Multi-Modal Data Fusion for Enhanced Analytics: Techniques for Integrating Structured and Unstructured Big Data

Serena Chen
Department of Data Science and Engineering, University of British Columbia, Canada
serena.chen@ubc.ca

Md. Murad Hossain
Modeling and Data Science Program, University of Turin, Canada
sadid.shimul@gmail.com

Abstract

With the exponential growth of data from diverse sources, organizations are increasingly looking to integrate and analyze multi-modal data to gain deeper insights. Structured data from databases and sensors can provide quantitative insights, while unstructured data from text, images, video and audio can provide contextual, qualitative understanding. Multi-modal data fusion enables a more comprehensive view by combining the breadth of unstructured data with the depth of structured data. This paper provides an overview of multi-modal data fusion techniques to enhance analytics. It covers methods like entity matching, linking and resolution that integrate structured and unstructured data at the entity-level. Techniques like feature extraction and sensor fusion that consolidate data at the feature-level are also discussed. The relative strengths and limitations of different techniques are considered in the context of analytics objectives. Challenges such as semantic alignment, data veracity, and cognitive load are examined. The paper concludes with best practices and future directions for multi-modal data fusion.

Keywords: Big Data, Multi-Modal Data, Fusion Techniques, Fusion, unstructured data

Introduction

The digital revolution has enabled an unprecedented exponential growth in data volume, variety, and velocity. As per IDC estimates, the global datasphere is expected to grow from 33 zettabytes in 2018 to 175 zettabytes by 2025. Data is being generated from diverse sources in different modalities and formats. Sensor data from Internet of Things (IoT) devices provides detailed quantitative measurements. Text documents, emails, social media posts, audio recordings and videos capture unstructured qualitative information. Images and multimedia encode perceptual knowledge [1]. The result is a complex amalgamation of multi-modal big data encompassing both structured and unstructured information. Structured data refers to quantitative information with a well-defined schema that is readily analyzable using standard queries and algorithms. It is typically generated from organizational databases, transactional systems, ERP applications, BI tools, spreadsheets, sensory devices and other systems of record. Structured data fields have precise semantics and relationships enabling systematic access. For example, relational databases contain interlinked tables with columns corresponding to different data attributes [2]. Time-series sensors record measurements at specific time intervals. GPS data contains latitude and longitude coordinate pairs. This allows structured data to be efficiently stored, indexed, processed and mined. The quantitative metrics also facilitate data-driven analytics, predictions, optimization, and automation. However, structured data lacks the contextual details and qualitative insights needed for a nuanced understanding of events, behaviors, and entities [3].

In contrast, unstructured data refers to information without a predefined schema or organization. It comprises text documents, emails, social media posts, audio recordings, videos, images, and other multimedia content [4], [5]. While unstructured data accounts for over 80% of organizational information, it poses multiple analytics challenges [6]. Ambiguities in natural language and the diversity of data formats make it difficult to systematically extract signals. But unstructured data provides vital contextual knowledge and qualitative insights that cannot be captured by structured data alone. The sentiments, opinions, cultural trends, relationships, and conceptual details within unstructured data enable a more holistic understanding of entities and events. With the concurrent explosion of both structured and unstructured big data, organizations are increasingly looking to integrate these heterogeneous data types to drive competitive advantage [7]. Neither structured nor unstructured data alone provides sufficient depth and breadth of understanding. Structured data enables quantitative measurement, monitoring, and optimization but lacks contextual details. Unstructured data provides qualitative intelligence but cannot easily be analyzed at a scale. The fusion of multi-modal data sources harnesses the best of both worlds - the depth of structured metrics and the breadth of contextual, unstructured signals [8], [9].

