Piscis in tabula: Turning Environmental Data into Actionable Reports

Beatrice Portelli[1,2,*], Alex Falcon[1], Daniele Lizzio Bosco[1,3], Paolo Tomè[2], Marco Galeotti[2], Snježana Zrnčić[4] and Giuseppe Serra[1]

[1]*University of Udine, Department of Mathematics, Computer Science and Physics, Italy*
[2]*University of Udine, Department of Agriculture, Animal, Environmental and Food Sciences, Italy*
[3]*University of Naples Federico II, Department of Biology, Italy*
[4]*Croatian Veterinary Institute Zagreb, Croatia*

Abstract

Modern mariculture operations generate large volumes of heterogeneous data, ranging from *in-situ* sensor readings and management logs to satellite imagery and meteorological forecasts. Although this information is essential for monitoring farm conditions and supporting decisions, it is often not readily available to operators. When available, it is typically conveyed through dashboards or raw tables that require substantial expertise and time to interpret. In this work, we introduce a web-based platform that organizes multimodal mariculture data as a vertical digital library, providing unified access, basic metadata organization, and integrated analytical tools. Beyond data aggregation, the system incorporates an AI Reporting Agent (AIRA) module that produces periodic narrative reports summarizing past conditions, highlighting relevant patterns, interpreting potential risks, and offering forward-looking suggestions informed by predictive models. We describe the architecture of the digital library and present early components of the proposed system, with a focus on the forecasting deep learning models and the report generation module. Preliminary quantitative experiments illustrate the feasibility of key analytical modules, while qualitative examples outline the potential of narrative reports to support operator understanding. This work positions digital libraries not only as repositories but as knowledge mediators, translating complex environmental datasets into actionable, human-centered narratives.

Keywords

digital libraries, mariculture, language generation, deep learning

1. Introduction

Environmental monitoring systems increasingly rely on heterogeneous and continuously expanding data sources, ranging from *in situ* measurements and laboratory analyses to the outputs of remote sensing systems and historical datasets [1, 2]. In fisheries and aquatic resource management, these data streams provide essential information for assessing the welfare of the animals, evaluating environmental pressures, and supporting operational decision-making [3, 4]. However, in practice, such data are often dispersed across systems, recorded at different temporal and spatial resolutions, inconsistently standardized, and difficult to interpret rapidly. As a result, operators frequently rely on manually crafted summaries, ad hoc interpretations of dashboards, and expert judgment, making routine reporting both time-consuming and variable in quality.

These challenges are particularly evident in marine fish farming, one of the world's fastest-growing food production sectors, including in the Mediterranean basin [5]. Its rapid expansion is accompanied by environmental and socio-economic constraints that may hinder long-term sustainability [6]. Among

IRCDL 2026: 22nd Conference on Information and Research Science Connecting to Digital and Library Science, February 19–20, 2026, Modena, IT

*Corresponding author.
portelli.beatrice@spes.uniud.it (B. Portelli); falcon.alex@spes.uniud.it (A. Falcon); lizziobosco.daniele@spes.uniud.it (D. Lizzio Bosco); paolo.tome@uniud.it (P. Tomè); marco.galeotti@uniud.it (M. Galeotti); zrncic@veinst.hr (S. Zrnčić); giuseppe.serra@uniud.it (G. Serra)
0000-0001-8887-616X (B. Portelli); 0000-0002-6325-9066 (A. Falcon); 0009-0002-7372-6518 (D. Lizzio Bosco); 0000-0003-4486-1403 (P. Tomè); 0000-0003-1378-8135 (M. Galeotti); 0000-0003-4084-5641 (S. Zrnčić); 0000-0002-4269-4501 (G. Serra)

the most critical issues are health and welfare risks linked to the emergence and spread of pathogens [7]. The Adriatic Sea follows the same trajectory, facing similar pressures as farms operate under increasingly variable environmental conditions. To address these issues, the Interreg project MARINET[1] aims to enhance the sustainability and resilience of Adriatic mariculture through continuous environmental monitoring, early detection of welfare risks, and the implementation of effective mitigation strategies. MARINET combines sensor networks, remote sensing, operational indicators, and diagnostic sampling to produce comprehensive, multi-scale environmental information for farm operators. However, while these data streams are increasingly available, synthesizing them into decision-ready knowledge remains a major bottleneck.

Similar challenges have been explored in the field of digital libraries, where large collections of heterogeneous digital objects must be made interoperable and intelligible through standardized metadata, controlled vocabularies, and structured indexing mechanisms. These systems demonstrate how data curation and good representation can transform raw digital artifacts into navigable, interpretable resources. Adopting analogous principles can help harmonize dispersed environmental data streams and support more automated, reliable decision-making workflows.

The present work addresses the gap between raw environmental data and domain-specific reporting by introducing an AI-assisted framework, called AIRA (AI Reporting Agent), capable of automatically generating structured environmental summaries tailored for fisheries management contexts. The proposed system integrates environmental time series with domain knowledge to produce concise, human-readable reports that highlight key conditions, deviations from expected ranges, and potential operational implications. The goal is not to replace expert interpretation but to reduce cognitive load, standardize routine assessments, and facilitate faster understanding.

