International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 09 | Sept - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Data Reviver-Iterative Data Resurgence Engine

Harshan Gowda B B[1], Prof. K Sharath [2]

[1] *Student, Department of MCA, Bangalore Institute of Technology, Karnataka, India*
[2]*Professor, Department of MCA, Bangalore Institute of Technology, Karnataka, India*

Abstract

In today's world, Cyber Security and Forensics are in great demand in this data recovery plays an important role. In certain scenarios, the data storage devices could be destroyed or damaged by the convict and this is where data recovery comes in to play and recover data from these artifacts. This field of data recovery is known to many criminals and they always try to bypass the current recovery techniques whereas the organizations always tries to discover newer and more reliable recovery techniques to tackle this situation. Windows Search maintains a single database of the files, emails, programmers and Internet history of all the users of a personal computer, providing a potentially valuable source of information for a forensic investigator, especially since some information within the database is persistent, even if the underlying data are not available to the system (e.g. removable or encrypted drives). However, when files are deleted from the system their record is also deleted from the database. Existing tools to extract information from Windows Search use a programmatic interface to the underlying database, but this approach is unable to recover deleted records that may remain in unused space within the database or in other parts of the file system.

Keywords—Data Recovery, Resilience, Iterative Processing, Data Integration, Optimization.

I. INTRODUCTION

In today's data-driven world, the loss or corruption of information can lead to significant setbacks, from operational disruptions to permanent damage to digital assets. Traditional recovery methods often fall short when faced with complex failures or partial data degradation. This growing need for more intelligent and robust recovery solutions has led to the development of **Data Reviver – Iterative Data Resurgence Engine**, a powerful system designed to breathe new life into compromised datasets.

At its core, Data Reviver uses an **iterative recovery methodology**, allowing it to refine and reconstruct lost or damaged data through multiple intelligent passes. Unlike one-time recovery tools, this engine continuously analyses patterns, detects anomalies, and reprocesses information to improve data accuracy with each iteration. This makes it particularly effective in scenarios where data integrity is partially compromised, and conventional recovery would fail or yield incomplete results.

The engine is equipped with advanced algorithms, redundancy analysis, and error-correction techniques that work together to maximize data salvageability. It supports a wide range of data environments—ranging from personal file systems to enterprise databases—and can adapt to the specific nature of the data source. By using contextual awareness and historical data patterns, Data Reviver not only restores what was lost but intelligently predicts and reconstructs what is missing.

Ultimately, Data Reviver stands out as a **resilient, intelligent, and adaptive solution** for modern data challenges. Whether used for digital forensics, disaster recovery, or routine data maintenance, it empowers users to ensure data continuity and reliability in the face of unpredictable data failures. Through its iterative resurgence process, the engine doesn't just recover data—it revives it

II. LITERATURE SURVEY

he increasing dependency on digital data across industries has made data integrity and recoverability vital concerns. Traditional data recovery methods often rely on linear, one-pass operations that may not be effective in handling complex, partially corrupted, or incomplete datasets. Over the past decades, researchers have explored iterative recovery methods in fields such as image processing, signal reconstruction, and data compression. These methods repeatedly refine the recovered output to improve accuracy, and their underlying principles serve as the foundation for the Data Reviver engine.

One influential approach is the use of iterative shrinkage and thresholding techniques, particularly in image restoration. These methods aim to minimize total variation (TV) or sparsity-based norms to eliminate noise and reconstruct lost data. Algorithms like ISTA (Iterative Shrinkage-Thresholding Algorithm) and its accelerated versions have proven effective in solving inverse problems where a single solution does not exist or is difficult to compute directly. Such strategies demonstrate the power of multiple-pass refinement over static methods, reinforcing the effectiveness of iteration-based models in recovery tasks.

The concept of iterative data refinement is also prominent in compressed sensing, where incomplete data can be fully reconstructed using sparse representations. Techniques like Iteratively Reweighted Least Squares (IRLS) further enhance recovery accuracy by adjusting weights in each iteration based on the output of the previous one. These approaches have gained traction in fields like medical imaging, where data quality is critical and direct reconstruction is often unreliable. Their success highlights the potential for applying similar iterative strategies in broader data recovery systems.