Multi-modal data fusion powers a wider range of analytics applications by enabling a unified view. Integrating structured and unstructured data facilitates both precise measurement as well as nuanced contextual understanding [10]. Some key analytics use cases powered by multi-modal data fusion include:

- Augmenting business metrics and operational data with customer feedback and market intelligence for strategic planning and product improvement .
- Enhancing structured risk models using earnings call transcripts, regulatory filings, and news events analysis.
- Improving predictive maintenance by correlating structured sensor data with unstructured technician maintenance logs and expert knowledge.
- Generating personalized recommendations by connecting user transactions and structured profile attributes with unstructured demographics, social media activity, and multimedia preferences.
- Detecting financial fraud by combining structured anomalies and transaction patterns with audio, text and network forensics.
- Deriving market insights by extracting structured semantic metadata from unstructured text, audio, video and image content.

Across domains, multi-modal data fusion enables both a panoramic perspective through the breadth of unstructured data as well as precise measurements and details through structured data. It provides the contextual augmentation needed for advanced analytics. This paper examines the techniques and technologies that enable meaningful fusion of structured and unstructured big data for actionable intelligence and enhanced decision making.

Background

Structured Data: Structured data refers to quantitative information with a predefined schema that enables easy storage, search and analysis. Common structured data sources include organizational databases, spreadsheets, financial/operational systems and sensory devices. Structured data fields have explicit meanings and relationships that allow systematic access and analysis [11]. Examples include customer contact records, sales transaction logs, sensor measurements of speed or temperature, and GPS coordinates. Structured data can be readily processed by algorithms and tools to generate reports, dashboards, alerts and predictive models. The precise metrics enable data-driven decision making and automation. However, structured data lacks contextual details to enable a nuanced understanding of events and behaviors [12].

Unstructured Data: In contrast, unstructured data refers to qualitative information without a predefined structure or schema. It includes text documents, social media posts, audio recordings, images, videos and other multimedia content. Unstructured data comprises over 80% of organizational information but is difficult to systematically analyze due to ambiguities in natural language and differences in formats. However, unstructured data provides vital contextual insights into customer sentiments, cultural trends, public perceptions, and other qualitative factors that structured data cannot capture [13]–[15]. Analyzing unstructured data often requires machine learning techniques like natural language processing, image recognition and speech analysis to extract meaningful signals. Unstructured data enables a more holistic and nuanced view of the world, complementing structured metrics.

Analytics Applications: Integrating structured and unstructured data enables a wider range of analytics applications to drive competitive advantage. Multi-modal data fusion augments

performance measurement and process optimization with deeper customer and market intelligence. Some key analytics use cases enabled by multi-modal data include:
- Customer sentiment analysis - Combine structured customer ratings with unstructured feedback, reviews and social media posts to understand satisfaction levels and pain points.
- Risk modelling - Enhance structured financial risk models with analysis of qualitative earnings call data, news events, and regulatory filings.
- Predictive maintenance - Correlate structured sensor data with unstructured maintenance logs and expert knowledge to predict equipment failures.
- Recommendation systems - Build holistic user profiles by connecting structured transactions with unstructured demographics, social media activity and text/multimedia preferences.
- Fraud detection - Identify structured anomalies as well as contextual red flags from text, audio and network patterns.
- Content analytics - Extract structured semantic metadata from unstructured text, images, audio and video to generate insights.

Multi-modal data fusion powers these applications by connecting the breadth of unstructured data with the depth of structured data. The subsequent sections examine fusion techniques for different data modalities and analytic objectives.

Multi-Modal Data Fusion Techniques

Multi-modal data fusion integrates discrete data types at different levels of abstraction to create integrated meaning. Key techniques can be categorized as entity-level, feature-level or decision-level fusion based on the consolidation target:

Entity-Level Fusion

Entity-level fusion associates different data objects that correspond to the same real-world entity. It enables a unified view by connecting heterogeneous data records referring to the same person, location, event or concept. Techniques include:
- Entity matching - Links records from multiple datasets referring to the same entity based on similarity of identifying attributes. This allows merging of disparate data sources.
- Entity resolution - Deduplicates entity references from different sources by data cleaning and integration.
- Entity linking - Disambiguates entity mentions unstructured data by linking to canonical definitions in a structured knowledge base. For example, linking “Washington” in text to the proper person vs. place entity.