The contributions of this work are the following:

- we develop a multimodal platform, inspired by digital libraries, for organizing, curating, and accessing environmental datasets collected in mariculture settings;
- we propose AIRA, an AI Reporting Agent that complements dashboards by providing narrative access to environmental information, transforming heterogeneous environmental signals into coherent narrative summaries, and offering a higher level of interpretability than dashboards alone;
- we demonstrate a prototype implementation and report quantitative evaluations of selected functional components, including predictive trend forecasting, and the generation of concise reports through AIRA.

The resulting methodology offers a scalable solution for transforming environmental measurements into decision-ready information. Although developed with fisheries applications in mind, the underlying principles are generalizable to other domains where timely synthesis of multi-source data is critical. This paper presents the conceptual design, implementation details, and evaluation of the proposed reporting system, highlighting its potential to streamline routine monitoring workflows and enhance situational awareness.

2. Related Work

Digital libraries have long provided structured frameworks for managing large collections of heterogeneous scientific [8] and environmental datasets [9]. Early systems emphasized standardized metadata, persistent identifiers, and robust indexing mechanisms to support reusability and long-term preservation [10, 11]. In the environmental sciences, digital libraries have been used to support data curation, interoperability, and discovery across distributed sensing infrastructures [12, 13]. These systems demonstrate the value of viewing environmental data as curated digital objects embedded in a broader knowledge infrastructure.

[1]https://www.italy-croatia.eu/web/marinet

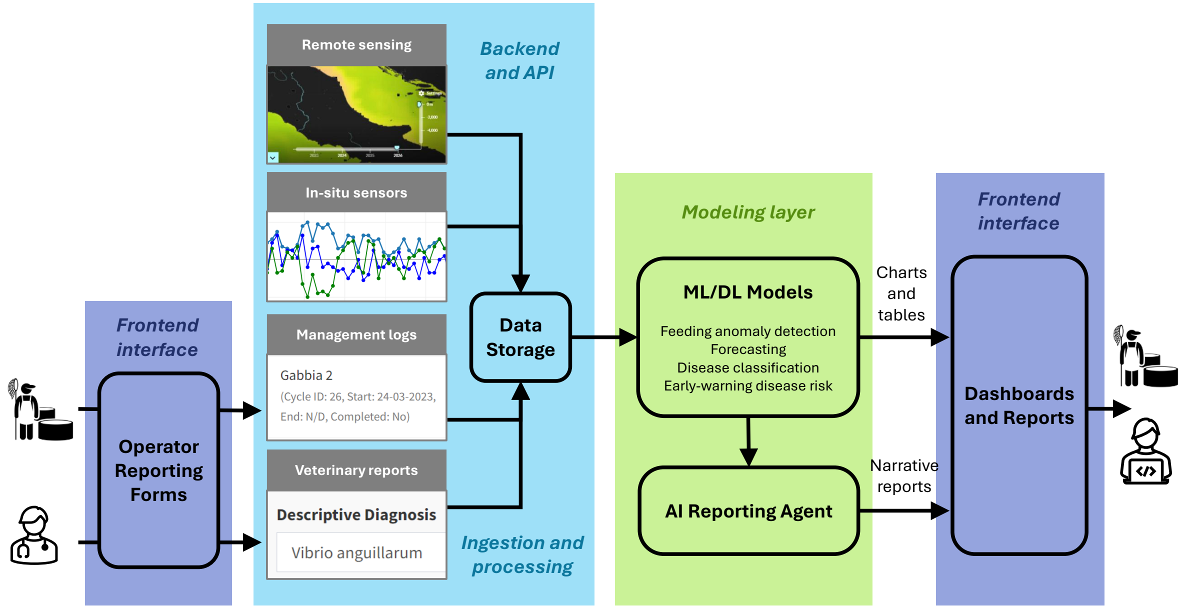


Figure 1: Overview of the proposed system highlighting the main architectural components.

Environmental monitoring and decision-support platforms have similarly tried to integrate multi-source data for marine and aquaculture applications. For example, integrated marine information systems have been developed to combine heterogeneous sensor streams and analysis tools for environmental monitoring and hazard assessment [14], and IoT-based decision support systems have been explored for real-time control and prediction tasks in fish farming [15]. These systems often provide dashboards, risk indices, or alerts that enhance situational awareness but largely rely on visual or tabular summaries of data, requiring domain expertise for interpretation.

Natural Language Generation (NLG) has been gaining increasing interest in the last few years thanks to the recent advances in the capabilities of Large Language Models (LLMs) and their availability to the general public (e.g. ChatGPT [16], Gemini [17]). Several studies focus on means of translating structured data into human-readable summaries [18, 19]. Early work in data-to-text generation applied NLG techniques to meteorological reporting and environmental indicators, demonstrating the value of linguistic descriptions of time series data [20]. More generally, text summarization and data-to-text tasks have been reviewed in the context of deep learning, identifying both the potential and the challenges in automatically generating textual summaries from structured inputs [21]. While modern LLM-based approaches may allow for flexible reporting, operational systems should rely on controlled, template-based generation to ensure factuality and consistency, particularly in domains with safety or regulatory implications.

To our knowledge, no prior system integrates multimodal environmental data management with interpretable AI-assisted narrative summaries tailored for mariculture. The framework presented in this work aims to bridge this gap by unifying structured data curation with lightweight, context-specific NLG to support operational decision-making.