Recent developments in Plug-and-Play (PnP) models have introduced a flexible framework for integrating iterative optimization with external denoising modules. In this paradigm, complex priors are replaced by learned denoisers such as deep neural networks, which can be "plugged" into the optimization process without changing the core algorithm.

This enables the system to learn from past data while maintaining a general-purpose iterative structure—an approach directly applicable to the Data Reviver engine, which may benefit from combining traditional error correction with modern machine learning techniques.

Van Cittert and Landweber iterations are classical techniques used in deblurring and reconstruction problems, where the input signal is iteratively corrected by comparing it against expected outcomes. These techniques have evolved to include dynamic updates and adaptive parameters that respond to changes in the quality of recovered data. Their strength lies in their simplicity and adaptability, offering a practical model for building scalable, real-time iterative systems. Such qualities are essential in the context of large-scale data recovery platforms.

More recent advances in deep learning have introduced architectures like DnCNN, FFDNet, and diffusion-based models, which incorporate iterative processing within neural networks themselves. These models learn to perform complex denoising or data completion tasks through layer-by-layer refinement, essentially mimicking iterative resurgence in a learned form. Incorporating similar mechanisms into the Data Reviver engine would allow it to not only recover data but also adapt its behavior based on the nature and severity of corruption.

In specialized domains such as hyperspectral imaging and satellite data restoration, iterative low-rank tensor approximations and hybrid statistical models have demonstrated impressive results. These methods account for spatial, spectral, and temporal dependencies across data dimensions and apply multi-level refinement techniques to achieve high-fidelity reconstruction. Their use of redundancy and contextual awareness could inspire similar strategies in the Data Reviver, especially for recovering structured or multidimensional datasets.

Finally, scalability and efficiency remain key concerns for any iterative system. Research in parallel and distributed processing—such as mesh or pyramid-based architectures—has shown that large-scale iterative recovery can be computationally feasible with the right hardware support. These advancements ensure that iterative engines like Data Reviver can function effectively in real-time environments or under heavy data loads, making them suitable for both enterprise and personal data recovery scenarios.

III. EXISTING SYSTEM

The current landscape of digital recovery and forensic tools reveals a significant imbalance between data recovery applications and full-fledged forensic suites. On one side, commonly available recovery tools such as Recuva, TestDisk, and EaseUS are designed to retrieve deleted or corrupted files from storage devices like hard drives, USB drives, or memory cards. These tools are often simple to use and effective for restoring files in personal or small-scale scenarios, but their functionality stops at the point of recovery. They do not provide the ability to manage cases, maintain evidence logs, or ensure the preservation of the chain-of-custody, all of which are critical in digital forensic investigations. In forensic contexts, simply recovering a file is not enough—the investigator must also demonstrate how the evidence was acquired, who accessed it, and that it has not been altered. This lack of forensic accountability in existing recovery tools severely limits their application in professional or legal investigations.

Another critical limitation of existing systems is the lack of integration between recovery, documentation, and reporting. In many current investigative workflows, examiners must use multiple disconnected tools: one for recovering files, another for creating case notes, and yet another for compiling formal reports. This fragmentation not only increases the time required to complete an investigation but also introduces the risk of inconsistencies and errors in documentation. Since evidence integrity is paramount, any discrepancies between recovered data and documented reports can compromise the validity of findings. Manual logging, which is often relied upon in such workflows, is especially vulnerable to omissions, inaccuracies, and human error, weakening the credibility of the investigation.

In summary, the existing systems in the field either fall short by being overly simplistic and recovery-focused without forensic features, or they overshoot by being highly complex, expensive, and resource-heavy, limiting their practical use for many investigators and students. The reliance on manual documentation and fragmented tools makes current workflows inefficient and error-prone, creating an urgent need for a balanced, affordable, and integrated solution. This gap in the existing system forms the foundation for the development of *Data Reviver*, a tool specifically designed to unify data recovery with forensic case management, evidence tracking, and automated reporting in a way that is reliable, secure, and user-friendly.