Entity-level fusion provides a critical consolidation of heterogeneous data at the foundation. It enables augmentation of structured records with contextual details and linkage across datasets for a 360-degree view. However, entity consolidation is also challenging due to data veracity issues and computational complexity for large datasets.

Table 1 compares the different techniques for entity-level fusion:

Technique	Description	Strengths	Limitations
Entity matching	Linking records from multiple datasets referring to the same real-world entity based on similarity of identifying attributes	- Allows merging of disparate data sources into unified view.- Scalable to large datasets with blocking, filtering and approximate matching techniques	- Computationally expensive.- Accuracy challenged by poor data quality, missing values.- Requires tuning similarity thresholds
Entity resolution	Deduplicating multiple entity references from different sources through data cleaning and integration	- Consolidates references to ensure a single version of truth for each entity.- Improves data quality by handling inconsistencies	- Quadartic complexity without optimization.- Ambiguity in disambiguating close matches
Entity linking	Disambiguating entity mentions in unstructured data by	- Annotates unstructured data with formal definitions of entities.-	- Knowledge base coverage limits context.- Ambiguity remains in

	linking to canonical definitions in a structured knowledge base	Facilitates integration with structured data	absence of sufficient context

Feature-Level Fusion

Feature-level fusion consolidates the data at the level of derived variables or features. Features extracted from diverse data types are combined into unified feature vectors for integrated mining and modelling. For example:
- Text mining - Extracting sentiment variables from unstructured text data.
- Audio/video analysis - Generating facial expressions or acoustic features from multimedia.
- Sensor fusion - Combining readings from multiple sensors into consolidated variables.

Feature-level fusion enables synthesis of higher-order variables that marshal the strengths of diverse data types. However, it can result in very high dimensionality and loss of granularity. Feature selection and dimensionality reduction are critical to avoid poor model performance or overfitting. The derived features should also be carefully engineered to capture all necessary information from the source variables.

Table 2 illustrates different feature extraction and fusion techniques:

Technique	Description	Examples
Text mining	Extracting semantic features from unstructured text data	- Sentiment analysis - Extracting sentiment polarity scores.- Topic modelling - Discovering abstract topics using LDA.- Embedding - Representing text in dense vector space
Audio analysis	Extracting acoustic features from audio data	- Speech recognition - Extracting transcripts.- Speaker diarization - Identifying speakers.- Prosody analysis - Extracting rhythm, intonation features
Video analysis	Extracting visual features from image/video data	- Object recognition - Detecting objects using CNNs.- Activity recognition - Classifying human activities in video.- Scene classification - Categorizing outdoor/indoor scenes
Sensor fusion	Combining sensor readings into higher-level features	- Merging lidar, camera, radar data for environmental perception in autonomous vehicles.- Fusing inertial sensor data for accurate motion tracking

Decision-Level Fusion

Decision-level fusion consolidates the outputs of independent models built on individual data types. For instance, predictive models trained separately on structured data variables and text features can be combined through:
- Voting - Models vote on the overall predicted class.
- Averaging - Continuous outputs like probabilities are averaged across models.
- Meta-Classifier - A higher-level model is trained on the base model outputs.

Decision-level fusion provides a late-stage consolidation of different inference results. However, errors or biases from the base models can get propagated to the integrated prediction. The base models may also end up redundant if built on similar features. Decision-level fusion works best when the data sources and models provide complementary signals.

Table 3 contrasts different decision-level fusion strategies:

Fusion Strategy	Description	Advantages	Disadvantages
Voting	Base models vote on overall predicted class	- Simple to implement. - Provides robustness against individual model failures	- All models must have the same classes. - Weights not tunable for model importance

Averaging	Continuous outputs like probabilities averaged	- Intuitive ensemble approach. - Handles different output scales	- Losses complementarity of outputs. - Sensitive to outliers
Meta-classifier	Higher-level model trained on base model outputs	- Learns complex interactions between models. - Weights tuned based on validation	- Prone to overfitting. - Interpretability challenging

The optimal fusion level depends on the problem context. Entity-level fusion is indispensable for a unified view and prevents propagation of errors. Feature-level fusion enables synergies through joint modelling but requires careful engineering. Decision-level fusion is best for late stage synthesis but can compound biases. A combinatorial approach across levels is often most robust for analytics.

