensures compatibility with proximity calculations conducted through the Haversine formula, which is essential for accurately determining environmental risks within a predefined radius. Bounding box and circular spatial queries were employed, both standardized to a 5-mile search area extending from the centroid of the hotel's input zip code. This spatial preprocessing guarantees that all geolocation calculations are uniform, enhancing the reliability of biodiversity threat detection across different hotel sites.

B. Data Cleaning, Structuring, and Storage

To maximize data integrity, a series of data cleaning operations were performed prior to database integration. Duplicate species records were identified and removed to avoid inflating biodiversity risk metrics. All species names were normalized using a consistent genus/species taxonomy format to ensure coherence across datasets. For environmental pressure datasets, such as marine and freshwater HCI scores, normalization techniques were applied to rescale raw environmental stress values onto a uniform 0–1 risk scale, thereby facilitating standardized comparisons across different ecosystem types.

All cleaned datasets were then transformed into structured formats such as CSV and JSON files. These files were ingested into PostgreSQL, the platform's primary relational database, with dedicated tables organized by risk category (e.g., iucn_species, marine_risks, terrestrial_hci). This structured relational model allowed efficient querying and modular database management, critical for supporting real-time risk visualization and reporting.

Backend integration was handled through the Flask application, where structured SQL queries retrieved biodiversity threat information directly from PostgreSQL. Retrieved data was transmitted to the React.js frontend and rendered interactively via Leaflet.js maps, enabling hotel users to visualize their local biodiversity risks dynamically.

C. Data Assumptions and Limitations

Several practical assumptions were necessary during the data collection and integration process. It was assumed that species recorded within a zip code could plausibly occur anywhere within the corresponding 5-mile analysis radius, providing a working model for local biodiversity presence without hyper-granular distribution data. In applying mitigation guidelines, it was assumed that the NJDEP recommendations were broadly applicable across various hotel categories unless explicitly stated otherwise in the documentation.

Marine, freshwater, and terrestrial risk values were derived from the Habitat Condition Index (HCI) datasets published by the World Bank, retrieved on February 15, 2024. These values were treated as temporally stable throughout the assessment period, acknowledging that real-time environmental variations—such as seasonal changes or urban development—were not captured due to the static nature of the dataset [14].

Species threat status and distribution data were obtained from the IUCN Red List of Threatened Species, retrieved on March 1, 2024. Since the IUCN database is periodically updated to reflect changes in species conservation status, the system may not reflect the most current listings at any given time [4].

These limitations underscore the challenges of using static data sources in dynamic ecological systems. Future versions of BiodivProScope may incorporate real-time data streams or periodic dataset refresh mechanisms to enhance temporal accuracy and support more responsive risk assessments.

Formatted: Font color: Black

7. Data Extraction and Analysis

A. Geolocation-Based Risk Querying

To retrieve localized biodiversity risks, BiodivProScope geocodes user-specified hotel locations using the Nominatim API from OpenStreetMap, which converts addresses or zip codes into WGS 84 latitude and longitude coordinates [22]. It then calculates the great-circle distance between hotel coordinates and geotagged environmental risk entries to perform precise 5-mile radius queries, using the Haversine formula [32]:

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\Delta\lambda}{2}\right)}\right) \quad (3)$$

where:

- d = distance between two points (in kilometers),
- r = radius of the Earth (mean radius = 6,371 km),
- ϕ_1, ϕ_2 = latitudes of the two points (in radians),
- λ_1, λ_2 = longitudes of the two points (in radians),
- $\Delta\phi = \phi_2 - \phi_1$,
- $\Delta\lambda = \lambda_2 - \lambda_1$.

These spatial queries are executed using PostgreSQL, leveraging its PostGIS extension for efficient geospatial indexing. Through this mechanism, relevant entries across multiple tables—such as iucn species, marine risks, terrestrial hci, and invasive species are retrieved rapidly, forming the basis for hotel-specific biodiversity risk assessments.

B. Risk Aggregation and Mitigation Insights

Following data retrieval, the platform aggregates risk information from various environmental datasets. Threatened or endangered species identified within the 5-mile radius are matched against the IUCN Red List to assess conservation urgency [4]. Marine, freshwater, and terrestrial ecosystem risks are determined using environmental intensity indicators sourced from the World Bank's Global Biodiversity dataset [14]. Additionally, the presence of invasive species is mapped by cross-referencing national and state-level databases, such as those maintained by the USDA and New Jersey authorities [35].