3. System Overview: A Digital Library for Mariculture

The proposed platform organizes multimodal mariculture data into a unified, library-inspired digital infrastructure designed to support monitoring, interpretation, and narrative reporting. The system integrates heterogeneous data sources, harmonizes them through a shared metadata schema, and exposes curated and aligned datasets to both analytical modules and the AI Reporting Agent (AIRA) pipeline. Figure 1 summarizes the main components.

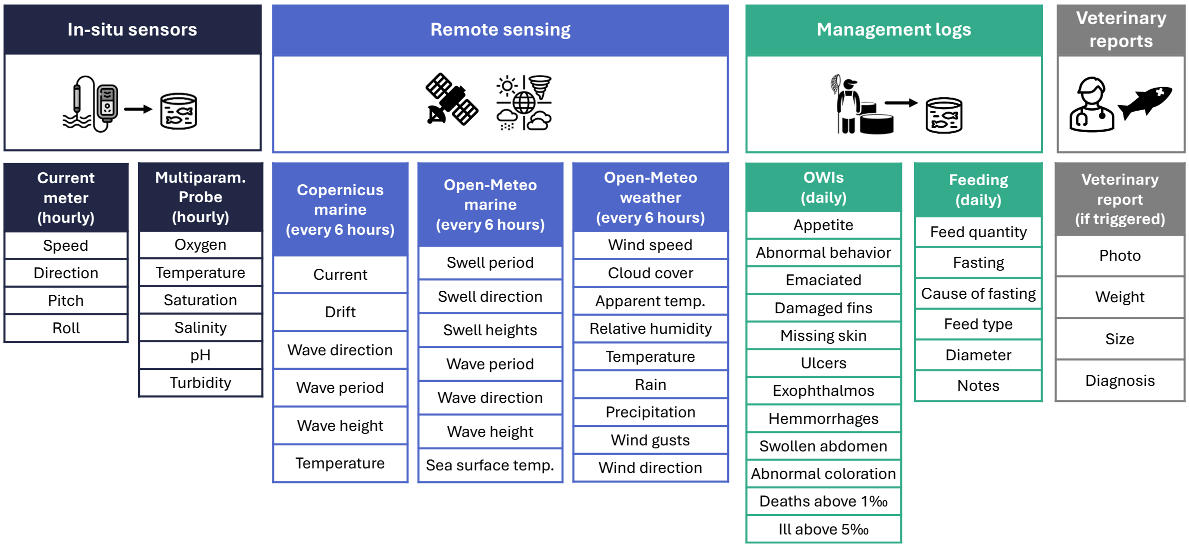


Figure 2: Overview of the main data sources (*in-situ* sensors, remote sensing, management logs, veterinary reports) and associated fields.

3.1. Data Sources

The platform consolidates a broad set of operational and environmental data routinely produced in mariculture settings. An overview of the main sources and associated data fields is presented in Figure 2. The system currently integrates four principal categories of information:

- **measurements from *in-situ* sensors**, such as multiparametric probes (dissolved oxygen, temperature, saturation, salinity, pH, turbidity) and current meters, providing high-frequency environmental measurements. These data are collected **hourly**, offering precise, cage-level insight into local environmental conditions.
- **remote sensing data**, consisting of meteorological and climate data, collected **every 6 hours**. These data include satellite-derived oceanographic variables from Copernicus services [22], as well as marine and weather information from Open-Meteo [23]. These sources complement *in-situ* measurements with broader spatial context and short- to medium-term environmental outlooks.
- operational and **management logs**: **daily** records of feeding, mortality counts, corrective actions, and structured Operational Welfare Indicators (OWI). These entries, collected via a form-based interface (web form), capture the perceived health status of the fish, abnormal behaviours, and any interventions performed at cage level.
- **veterinary sampling and imaging**: periodic fish-level examinations generate diagnostic reports and a catalog of images used for AI-based lesion analysis. Sampling events are triggered when daily inspections indicate a significant increase in morbidity or mortality (see OWIs, “Deaths above 1‰” and “Ill above 5‰”).

The resulting dataset is highly heterogeneous in format, scale, and resolution. The system addresses this through a dedicated pipeline that performs cleaning, harmonization, and temporal alignment, ensuring that data layers can be jointly analyzed.

3.2. Metadata and versioning

To support discoverability and long-term curation, the platform adopts a data model inspired by digital-library principles. Each resource (e.g., sensor time series, a sampling report, or satellite-derived image) is treated as a structured digital object enriched with descriptive metadata (source, measurement units), administrative metadata (provenance, ingestion timestamp), and semantic metadata (links to fishfarm location, cage, or specific events).

This schema enhances reusability and facilitates cross-modal linking, enabling queries such as aligning environmental anomalies with welfare indicators or contextualizing veterinary diagnoses with preceding environmental conditions.

All data are versioned, allowing analyses and generated reports to refer to stable historical snapshots.

3.3. Platform Architecture

The platform is implemented as a Progressive Web Application (PWA) backed by a structured relational database. The adoption of a PWA allows deployment in heterogeneous and resource-constrained environments, supporting cross-device access and facilitating routine use by non-technical operators while maintaining a unified entry point to the underlying digital library. The architecture comprises the following four main components.

Backend and API layer. Built on a MariaDB DBMS accessed through a Symfony/Doctrine stack, ensuring ACID compliance, referential integrity, and efficient handling of large structured datasets. A REST API provides secure access to curated data, metadata, and analytical endpoints.