Challenges in Multi-Modal Data Fusion

While offering enriched insights, integrating diverse data types also poses multiple challenges including:

Semantic Alignment With heterogeneous, multi-modal data, a key difficulty is aligning the semantics to enable meaningful integration [16]. Structured data fields often lack contextual details captured in unstructured data. Low level features extracted from text or multimedia may not correspond to high-level abstract concepts. Automated semantic matching and mapping techniques like neural embeddings are still imperfect and often need human expertise .

Data Veracity: Data obtained by fusing multiple sources can suffer from inaccuracies inherited from the original sources. Errors may get introduced during incomplete entity linking or inaccurate feature extraction. This poses data quality challenges for downstream analytics. Provenance tracking, uncertainty modelling and outlier detection are necessary to ascertain veracity.

Cognitive Load: Joint analysis of mixed data types increases the cognitive load for data scientists. Understanding low level multimedia features along with abstract structured data requires broad expertise. Diverse data formats also increase the software and infrastructure complexity. Specialized multi-modal analytics tools are needed to ease cognitive burdens [2], [17].

The fusion process entails iterative refinements using validation data to ensure accurate mapping, alignment and integrated metrics generation. Governance frameworks managing data lifecycles, models and decisions are also critical to address ethical concerns with fusion techniques.

Best Practices for Multi-Modal Data Fusion

Based on the opportunities and challenges, some recommended best practices for fusing structured and unstructured data include:

- Perform fusion iteratively with continuous validation of mapping and alignments. Involve both data engineers and domain experts.
- Build a knowledge graph to represent the consolidated entity-level schema. Use unique persistent identifiers for entities.
- Leverage transfer learning approaches to align vector space representations across modalities.
- Employ explainable AI techniques to generate human-interpretable features. Assess cognitive load.
- Implement rigorous data provenance tracking and uncertainty modelling for veracity.
- Develop specialized analytics tools that facilitate multi-modal feature engineering and modelling.
- Follow modelops and responsible AI practices to address bias, fairness and accountability.

The field continues to evolve with advances in representation learning, causality modelling and human-AI interaction. Next we discuss future directions.

Future Outlook

Key developments on the horizon for multi-modal data fusion include:

- Improved semantic contextualization using large language models and graph embeddings.
- Reinforcement learning and adversarial techniques for automated feature alignment.

- Multimodal generative models like conditional GANs to synthesize fused representations.
- Causal modelling to infer causative relationships between modalities.
- Intelligent interfaces and visual analytics to streamline multi-modal analysis.
- Edge and neuromorphic computing to enable real-time sensor fusion.
- Blockchain and distributed ledger solutions to manage veracity in fusion.

As technologies mature, multi-modal data fusion will become more automated, real-time and scalable. Next-generation platforms will allow non-technical users to achieve unified insights from diverse big data. The future remains promising for augmented intelligence through synthesis of broad and deep data.

Conclusion

The fusion of multi-modal data sources provides the essential breadth and depth needed for comprehensive analytics and intelligence generation. Structured data enables quantification, measurement, monitoring and precise analysis [18], [19]. Unstructured data provides the qualitative context, sentiments, associations and details absent from structured data. By consolidating their complementary strengths, multi-modal data fusion facilitates both a panoramic perspective as well as focused insights. It overcomes the inherent limitations of analyzing structured or unstructured data in isolation. As elucidated through this paper, a variety of techniques at the entity, feature and decision levels exist for integrating heterogeneous data types [20]. Entity-level fusion through matching, resolution and linking establishes consolidated views of real-world entities by connecting references across datasets. This interlinking of knowledge is the foundation required for unified analytics [21]. Feature-level fusion enables synthesis of higher-order derived variables that marshal the signals within different modalities. Decision-level fusion supports late stage consolidation of inferred knowledge. A combinatorial approach across the levels provides a robust framework for analytics over structured and unstructured data [22].