The aggregated results are then organized and presented to the user through a structured display that highlights key elements: risk type (marine, freshwater, IUCN species), involved species, their precise geographic relevance, and suggested mitigation strategies. Where appropriate, the platform supplements static recommendations with dynamic suggestions generated by its integrated Retrieval-Augmented Generation (RAG) model, which sources real-world mitigation actions from an embedded vector database of conservation reports.

C. Report Generation and Export Functionality

To enhance usability and facilitate compliance reporting, BiodivProScope enables users to export biodiversity risk assessments in multiple formats. Generated reports are available in PDF format for official submissions or archival purposes. Additionally, CSV and Excel formats are offered to accommodate internal reviews, external consultant sharing, or further data analysis.

Each downloadable report includes a detailed breakdown of:

- Identified risk types and corresponding species,
- Assigned threat levels (High, Moderate, Low),
- Recommended mitigation strategies, categorized by ecosystem type.

This multi-format export capability ensures that users can effectively document their environmental risk profiles and adopt informed conservation measures tailored to the unique ecological context of each hotel location.

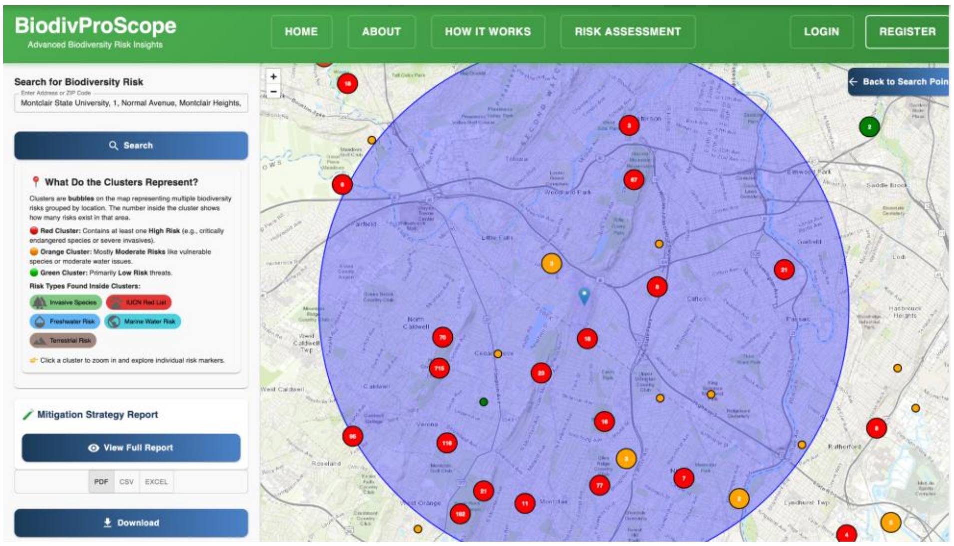
---

8. Results

A. Interface Overview: Interactive Map and Mitigation Reporting

The BiodivProScope platform features a full-width, dynamic map interface (Figure 2) built using Leaflet.js, enabling users to visualize local biodiversity risks surrounding hotel locations. The map supports zooming, panning, and marker clustering to efficiently manage dense risk data points. Custom-designed icons represent different risk types—including IUCN threatened species, invasive organisms, freshwater risks, marine risks, and terrestrial threats—facilitating intuitive understanding of various environmental pressures.

Upon selecting a marker, users can access real-time popups displaying detailed risk information, such as species name, threat level, geographic coordinates, and corresponding mitigation strategies. The popup interface is designed with collapsible accordion-style sections that categorize risks by type and severity to enhance information clarity. The Mitigation Report panel is accessible through the map or the dashboard, presenting numbered mitigation actions grouped by risk type, alongside species-specific recommendations. Export options include PDF downloads, preserving visual and structural consistency with the dashboard's full-screen layout.



"Figure 2: RiskMap interface showing biodiversity risks within a 5-mile radius of a selected hotel location."

B. Functionality Testing: Authentication, Map Interaction, and Reporting

Comprehensive functionality testing was conducted to ensure system reliability across major workflows. The authentication module was validated through successful login and registration flows, employing session-based authentication. Edge cases such as invalid credentials and session expiration were handled via custom alert messages to maintain user experience integrity.

In terms of geospatial functionality, map interaction features were tested by submitting hotel locations using New Jersey zip codes, such as 07663 (Figure 1). The system correctly filtered risks within the designated 5-mile radius using Haversine-based queries. Performance was validated under conditions of frequent zooming, panning, and scroll-based pagination, demonstrating stable load times and seamless user interaction even with dense clusters of species markers.