Disadvantages

- Iterative processes often require multiple passes over the data, leading to increased computational load and longer processing times, especially for large-scale datasets.
- When initial data is highly corrupted, the iterative mechanism may reinforce inaccuracies or noise, resulting in compounded errors and unreliable recovery outcomes.
- The effectiveness of the engine can depend heavily on careful tuning of parameters like iteration count, thresholds, and convergence criteria, which may require technical expertise.
- Scaling the system to handle massive or real-time data streams can be challenging, as iterative models may struggle with performance bottlenecks without optimized infrastructure.

IV. PROPOSED SYSTEM

The proposed system, Data Reviver – Forensic Recovery Tool, is designed to overcome the shortcomings of existing recovery applications and forensic suites by providing a unified, efficient, and user-friendly solution that integrates data recovery with forensic case management. Unlike traditional recovery tools that stop at restoring deleted files, the proposed system introduces a structured framework where recovered files are directly associated with a forensic case, ensuring that every action is logged and evidence integrity is preserved. By combining file recovery with automated evidence logging, manifest creation, and professional report generation, the system streamlines the entire investigative workflow into a single platform, eliminating the need for multiple disconnected tools and reducing the chances of errors caused by manual documentation.

The proposed system also emphasizes usability and accessibility, offering an intuitive WinForms-based graphical user interface that simplifies the tasks of case creation, evidence management, and report generation. Investigators

can easily view available drives, select files for recovery, and add them to a forensic case through an organized panel structure. Unlike complex forensic suites that require extensive training, Data Reviver is designed to be used by both professional forensic examiners and IT staff with limited forensic experience, making it more versatile and accessible across different user groups.

Furthermore, the proposed system includes automated report generation as a core functionality. At the end of an investigation, examiners can export professional reports containing details of the case, evidence, and logged actions. These reports are structured, consistent, and suitable for submission in legal proceedings, academic research, or corporate reviews. By automating this process, the system reduces the workload on investigators and ensures that all reports adhere to a standardized format, eliminating inconsistencies caused by manual preparation.

The modular design of the system allows for scalability and future enhancements, including support for additional file systems, integration with external forensic databases, and cloud synchronization for collaborative investigations. This adaptability ensures that the tool is not limited to its initial scope but can evolve to meet the needs of emerging forensic challenges.

In essence, the proposed system seeks to deliver a comprehensive digital forensic solution that balances functionality, usability, and security. By unifying recovery, evidence logging, case management, and reporting within a single application, Data Reviver eliminates inefficiencies, reduces errors, and ensures that digital investigations are conducted in a reliable, accountable, and legally admissible manner. Its affordability, simplicity, and extensibility make it suitable for a wide range of users from law enforcement agencies and corporate IT departments to academic researchers and students thereby addressing the critical gaps left by existing systems.

Advantages:

- Iterative recovery methods can gradually improve data quality by refining results over multiple passes, leading to higher accuracy compared to one-time recovery techniques.
- They are highly adaptable and can be tailored to handle different types of data corruption, including missing values, noise, and structural damage.
- These systems often integrate well with advanced algorithms like machine learning or denoising models, enhancing their ability to intelligently reconstruct complex datasets.
- Iterative engines allow for partial recovery and continuous improvement, making them suitable for dynamic environments where data may be recovered in stages.