While promising, multi-modal data fusion poses multiple challenges as well. Semantic alignment across modalities remains difficult owing to differences in representations. Techniques like neural embeddings have advanced representations learning but still require contextual fine-tuning [23]. Data veracity is challenged as inaccuracies can get introduced during entity linking or feature extraction. Rigorous validation, uncertainty modelling and provenance tracking is imperative to ensure high quality fusion. The increased data diversity also heightens the cognitive load for analysts, necessitating advances in human-AI interaction. Automated fusion processes need to apply robust data validation, explain ability and responsible AI practices.

Ongoing advances in Representation learning using large language, graph and multimodal models facilitate better semantic contextualization and alignment. Transfer learning and adversarial techniques can allow automated adaptation of features across modalities. Multimodal generative models like conditional GANs can synthesize fused representations. Causal modelling approaches infer causative mechanisms between modalities to select optimal integration techniques [24]. Human-centered techniques like interactive visual analytics and natural language interfaces ease cognitive burdens. The emergence of edge and quantum computing will enable real-time, low-latency analytics over fused data. Blockchain and distributed ledger architectures provide immutable provenance for enhanced veracity. As these technologies mature, multi-modal data fusion will become more flexible, automated, and scalable. It has the potential to become the digital analytics paradigm for the foreseeable future, enabling organizations to achieve augmented insights and intelligence [25]. With careful application of emerging techniques, end users without extensive technical expertise can be empowered to analyze disparate data meaningfully. This democratization promises to unlock tremendous latent value from siloed structured and unstructured data [26].

In essence, multi-modal data fusion is set to fundamentally transform analytics, intelligence generation and data-driven decision making. It overcomes the constraints of single data types to provide a holistic digital representation [27]. The synergistic combination of depth, breadth, structure and context sets the stage for the next level of actionable knowledge discovery. Realizing this vision requires diligent efforts to develop robust technologies while addressing ethics, privacy and responsible AI concerns. With sound implementation, multi-modal data fusion can usher in the next frontier of analytics, delivering profound benefits for individuals, organizations and society. The future remains promising but prudence paramount.

References

[1] T. Adali, Y. Levin-Schwartz, and V. D. Calhoun, "Multi-modal data fusion using source separation: Two effective models based on ICA and IVA and their properties," *Proc. IEEE Inst. Electr. Electron. Eng.*, vol. 103, no. 9, pp. 1478–1493, Sep. 2015.

[2] S. Bruge *et al.*, "Multi-modal data fusion for Cardiac Resynchronization Therapy planning and assistance," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2015, pp. 2391–2394, 2015.

[3] J. Simanek, V. Kubelka, and M. Reinstein, "Improving multi-modal data fusion by anomaly detection," *Auton. Robots*, vol. 39, no. 2, pp. 139–154, Aug. 2015.

[4] M. Elshehaly, D. Gračanin, M. Gad, H. G. Elmongui, and K. Matković, "Interactive fusion and tracking for multi-modal spatial data visualization," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 251–260, Jun. 2015.

[5] A. Nassar and M. Kamal, "Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations," *IJRAI*, vol. 11, no. 8, pp. 1–11, Aug. 2021.

[6] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Context-aware query performance optimization for big data analytics in healthcare," in *2019 IEEE High Performance Extreme Computing Conference (HPEC-2019)*, 2019, pp. 1–7.

[7] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Federated query processing for big data in data science," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 6145–6147.

[8] M.-H. Amri, Y. Becis, D. Aubry, and N. Ramdani, "Indoor human/robot localization using robust multi-modal data fusion," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, WA, USA, 2015.

[9] M. Kamal and T. A. Bablu, "Machine Learning Models for Predicting Click-through Rates on social media: Factors and Performance Analysis," *IJAMCA*, vol. 12, no. 4, pp. 1–14, Apr. 2022.