Ingestion and processing pipeline. Automated routines manage data acquisition from sensors, external APIs (Copernicus, Open-Meteo), operator-submitted forms, and sampling uploads. The pipeline enforces validation, harmonization, metadata enrichment, and versioning, ensuring that all incoming resources conform to the platform's data model.

Modeling layer. The modeling layer provides analytical and predictive capabilities built on top of the data repository. It currently incorporates a variety of modules (in various stages of development), including:

- feeding anomaly detection, implemented through classical machine learning models such as Logistic Regression, Random Forests, and Support Vector Machines;
- short-term forecasting of *in-situ* environmental conditions, based on lightweight statistical approaches (e.g., ARIMA, SARIMA, and SARIMAX) and zero-shot inference from foundational models (e.g., Prophet [24], Chronos2 [25]);
- image-based disease classification, using compact convolutional neural networks derived from ResNet-18 [26] for lesion detection and health-status assessment based on the sampled fish;
- early-warning disease risk estimation, which integrates environmental parameters with OWIs and cage-level symptoms to estimate the probability of specific disease conditions emerging or developing. This module, based on classical machine learning models, combines environmental patterns with phenotypic signals to highlight the most plausible disease scenarios and support timely intervention.

These components produce structured outputs (such as anomaly scores, trend descriptors, or risk indicators) that can be directly displayed to the operators via dashboards, or further processed by AIRA for generating narrative reports.

The next sections will focus on the forecasting component, which is currently the most production-ready, and demonstrate its interaction with AIRA as an example of how analytical models interface with downstream reporting tools.

Frontend interface for operators and analysts. The cross-platform PWA offers dashboards for real-time monitoring, tools for exploratory data analysis, and direct access to automatically generated narrative reports. The interface is optimized for non-expert users in operational settings, balancing readability with the ability to inspect details when needed. Beyond data visualization, the PWA acts as an operational interface for controlled data ingestion and consistency checks, enabling farm operators to record structured observations, management actions, and welfare indicators directly in the field, thus ensuring coherence and traceability across the entire data lifecycle.

4. Technical Details

To evaluate the effectiveness of the forecasting component and illustrate its interaction with the AIRA, we conduct a series of experiments using a subset of data extracted from the MARINET platform. We focus on measurements acquired by the *in-situ* multiparametric probes installed at three representative fish-farm sites (hereafter LOC1, LOC2, and LOC3). The data covered a period of 292 days for LOC1 and LOC2 (from February 2 to November 25, 2025) and 262 days for LOC3 (from March 5 to November 21, 2025).

For each site, we consider the following environmental variables to predict: dissolved oxygen, pH, salinity, oxygen saturation, temperature, and turbidity. In addition, we included the readings of the current meter as additional input-only variables (current speed, direction, pitch, and roll).

Since the sensors were sampled at slightly different timestamps and contained occasional missing values, all time series were preprocessed through a uniform resampling procedure. The raw measurements were aligned to a regular hourly grid and missing values were filled via interpolation. Each multivariate series was then divided into blocks of five consecutive weeks. The first four weeks were used as historical context for the models, while the fifth week was used for evaluation, and represents the forecasting target.

4.1. Forecasting Module

To evaluate the forecasting component of the system, we employ Chronos-2 [25], a time-series foundational model capable of producing short-term predictions from brief multivariate context windows. Chronos-2 supports *zero-shot* forecasting: it can generate future trajectories without any task-specific fine-tuning, relying solely on representations learned during pre-training on large and diverse time-series corpora.

This property is particularly valuable in operational mariculture settings, where locally available training data are often scarce, site-specific, and updated irregularly. Under these constraints, retraining or fine-tuning large models on farm-level data can lead to overfitting, model drift, or inconsistent behaviour across sites. For these reasons, our methodology deliberately relies on Chronos-2 in *zero-shot* mode, which ensures stable performance while remaining compatible with realistic deployment scenarios.

4.2. AI Reporting Agent (AIRA)

AIRA constitutes the natural-language layer of the proposed system and is responsible for transforming numerical model outputs into concise, operator-friendly narrative summaries. In this work, AIRA is implemented using the Gemini 3 Pro model[2], combined with a template-guided prompting strategy to ensure factual consistency and controlled vocabulary.

AIRA operates downstream of the other models in the platform's "Modeling layer", using several tailored prompts that ingest relevant data and generate topical sub-reports. More broadly, AIRA is designed to reduce the cognitive load associated with interpreting heterogeneous datasets and model-derived indicators. While dashboards and visualizations already facilitate access to individual signals, operators must still integrate multiple sources of data (e.g., meteorological forecasts, OWIs, and occasional laboratory sampling or imaging results). AIRA consolidates these inputs into coherent textual overviews that highlight current conditions, contextualize recent events, and anticipate emerging risks.

In this work, we focus on its interaction with the forecasting module. Given the short-term predictions generated by Chronos2, and the relevant contextual metadata (such as site-specific sensor threshold), AIRA synthesizes these elements into structured narrative reports. This interaction provides a concrete example of how analytical components integrate with reporting tools within the MARINET platform.

In the present study, AIRA is evaluated in two roles: as a consumer of model outputs, where it must correctly interpret the trends and magnitudes produced by Chronos2; and as a reporting agent,

[2]https://deepmind.google/models/gemini/pro/

responsible for generating short, understandable summaries for the human operators.