For reporting, dynamically generated mitigation reports (Figure 2) were validated to include comprehensive data—risk types, affected species, threat levels, geographic coordinates, and recommended mitigation actions—retrieved through PostgreSQL and ChromaDB queries.

Reports were successfully exported in both CSV and Excel formats, offering users flexibility for sustainability audits and internal reviews (Figure 3).

Full Mitigation Action Report

Total Environmental Risks in your area : 267

High Risk

Invasive Species

Freshwater Risk

Terrestrial Risk

1. Terrestrial Risk Level: 1.00 (Latitude: 41.1250, Longitude: -74.1250)

Action: 1. Enforce strict zoning and habitat protection measures. 2. Establish ecological corridors. 3. Implement buffer zones near sensitive areas like wetlands or reserves. 4. Restrict land conversion near protected regions. 5. Conduct monthly biodiversity impact monitoring. 6. Follow NJAC 7:50-6.21(b) for critical wildlife habitats. 7. Implement habitat restoration programs within hotel grounds, including reforestation with native trees. Control soil erosion using natural barriers and maintain corridors for wildlife movement. 8. Partner with conservation organizations to restore degraded lands by planting native trees and removing invasive species. Provide biodiversity education to guests through interpretive signage, nature trails, and interactive garden maps.

Moderate Risk

Invasive Species

1. *Poison hemlock* (Latitude: 40.8279, Longitude: -74.2457)
2. *Japanese maple* (Latitude: 40.8289, Longitude: -74.2801)
3. *Sweet breath of spring* (Latitude: 40.8110, Longitude: -74.2395)
4. *Paper-mulberry* (Latitude: 40.8078, Longitude: -74.1912)

"Figure 3: Sample mitigation report for a location in Montclair, NJ."

C. Summary of Risk Outputs and Mitigation Effectiveness

A biodiversity risk summary table was developed to present aggregated outputs per hotel location. Key columns included Zip Code, Risk Type, Species Count, Threat Level, Sample Species, and Top Mitigation Action. Each row corresponded to a distinct New Jersey hotel location tested during the development phase.

Additionally, mitigation action effectiveness was evaluated by grouping recommendations according to risk type and threat level. The system dynamically retrieved appropriate mitigation entries from ChromaDB, ensuring that the generated strategies were both contextually relevant and actionable. These outputs validated the operational readiness of BiodivProScope in assisting hotel operators with environmental risk compliance and biodiversity conservation efforts.

---

4. Discussion (IX)

A. Usability for Hotel Staff: Balancing Simplicity and Depth

BiodivProScope was specifically designed with non-technical users in mind, including hotel managers and environmental compliance officers. The system's dashboard and interactive map interface prioritize intuitive navigation, enabling users to access biodiversity risk assessments without requiring technical expertise. Risk outputs are presented in natural language, supplemented with easily recognizable icons that minimize interpretation efforts. While this simplicity enhances user adoption and engagement, it introduces a trade-off: users have limited visibility into the technical scoring mechanics underpinning the risk assessments. Although transparent to the backend processes, detailed scoring methodologies may remain abstract to end users unless specifically requested.

Mitigation actions are also structured to maximize clarity. Recommendations are delivered in a direct, step-by-step format, categorized by risk type (IUCN species, invasive species, freshwater risks, marine risks, and terrestrial risks) and by threat level. The interface utilizes an accordion layout and color-coding schemes to enhance rapid comprehension and decision-making.

B. Flexibility in Updating Datasets

The backend infrastructure of BiodivProScope was designed to accommodate periodic updates to biodiversity datasets. The PostgreSQL database schema supports Extract, Transform, Load (ETL) pipelines that allow for the ingestion of updated species lists, pollution metrics, and conservation statuses as they become available. For mitigation strategies, ChromaDB embeddings can be seamlessly updated as new regulatory PDFs or structured JSON files are parsed and embedded.

Additionally, the system's architecture anticipates the need for data governance and maintenance features. Admin-level tools can be implemented to allow backend managers to oversee dataset updates, ensuring that the tool remains compliant with evolving biodiversity regulations, such as those issued by the NJDEP.

C. Relevance of Risk Types for Environmental Compliance

BiodivProScope's risk classification system was designed with a regulation-centered architecture, focusing specifically on New Jersey's biodiversity compliance requirements. Risk types—including invasive species threats, marine pollution risks, and freshwater ecosystem pressures—were selected based on authoritative regulatory documents published by the IUCN, World Bank, and NJDEP.