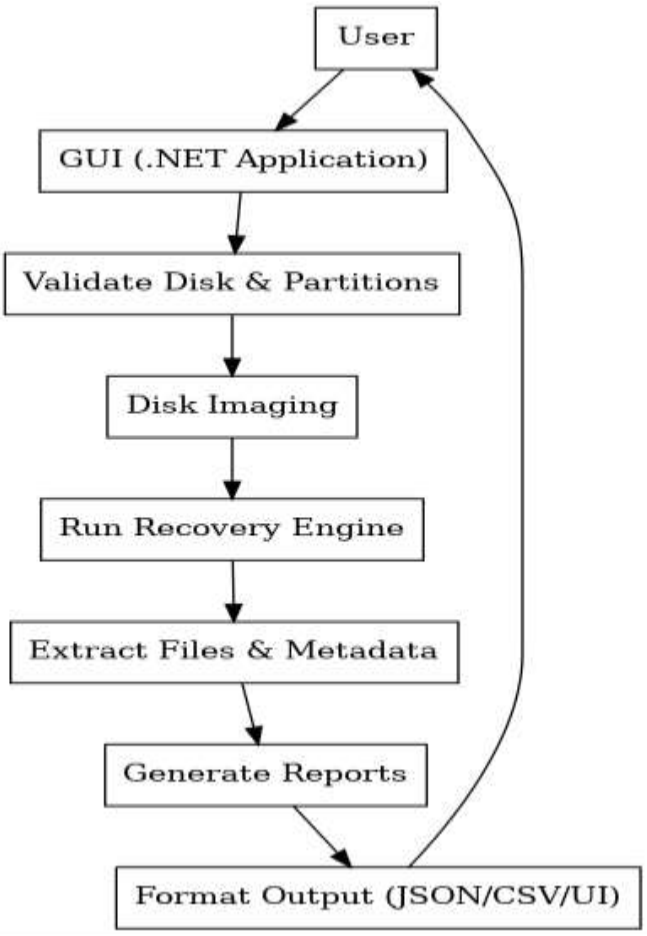

Fig 1: Proposed Model

V. IMPLEMENTATIONS

System Architecture:

The Data Reviver – Iterative Data Resurgence Engine is architected as a modular, scalable. The system is designed to separate key components including user authentication, data ingestion, iterative recovery processing, multi-modal integrity analysis, and final output generation. It supports flexible deployment across cloud or on-premise environments and is optimized for performance and fault tolerance in handling degraded or partially lost data.

Authentication and User Management:

The system employs **OAuth2 with JWT tokens** to ensure secure user authentication and authorization. Users can register, log in, and manage API tokens, with passwords securely hashed using **passlib** and **bcrypt** to protect sensitive credentials. For input handling, Data Reviver supports uploads of various data types (e.g., CSV, JSON, XML, audio, or video files) through a user-friendly interface, with strict validation enforced via **Pydantic models** to verify file formats and size limits, ensuring only supported and safe data is processed by the engine.

Multi-Modal Analysis Pipeline:

The system processes input data through multiple parallel stages tailored to the data type and corruption pattern. For example, structured data undergoes schema validation and statistical imputation, textual data is analyzed using transformer-based contextual reconstruction models, and audio

or media files are examined with signal processing and anomaly detection algorithms. Each processing path generates intermediate recovery outputs with associated confidence metrics and error estimates, enabling fine-grained tracking of data integrity across different dimensions

Result Fusion and Generation:

Outputs from the various recovery modules are intelligently merged using iterative alignment and conflict resolution techniques. The system consolidates overlapping or contradictory data segments by applying weighted confidence scoring and temporal or structural consistency checks. The final result is a unified, comprehensive dataset restoration report presented in a structured JSON format, which includes reconstructed data, quality indicators, and detailed logs of the recovery process, allowing users to evaluate the completeness and reliability of the restored information.

Error Handling and Security:

Robust error handling and security are integral to the Data Reviver engine. The system enforces strict input validation to ensure only supported data formats and sizes are processed, minimizing the risk of corruption or system failure. Comprehensive exception handling captures and logs errors at every stage of the recovery pipeline, enabling prompt troubleshooting and system resilience. Security measures include CORS configuration for safe cross-origin requests, environment-based management of sensitive secrets and API keys, and encryption of data both in transit and at rest. These practices ensure reliable operation and secure integration within enterprise environments, safeguarding user data and maintaining system integrity

VI. CONCLUSIONS

One potential future enhancement is the integration of advanced machine learning models, particularly deep learning architectures, to further improve the accuracy and adaptability of the recovery process. By training on diverse datasets and corruption scenarios, these models can learn complex patterns of data degradation and more effectively predict missing or corrupted segments. Incorporating techniques such as reinforcement learning or self-supervised learning could also enable the engine to autonomously refine its strategies over time, reducing the need for manual parameter tuning and enhancing recovery performance across varying data types.