The generated reports are structured around three core functions:

1. Current status and trends. AIRA summarizes the present state of the monitored variables, identifies anomalies using expert thresholds, and describes short-term trends extracted from high-frequency sensor data.
2. Notable recent events. It highlights relevant occurrences from the past week, such as welfare observations or corrective actions, linking them to temporally aligned environmental measurements when appropriate.
3. Early warnings for upcoming conditions. Using the predictions from the forecasting module, AIRA provides short-term anticipatory insights, helping operators plan interventions before conditions deteriorate.

Building on recent advances in LLMs, different prompts are designed to query the LLM and generate sub-reports that remain consistent with the underlying data while simplifying their interpretation for fish farmers. In this section, we present the template developed specifically for producing a sub-report focused on forecasting and analyzing the conditions expected for the upcoming week. The prompt, shown in Figure 3, incorporates several key elements, including a description of the input data and their structure, the expert-defined thresholds associated with each sensor, a clear specification of the task, and detailed instructions on how the output should be structured.

5. Preliminary Results

The following preliminary results evaluate the short-term forecasting capabilities of Chronos2 and illustrate how AIRA converts these predictions into structured, human-readable summaries.

5.1. Forecasting Module

We evaluate Chronos2 in terms of Mean Absolute Percentage Error (MAPE), averaged across all timestamps within the horizon and across all time blocks. Because forecasting accuracy typically decreases as the prediction horizon increases, we report MAPE at three horizons: one, three, and seven days. Assessing long-range predictions is particularly important for evaluating trend-level consistency (i.e., whether the model correctly captures rising, decreasing, or stable patterns), as these qualitative descriptors are directly consumed by AIRA during report generation. The proposed system provides both predictions and their associated confidence intervals, which can also be shown to human operators in the dashboards. Figures 4a and 4b show examples of the model's predictions including the confidence intervals.

Table 1 reports the MAPE across the six target sensors, the three experimental sites (LOC1, LOC2, and LOC3) and three prediction horizons. Overall, Chronos2 achieves low MAPE for most probes, often below 4% at 1 and 3 days and remaining below 1% for pH even at a seven-day horizon. These results indicate that the model provides stable and reliable short-term forecasts for key environmental indicators such as oxygen, pH, temperature, and saturation.

As expected, performance decreases for variables that exhibit higher volatility or are prone to measurement noise. Notably, salinity at LOC1 and turbidity at LOC1 and LOC2 show higher errors, especially for the longest horizon. In the case of turbidity, this degradation aligns with known sensor behavior: turbidity probes require frequent manual cleaning, and when maintenance is irregular the readings may drift or become unreliable. This introduces additional noise into the historical data and consequently reduces forecasting accuracy. Nonetheless, even under such conditions, the model captures general qualitative trends, which remain sufficient for downstream narrative generation.

5.2. AIRA Output

Building on the short-term forecasts and associated confidence intervals presented in the previous section, we evaluated how AIRA translates these quantitative predictions into operator-friendly narrative

You are an expert data analyst specializing in environmental monitoring for fish farms. I will provide three inputs:
1. Historical data (28 days, hourly): A CSV-like table where each column is a variable measured by probes (e.g., temperature, salinity, oxygen, pH, etc.), and each row corresponds to a timestamp.
2. Forecasted data (7 days, hourly): A CSV-like table with the same variables and structure, but containing forecasted values.
3. Variable thresholds: For each variable, four numbers (c, l, g, h) describing six situations depending on the measured values (v):
- If v is lower than c: critical situation
- If v is within c and l: low value, should be checked
- If v is within l and g: good value
- If v is within g and h: high value, should be checked
- If v is greater than h: critical situation

[**Your task**]
Using all three inputs, produce a concise, actionable report for fish farmers. To create this report, you must analyze:
1. Status and trends in the historical data
- Identify recent trends, anomalies, or abrupt changes for each variable.
- Highlight any slow drifts that might indicate an approaching problem.
2. Forecast analysis for the next 7 days
- Highlight the timing of any expected threshold crossings.
- Identify potential risks early (e.g., continuous warming trend leading to critical temperature on day 5).
3. Recommended preventive actions
- If any variable risks entering a critical range, propose specific interventions fish farmers can take now (e.g., increase aeration, adjust water intake, reduce feeding, etc.).
- Include short explanations of why these actions matter.

[**Output format**]
Provide a structured report with the following sections:
- Recent Status & Trends
- Risks & Threshold Crossings
- Recommended Actions
Keep the report short, clear, and directly usable by fish farm managers.

[**Input**]
Historical data provided in file: { *historical_data.csv* }
Forecasted data provided in file: { *forecasted_data.csv* }
Variable thresholds provided in file: { *thresholds.csv* }

