Towards Multimodal Knowledge Graphs for Data Spaces

Atiya Usmani*
Insight SFI Research Centre for Data Analytics, Data Science Institute, University of Galway
Galway, Ireland
a.usmani1@universityofgalway.ie

M. Jaleed Khan*
SFI Centre for Research Training in AI, Data Science Institute, University of Galway.
Galway, Ireland
m.khan12@universityofgalway.ie

John G. Breslin
Insight SFI Research Centre for Data Analytics, Data Science Institute, University of Galway.
Galway, Ireland
john.breslin@universityofgalway.ie

Edward Curry
Insight SFI Research Centre for Data Analytics, Data Science Institute, University of Galway.
Galway, Ireland
edward.curry@universityofgalway.ie

ABSTRACT

Multimodal knowledge graphs have the potential to enhance data spaces by providing a unified and semantically grounded structured representation of multimodal data produced by multiple sources. With the ability to integrate and analyze data in real-time, multimodal knowledge graphs offer a wealth of insights for smart city applications, such as monitoring traffic flow, air quality, public safety, and identifying potential hazards. Knowledge enrichment can enable a more comprehensive representation of multimodal data and intuitive decision-making with improved expressiveness and generalizability. However, challenges remain in effectively modelling the complex relationships between and within different types of modalities in data spaces and infusing common sense knowledge from external sources. This paper reviews the related literature and identifies major challenges and key requirements for effectively developing multimodal knowledge graphs for data spaces, and proposes an ontology for their construction.

CCS CONCEPTS

• **Computing methodologies → Visual content-based indexing and retrieval**; **Ontology engineering**; • **Information systems → Information integration**.

KEYWORDS

data spaces, knowledge graphs, multimodal data, smart city

ACM Reference Format:
Atiya Usmani, M. Jaleed Khan, John G. Breslin, and Edward Curry. 2023. Towards Multimodal Knowledge Graphs for Data Spaces. In *Companion Proceedings of the ACM Web Conference 2023 (WWW '23 Companion), April 30–May 04, 2023, Austin, TX, USA*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3543873.3587665

*Both authors contributed equally to this research.

WWW '23 Companion, April 30–May 04, 2023, Austin, TX, USA

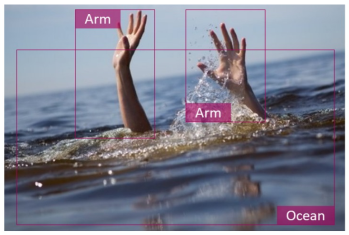
ACM ISBN 978-1-4503-9419-2/23/04.
https://doi.org/10.1145/3543873.3587665

1 INTRODUCTION

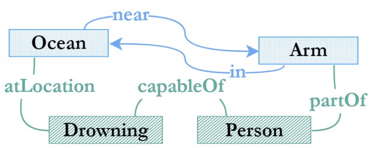
Data spaces are a growing paradigm in decentralized data management that facilitate collaboration and sharing of multimodal data for enhanced data transparency and open innovation. They offer an environment for businesses and organizations to store, manage, and analyze different types of data, gaining new insights and making better-informed decisions. [11] However, modeling the complex relationships between different types of data effectively can be challenging, along with addressing issues such as data integration and semantic heterogeneity. With the vast amounts of multimodal data generated daily, data spaces are expected to play a crucial role in the future of data management in general and smart cities in particular.

Knowledge graphs (KGs) are essential in data spaces as they offer a flexible and extensible way to store, organize, and represent data. [18] Multimodal KGs are becoming increasingly significant as they integrate different modalities, including text, image, audio, and video data, into a single graph, allowing for a more comprehensive representation of complex data. For example, a multimodal KG can model relationships between various data sources, such as people, vehicles, buildings, and the environment, providing a unified view of the data. Real-time analysis of this data can help identify patterns, such as identifying accident-prone areas in traffic flow and detecting unusual activities to alert authorities.