[10] A. Eshaghi *et al.*, "Classification algorithms with multi-modal data fusion could accurately distinguish neuromyelitis optica from multiple sclerosis," *NeuroImage Clin.*, vol. 7, pp. 306–314, Jan. 2015.

[11] P. Chavali and A. Nehorai, "Hierarchical particle filtering for multi-modal data fusion with application to multiple-target tracking," *Signal Processing*, vol. 97, pp. 207–220, Apr. 2014.

[12] P. Budikova, M. Batko, D. Novak, and P. Zezula, "Inherent fusion," *J. Database Manag.*, vol. 27, no. 4, pp. 1–23, Oct. 2016.

[13] R. Arianto, H. L. H. S. Warnars, F. L. Gaol, and A. Trisetyarso, "Mining unstructured data in social media for natural disaster management in Indonesia," in *2018 Indonesian Association for Pattern Recognition International Conference (INAPR)*, Jakarta, Indonesia, 2018.

[14] H. Yue, "Unstructured healthcare data archiving and retrieval using Hadoop and Drill," *Int. J. Big Data Anal. Healthc.*, vol. 3, no. 2, pp. 28–44, Jul. 2018.

[15] K. Sambrekar, V. S. Rajpurohit, and J. Joshi, "A proposed technique for conversion of unstructured Agro-data to semi-structured or structured data," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, 2018.

[16] G. Li and P. Bours, "A mobile app authentication approach by fusing the scores from multi-modal data," in *2018 21st International Conference on Information Fusion (FUSION)*, Cambridge, United Kingdom, 2018.

[17] L. M. García-Moreno, J. P. Díaz-Paz, H. Loaiza-Correa, and A. D. Restrepo-Girón, "Dataset of thermal and visible aerial images for multi-modal and multi-spectral image registration and fusion," *Data Brief*, vol. 29, no. 105326, p. 105326, Apr. 2020.

[18] C. Zhu, Q. Li, L. Kong, X. Wang, and X. Hong, "Associated index for big structured and unstructured data," in *Web-Age Information Management*, Cham: Springer International Publishing, 2015, pp. 567–570.

[19] F. Liu and N. Wu, "SentiLDA — an effective and scalable approach to mine opinions of consumer reviews by utilizing both structured and unstructured data," in *Big Data Analytics and Knowledge Discovery*, Cham: Springer International Publishing, 2016, pp. 82–96.
[20] X. Liu, P. V. Singh, and K. Srinivasan, "A structured analysis of unstructured big data by leveraging cloud computing," *Mark. Sci.*, vol. 35, no. 3, pp. 363–388, May 2016.
[21] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Big data in cloud computing review and opportunities," *arXiv preprint arXiv:1912.10821*, 2019.
[22] A. Majumdar and J. Sowa, "- extracting useful information from multivariate temporal data," in *Big Data and Business Analytics*, Auerbach Publications, 2016, pp. 194–213.
[23] S. Venkatraman, School of Engineering, Construction and Design (IT), Melbourne Polytechnic, VIC 3072, Australia, K. F. S. Kaspi, and R. Venkatraman, "SQL Versus NoSQL Movement with Big Data Analytics," *Int. J. Inf. Technol. Comput. Sci.*, vol. 8, no. 12, pp. 59–66, Dec. 2016.
[24] M. van Rijmenam, T. Erekhinskaya, J. Schweitzer, and M.-A. Williams, "Avoid being the Turkey: How big data analytics changes the game of strategy in times of ambiguity and uncertainty," *Long Range Plann.*, vol. 52, no. 5, p. 101841, Oct. 2019.
[25] M. Fahmideh and G. Beydoun, "Big data analytics architecture design—An application in manufacturing systems," *Comput. Ind. Eng.*, vol. 128, pp. 948–963, Feb. 2019.
[26] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Approximate query processing for big data in heterogeneous databases," in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 5765–5767.
[27] M. Minelli, M. Chambers, and A. Dhiraj, *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*. John Wiley & Sons, 2013.