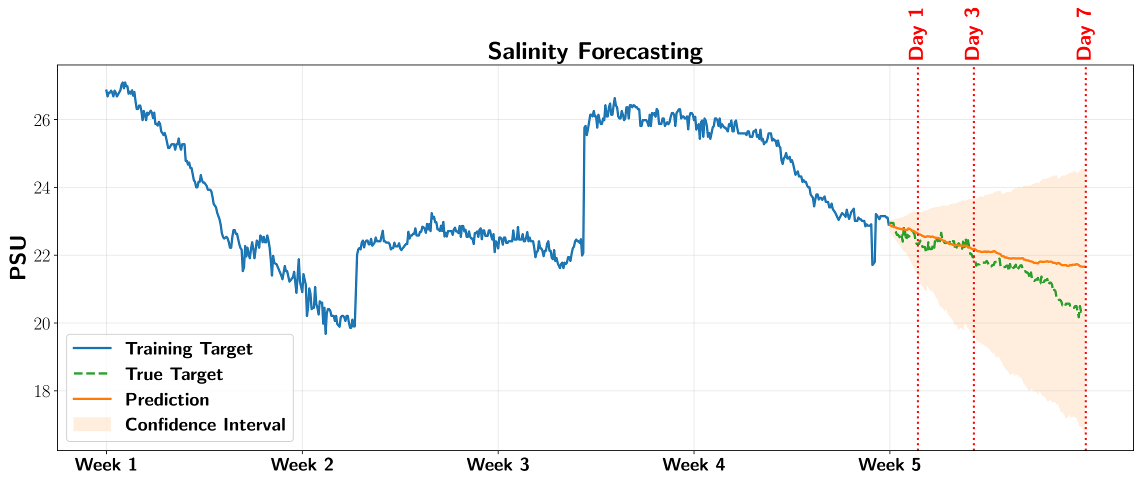
Figure 3: Prompt used by AIRA for forecasting analysis. *historical_data.csv*, *forecasted_data.csv*, and *thresholds.csv* are the paths to the input files.

reports.

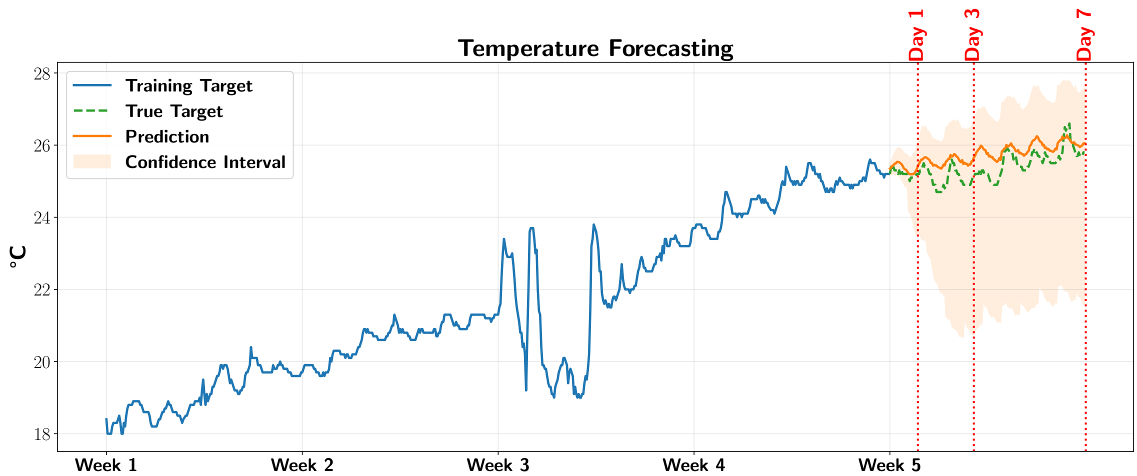
To evaluate report quality, we conducted a manual analysis on a subset of the reports. We verified that each report respected the following conditions: accurately describing the past status of the sensors, highlighting undesired trends relative to expert-defined thresholds, and including recommended actions that are meaningful and operationally relevant.

For each block of probe measurements, the forecasting module was used to predict the behavior of each variable over the following seven days. These forecasts, together with the expert-defined thresholds, were provided to AIRA via the prompt shown in Figure 3. An example of a generated report, based on the first block of measurements from LOC3, is presented in Figure 5. The example illustrates how AIRA synthesizes historical observations, predicted trends, and thresholds into a concise, human-readable summary suitable for operational decision-making.

The resulting report immediately highlights two key issues in the past weeks of measurements: persistently low water temperatures and an anomalous fluctuation in pH levels. Both findings are consistent with the data: the average temperature is 14.2°C (standard deviation: 0.3), which experts classify as low (12-16°C). The pH readings also contain anomalies, including 56 values of 0 and 1 value of 14, while the critical thresholds are 7.4 (low) and 8.7 (high). In addition, the report correctly identifies



(a)



(b)

Figure 4: Time series prediction of salinity and temperature over a one-week horizon with confidence intervals. The shaded area represents the model's predicted uncertainty. (a) predicted salinity with confidence intervals; (b) predicted temperature with confidence intervals. These visualizations demonstrate the ability of the system to provide accurate forecasts and early warnings, especially in the short term.

the main concern in the forecast, namely that water temperature is not expected to increase over the next several days, potentially prolonging suboptimal conditions for the fish. Finally, the suggested actions are appropriate: low temperatures often reduce feeding activity, increasing the risk of uneaten feed degrading water quality, while the extreme pH values observed in the historical data are likely caused by momentary sensor malfunctions and warrant routine inspection of the probes.

Similarly, all analyzed reports complied with our qualitative evaluation protocol, demonstrating that the combination of detailed prompts and robust foundational models can reliably transform quantitative forecasts and heterogeneous inputs into coherent, actionable narrative summaries.