This regulatory alignment provides added value for hotel users. The platform enables businesses to demonstrate environmental due diligence during sustainability audits, while also generating actionable insights that address real-world biodiversity concerns near their properties. Through its structured risk identification and mitigation framework, BiodivProScope helps hotel operators bridge the gap between policy compliance and ecological responsibility.

D. System Limitations: API Restrictions, Data Granularity, and Deployment Constraints

BiodivProScope currently faces two key limitations. First, some ecological datasets—particularly species distribution records—lack zip-code-level specificity, which may cause biodiversity risks to be mapped slightly beyond their actual locations, requiring careful interpretation by users

Finally, the current system prototype is not optimized for full-scale commercial deployment. While effective for localized use in New Jersey, broader deployment would require additional system enhancements, including the addition of security protocols, load balancing layers, and horizontal scaling capabilities to accommodate increased user concurrency across multiple states.

E. Future Directions: Adaptive Models, Real-Time Integration, and User Feedback

Future enhancements to BiodivProScope aim to incorporate adaptive machine learning models to improve mitigation strategy recommendations. Planned developments include training Random Forest models and Retrieval-Augmented Generation (RAG) pipelines to deliver personalized mitigation actions based on user location, risk history, and evolving environmental contexts.

Additionally, the system may integrate federated API access to real-time biodiversity databases such as GBIF, employing caching strategies to ensure responsiveness without violating licensing constraints. A push notification system is envisioned to alert hotel users to critical updates, such as seasonal pollution events or emerging invasive species threats.

Finally, user feedback integration will play a critical role in iterative system improvements. A feedback and rating mechanism will allow users to evaluate mitigation suggestions, providing valuable data to fine-tune recommendation models and enhance the overall user experience.

---

5. Conclusion (X)

A. Significance of the BiodivProScope Tool

The BiodivProScope project addresses a critical intersection between biodiversity conservation and regulatory business compliance, with a focus on the hospitality sector in New Jersey. By providing an intuitive digital platform capable of identifying local biodiversity risks within a 5-mile radius of hotel locations, BiodivProScope effectively bridges the gap between ecological science and practical decision-making. Its strategic alignment with emerging environmental compliance mandates makes it a timely and impactful contribution to the growing ecosystem of sustainability-focused digital infrastructure tools available to the commercial sector.

B. Major Achievements: Risk Identification, Visualization, and Reporting

The platform successfully integrates geospatial proximity calculations using the Haversine formula to isolate species and ecological threats relevant to hotel-adjacent areas. Biodiversity scoring logic, grounded in species conservation statuses and environmental risk indicators, was developed to prioritize risks into High, Moderate, and Low categories.

An interactive visualization system was built using Leaflet.js, allowing hotel managers to engage with real-time biodiversity risk maps enhanced with cluster markers and risk-specific icons. This design facilitates an intuitive understanding of local environmental vulnerabilities. Furthermore, the system generates dynamic mitigation reports, exportable in PDF, CSV, and Excel formats. These reports were structured to support regulatory reporting requirements, offering hotel operators a streamlined approach to documenting their biodiversity compliance efforts.

C. Contribution to Hotel Sustainability Compliance

Beyond risk identification, BiodivProScope provides proactive mitigation strategies that empower hotel management teams to act responsibly in preserving surrounding ecosystems. By automating biodiversity risk assessments and aligning mitigation guidance with New Jersey Department of Environmental Protection (NJDEP) recommendations, the platform supports the broader sustainable transformation of the hospitality industry. It delivers transparency, regulatory accountability, and actionable insights, essential for fostering trust with both government regulators and environmentally conscious clientele.

D. Recommendations for Future Development

Building on the current system, future enhancements should focus on improving the spatial resolution of the underlying datasets. Incorporating higher-resolution species presence data at the sub-zip code or site-specific level would enable more accurate local risk assessments. Partnerships with local environmental agencies and conservation organizations could facilitate access to finer-grained ecological data.

To address API access and licensing limitations, future iterations should explore federated querying approaches using open-source databases such as the Global Biodiversity Information Facility (GBIF), supported by rate-limiting and caching mechanisms for efficient real-time operations.

The mitigation recommendation engine should evolve from its current rule-based framework toward a fully adaptive machine learning-driven model, incorporating feedback loops and real-time ecological context to deliver personalized, context-sensitive conservation strategies. Finally, expanding BiodivProScope's scope beyond New Jersey to other U.S. states—or even internationally—would involve modularizing the system's architecture for easier adaptation to varying regional biodiversity compliance frameworks and industry-specific environmental regulations.