Another avenue for enhancement involves expanding the multi-modal analysis pipeline to support a broader range of data formats and sources, including real-time streaming data, sensor outputs, and complex relational databases. This would require developing specialized modules capable of handling the unique characteristics of these formats, such as temporal dependencies or hierarchical relationships. Additionally, introducing distributed and parallel processing capabilities would allow the system to scale efficiently, enabling faster recovery of large datasets and supporting enterprise-level deployments with high throughput requirements.

VII. REFERENCES

[1] S. Tomer, A. Apurva, P. Ranakoti, S. Yadav, and N. R. Roy, "Data recovery in forensics," (IC3TSN), 2017 International Conference on Computing and Communication Technologies for Smart Nation Gurgaon, 10.1109/IC3TSN.2017.8486093. India, pp. 188–192, 2017. doi:

[2] H. Chivers and C. Hargreaves, "Forensic data recovery from the Windows Search database," Digital Investigation, vol. 7, pp. 114–126, 2011. doi: 10.1016/j.diin.2011.01.001.

[3] R. R. Ali, K. M. Mohamad, S. Jamel, and S. K. A. Khalid, "A review of digital forensics methods for JPEG file carving," Journal of Theoretical and Applied Information Technology, vol. 96, no. 17, pp. 5841–5855, Sep. 2018.

[4] B. C. Ogazi-Onyemaechi, A. Dehghantanha, and K. K. R. Choo, "Performance of Android Forensics Data Recovery Tools," Contemporary Digital Forensic Investigations of Cloud and Mobile Applications, Chapter 7, pp. 91–110, Elsevier, 2017.

[5] H. I. Sahib, N. H. Ab Rahman, A. K. Al-Qaysi, and M. L. Attiah, "Comparison of data recovery techniques on master file table between Aho-Corasick and logical data recovery based on efficiency," TELKOMNIKA Telecommunication, Computing, Electronics and Control, vol. 19, no. 1, pp. 73–78, Feb. 2021. doi: 10.12928/TELKOMNIKA.v19i1.16276.

[6] S. N. Varayogula, K. Dodiya, P. Lakhalani, and A. Chawla, "Computer Forensics Data Recovery Software: A Comparative Study," International Journal of Innovative Research in Computer Science & Technology (IJIRCST), vol. 10, no. 2, pp. 513-518, Mar. 2022, doi: 10.55524/ijircst.2022.10.2.101.

[7] V. Vijayan, Android Forensic Capability and Evaluation of Extraction Tools, M.Sc. Thesis, Edinburgh Napier University, Scotland, Apr. 2012. [7] M. R. Lehrfeld, "Insider Risk: Finding Sensitive Files in the Enterprise using a PC's Master File Table," 2018 ASCUE Proceedings, East Tennessee State University, pp. 70–77, 2018.

[8] J. A. M. Jeyaseeli and C. Shanthi, "A smart techniques to extract the deleted data form the android application," International Journal of Health Sciences, vol. 6, Special Issue, pp. 2864–2871, 2022. doi: 10.53730/ijhs.v6nS1.5284.

[9] M. F. Abdillah and Y. Prayudi, "Data Recovery Comparative Analysis Using Open-based Forensic Tools Source on Linux," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 13, no. 9, pp. 633–639, 2022.

[10] M. Breeuwsma, M. de Jongh, C. Klaver, R. van der Knijff, and M. Roeloffs, "Forensic Data Recovery from Flash Memory," Small Scale Digital Device Forensics Journal, vol. 1, no. 1, pp. 1–17, June 2007.