6. Conclusions and Future Work

This work introduced an AI-assisted framework for generating structured environmental summaries that integrate heterogeneous time-series data, domain thresholds, and predictive modelling outputs. This work contributes to the broader goals of the MARINET project, which integrates heterogeneous

Table 1
Mean Absolute Percentage Error (MAPE) of time series predictions for each variable across different sites (LOC1, LOC2, LOC3) and prediction horizons (1, 3, and 7 days). Lower values are better.

variable	oxygen	pH	salinity	saturation	temperature	turbidity
LOC1 @1	2.448	0.138	3.339	2.313	1.311	110.713
LOC1 @3	2.672	0.218	9.185	2.132	2.045	961.340
LOC1 @7	5.376	0.274	3335.187	2.940	3.390	686.350
LOC2 @1	3.386	0.564	2.505	3.582	1.053	31.340
LOC2 @3	6.193	0.853	7.304	6.791	2.343	591.896
LOC2 @7	9.193	1.199	14.800	8.110	5.209	748.540
LOC3 @1	1.633	0.215	0.639	1.777	0.671	27.643
LOC3 @3	2.789	0.296	0.883	3.060	1.168	44.419
LOC3 @7	3.893	0.307	1.457	4.207	2.094	59.894

1. Recent Status & Trends
- **Temperature (Chronic Issue)**: Water temperature has been consistently in the **Warning/Low** range (13.5°C – 14.5°C) for the past 28 days. While there is a very slow warming trend, it has not reached the optimal range (>16°C).
- **pH (Anomaly Detected)**: Major anomalies were detected in the historical data (approx. 10-12 days ago), where values fluctuated to impossible extremes (0 and 14). While the current value is stable at 8.25 (Good), the history suggests a sensor malfunction.
- **Other Variables**: Oxygen, Salinity, Saturation, and Turbidity are stable and currently within Good ranges. Oxygen levels show a healthy slight increasing trend.

2. Risks & Threshold Crossings (Next 7 Days)
- **Persistent Low Temperature**: The forecast indicates temperature will remain in the **Warning/Low** category (14.5°C – 14.7°C) for the entire 7-day period.
Timing: Continuous (Day 1 through Day 7).
Risk: No critical drop (<12°C) is predicted, but the water will fail to reach the growth-optimal threshold (>16°C).
- **No Critical Emergencies**: No variable is forecasted to cross into a Critical zone (e.g., Oxygen < 4.5 mg/L or pH < 7.4).

3. Recommended Actions
- **Manage Feeding (Priority: High)**
Action: Reduce feeding rates immediately.
Reason: At ~14.5°C, fish metabolism is significantly slower than at optimal temperatures. Excess feed will go uneaten, wasting money and potentially spoiling water quality (ammonia spikes).
- **Verify Sensors (Priority: High)**
Action: Inspect and calibrate pH probes.
Reason: The extreme historical readings (pH 0-14) likely indicate sensor failure. Confirm that the current "Good" reading is accurate and not a sensor artifact.
- **Health Surveillance**
Action: Monitor stock for signs of cold-water stress or fungal infections.
Reason: Prolonged exposure to suboptimal temperatures can suppress the fish immune system.

Figure 5: Output of AIRA on a forecasting scenario

data streams, ranging from sensor measurements to remote sensing and manual veterinary sampling. By combining short-term forecasting with controlled natural-language generation, the system aims to reduce the cognitive load associated with interpreting complex environmental data streams and to support actionable decision-making in aquaculture settings.

The experimental evidence demonstrates the feasibility of the approach. Chronos2 showed promising performance in short-term forecasting of water-quality indicators, while AIRA effectively synthesized historical measurements, predicted trends, and expert thresholds into coherent textual reports. The resulting narratives accurately describe current conditions, highlight deviations from expected ranges, and outline plausible interventions. These findings show the potential of LLM-based reporting to support real-world aquaculture operations, combining templates, domain rules, and LLM-based reasoning.

Several research directions emerge from this work. Future efforts within the MARINET project will focus on expanding the input modalities of the AIRA modules, integrating all the other data available in the framework (e.g., OWIs and satellite data). In addition we plan to incorporate operator-in-the-loop feedback mechanisms, to ensure that the increasingly detailed generated reports align with the practical needs and expertise of field operators. Finally, we intend to extend the system toward multilingual reporting to support the diverse workforce of Adriatic mariculture and facilitate broader adoption.

In conclusion, this contribution aligns with the broader vision of narrative digital libraries for environmental monitoring [9], where heterogeneous data streams are organized, interpreted, and transformed into coherent textual knowledge. Although the present study focuses on aquaculture, the underlying approach could be extended to assist field operators in other environmental domains, such as forestry and water management, providing concise, actionable insights from complex, multi-source data.

Acknowledgments

This work was supported by INTERREG VI-A "MARINET - MARICULTURE NETWORK: Implementazione di nuove tecnologie per un'acquacoltura diversificata e sostenibile rivolta a una società sana e a regioni competitive" - ID: ITHR0200334 - CUP: G33C23000660007, and by the Department Strategic Plan (PSD) of the University of Udine – Interdepartmental Project on Artificial Intelligence (2020-25).

Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT-5 in order to: Improve writing style. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] R. Massey, L. T. Berner, A. C. Foster, S. J. Goetz, U. Vepakomma, Remote sensing tools for monitoring forests and tracking their dynamics, in: Boreal forests in the face of climate change: Sustainable management, Springer, 2023, pp. 637–655.