6. Acknowledgements (XI)

The authors sincerely thank their instructor, Dr. Vaibhav Anu, for his invaluable guidance and direction throughout the course of this project. The authors also gratefully acknowledge their sponsors, Dr. Pankaj Lal, Dr. Manveer Mann, Kripa Shrestha, and Cassidy Rucker, for their contributions to this research and for their continuous support in making this project possible.

7. References

[1] D. Díaz et al., "Biodiversity loss: A threat to human well-being," *Science*, vol. 356, no. 6335, pp. 270–275, 2017.

[2] Secretariat of the Convention on Biological Diversity, *Global Biodiversity Outlook 5*, Montréal, QC, 2020.

[3] New Jersey Department of Environmental Protection, *Hotel and Lodging Sustainability Guidance for Biodiversity Awareness*, Trenton, NJ, 2022.

[4] International Union for Conservation of Nature (IUCN), *The IUCN Red List of Threatened Species, 2024*

[5] US EPA, "EnviroMapper for Envirofacts," 2024.

[6] GBIF, "Global Biodiversity Information Facility," 2024.

[7] R. Froese and K. Kesner-Reyes, "AquaMaps: Predictive Maps of Fish Species Distribution," 2020.

[8] V. Agafonkin, "Leaflet.js," 2024.

[9] P. Ramsey, "PostGIS: Spatial Database Extension for PostgreSQL," OpenGeoTech, 2020.

[10] A. M. MacEachren et al., "Visualizing Geospatial Information Uncertainty," *Cartography and Geographic Information Science*, vol. 32, no. 3, pp. 143–160, 2005.

[11] M. Goodchild, "Citizens as sensors: the world of volunteered geography," *GeoJournal*, vol. 69, no. 4, pp. 211–221, 2007.

[12] K. S. Davis et al., "Citizen Science GIS Projects for Environmental Monitoring," *Environmental Practice*, vol. 19, no. 3, pp. 133–147, 2017.

[13] B. S. Halpern et al., "Cumulative human impacts on marine ecosystems," *Science*, vol. 319, no. 5865, pp. 948–952, 2008.

[14] World Bank, "Global Biodiversity Dataset," 2024.

[15] IUCN, *Ecological Risk and Conservation Policy*, Report No. REP-2011-015, 2011.

[16] Green Key Global, "Green Key Eco-Rating Program," 2024.

[17] EarthCheck, "EarthCheck Certified Program," 2024.

[18] U.S. Department of Agriculture (USDA), *National Invasive Species Information Center*

[19] New Jersey Department of Environmental Protection (NJDEP), *Water Monitoring and Standards Program Data*, Trenton, NJ, 2021.

[20] C. Venter et al., “Global terrestrial Human Footprint maps for 1993 and 2009,” *Scientific Data*, vol. 3, no. 1, pp. 1–10, 2016.

[21] New Jersey Future, *Planning for Biodiversity at the Municipal Scale*, 2018

[22] OpenStreetMap Foundation, “Nominatim,”

[23] Chroma, “ChromaDB: Open-source Embedding Database for AI Applications

[24] React Documentation, “React – A JavaScript library for building user interfaces.”

[25] Flask Documentation, “Flask: Python Microframework.”

[26] PostgreSQL Global Development Group, “PostgreSQL: The World’s Most Advanced Open Source Relational Database.”

[27] Vladimir Agafonkin, “Leaflet: an open-source JavaScript library for mobile-friendly interactive maps.”

[28] GitHub, “GitHub Docs – Collaborative software development platform.

[29] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[30] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks,” in *Proc. of EMNLP*, 2019.

[31] New Jersey Department of Environmental Protection (NJDEP), *Guidelines for Biodiversity and Habitat Mitigation in Hotel Operations*, Trenton, NJ, 2022.

[32] R. Sinnott, *Geographic Information Systems and Science*, 3rd ed. Hoboken, NJ: Wiley, 2011, ch. 5.

[33] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[34] P. Lewis et al., “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,” in *Proc. 34th Conf. Neural Information Processing Systems (NeurIPS)*, Vancouver, Canada, 2020.

[35] Beyond Green, *Global Impact Report 2023*, Post Ranch Inn, Dec. 2023.

[36] T. L. Saaty, *The Analytic Hierarchy Process*. New York, NY, USA: McGraw-Hill, 1980.

[37] O. S. Vaidya and S. Kumar, “Analytic hierarchy process: An overview of applications,” *European Journal of Operational Research*, vol. 169, no. 1, pp. 1–29, Feb. 2006.