[2] M. Pause, C. Schweitzer, M. Rosenthal, V. Keuck, J. Bumberger, P. Dietrich, M. Heurich, A. Jung, A. Lausch, In situ/remote sensing integration to assess forest health—a review, Remote Sensing 8 (2016) 471.

[3] M. O. Barreto, S. Rey Planellas, Y. Yang, C. Phillips, K. Descovich, Emerging indicators of fish welfare in aquaculture, Reviews in Aquaculture 14 (2022) 343–361.

[4] A. Fitzgerald, C. C. Ioannou, S. Consuegra, A. Dowsey, C. García de Leaniz, Machine vision applications for welfare monitoring in aquaculture: challenges and opportunities, Aquaculture, Fish and Fisheries 5 (2025) e70036.

[5] N. Papageorgiou, P. D. Dimitriou, M. L. Moraitis, F. Massa, D. Fezzardi, I. Karakassis, Changes of the mediterranean fish farm sector towards a more sustainable approach: A closer look at temporal, spatial and technical shifts, Ocean & Coastal Management 214 (2021) 105903.

[6] F. Massa, L. Onofri, D. Fezzardi, Aquaculture in the mediterranean and the black sea: a blue growth perspective, in: Handbook on the economics and management of sustainable oceans, Edward Elgar Publishing, 2017, pp. 93–123.

[7] M. G. Leandro, J. Zornu, A. Le Breton, N. Chérif, B. Basurco, D. Furones, A. Muniesa, A. Toffan, M. Dalla Pozza, E. Franzago, et al., Quantification of biosecurity measures in mediterranean european seabass and gilthead sea bream farms, Aquaculture 596 (2025) 741898.

[8] C. L. Borgman, P. T. Darch, A. E. Sands, I. V. Pasquetto, M. S. Golshan, J. C. Wallis, S. Traweek, Knowledge infrastructures in science: data, diversity, and digital libraries, International Journal on Digital Libraries 16 (2015) 207–227.

[9] E. G. Barros, A. H. Laender, M. A. Gonçalves, F. A. Barbosa, A digital library environment for integrating, disseminating and exploring ecological data, Ecological Informatics 3 (2008) 295–308.

[10] W. Y. Arms, Digital libraries, MIT press, 2001.

[11] C. A. Lynch, Institutional repositories: essential infrastructure for scholarship in the digital age, portal: Libraries and the Academy 3 (2003) 327–336.

[12] C. L. Borgman, J. C. Wallis, M. S. Mayernik, A. Pepe, Drowning in data: digital library architecture to support scientific use of embedded sensor networks, in: Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries, 2007, pp. 269–277.

[13] C. L. Borgman, J. C. Wallis, N. Enyedy, Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries, International Journal on Digital Libraries 7 (2007) 17–30.

[14] G. Pieri, M. Cocco, O. Salvetti, A marine information system for environmental monitoring: Argo-mis, Journal of Marine Science and Engineering 6 (2018) 15.

[15] D. Dhinakaran, S. Gopalakrishnan, M. Manigandan, T. Anish, Iot-based environmental control system for fish farms with sensor integration and machine learning decision support, arXiv preprint arXiv:2311.04258 (2023).

[16] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al., Gpt-4 technical report, arXiv preprint arXiv:2303.08774 (2023).

[17] G. Comanici, E. Bieber, M. Schaekermann, I. Pasupat, N. Sachdeva, I. Dhillon, M. Blistein, O. Ram, D. Zhang, E. Rosen, et al., Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities, arXiv preprint arXiv:2507.06261 (2025).

[18] Y. Lin, T. Ruan, J. Liu, H. Wang, A survey on neural data-to-text generation, IEEE Transactions on Knowledge and Data Engineering 36 (2023) 1431–1449.

[19] M. Sharma, A. K. Gogineni, N. Ramakrishnan, Neural methods for data-to-text generation, ACM Transactions on Intelligent Systems and Technology 15 (2024) 1–46.

[20] A. Cascallar-Fuentes, J. Gallego-Fernández, A. Ramos-Soto, A. Saunders-Estevez, A. Bugarin-Diz, Automatic generation of textual descriptions in data-to-text systems using a fuzzy temporal ontology: Application in air quality index data series, Applied Soft Computing 119 (2022) 108612.

[21] M. E. Saleh, Y. M. Wazery, A. A. Ali, A systematic literature review of deep learning-based text summarization: Techniques, input representation, training strategies, mechanisms, datasets, evaluation, and challenges, Expert Systems with Applications 252 (2024) 124153.

[22] Copernicus marine data store, https://data.marine.copernicus.eu/products, 2025.

[23] Open-meteo api, https://open-meteo.com/, 2025.

[24] S. J. Taylor, B. Letham, Forecasting at scale, The American Statistician 72 (2018) 37–45.

[25] A. F. Ansari, O. Shchur, J. Küken, A. Auer, B. Han, P. Mercado, S. S. Rangapuram, H. Shen, L. Stella, X. Zhang, M. Goswami, S. Kapoor, D. C. Maddix, P. Guerron, T. Hu, J. Yin, N. Erickson, P. M. Desai, H. Wang, H. Rangwala, G. Karypis, Y. Wang, M. Bohlke-Schneider, Chronos-2: From univariate to universal forecasting, 2025. URL: https://arxiv.org/abs/2510.15821.

[26] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2015. URL: https://arxiv.org/abs/1512.03385.