Scene graphs extract and represent objects, attributes, and relationships within multimedia data [4], providing a structured representation of unstructured data for integration into a multimodal KG. To address potential bias towards under-represented or unseen concepts, external common sense knowledge [29] can be incorporated to enrich scene graphs and make them more expressive and generalizable. Multimodal KGs and knowledge enrichment in data spaces have vast applications across multiple industries. In smart cities [45], they can enhance infrastructure planning, optimize traffic management and improve public safety by integrating data from various sources. Machine learning techniques and event processing approaches can then be applied to analyze this data. In healthcare [54], they facilitate the integration of patient data and structured clinical knowledge, supporting better diagnosis and treatment of diseases. Furthermore, multimodal KGs can be used for personalization in recommender systems, marketing, and education and



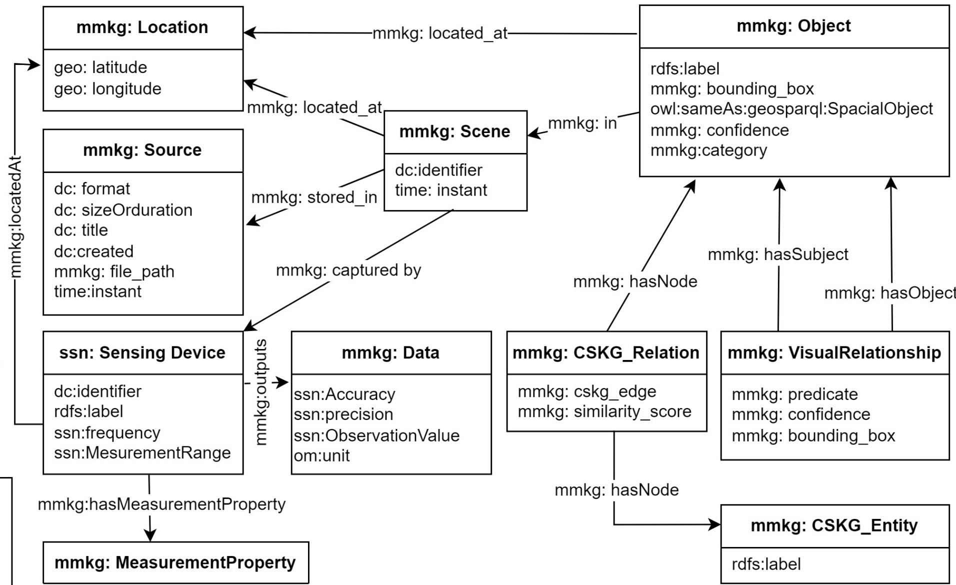
(a) Motivational Example: Image



(b) Motivational Example: Enriched SG

```
ASK{
EXISTS{?s1 rdfs:label "Person"} &&
EXISTS{?s2 rdfs:label "Ocean"} &&
EXISTS{?s rdfs:label "Drowning"}
}
```

(c) Motivational Example: Query



(d) Proposed MMKG Ontology

Figure 1: A Motivational Example and Proposed Ontology for Multimodal Knowledge Graphs for Data Spaces.

can benefit other industries such as finance, manufacturing, and education.

In this paper, we propose leveraging multimodal KGs to represent and reason about multimodal data streams in data spaces for smart cities and enhancing their expressiveness with external common sense knowledge. Figure 1(a-c) provides a motivational example that showcases how external background knowledge about arms of the person and the ocean can be leveraged to deduce that the person is drowning and requires immediate rescue. In Section 2, we reviewed related literature on data spaces, multimodal data representation and enrichment, and smart city ontologies. Section 3 presents the challenges associated with data spaces that multimodal KGs and knowledge enrichment can address, along with the key requirements for creating a multimodal KG and the proposed ontology. This is followed by conclusion in Section 4.

2 BACKGROUND

In this section, we review and summarize the recent literature on data spaces, multimodal data representation and knowledge enrichment, and smart city ontologies.

Dataspaces. Initially introduced by Halvey et al. [24], data spaces, unlike traditional database management systems, provide basic search functionality instead of full data control. Local data sources store the data, with semantic mappings or integration happening when needed. KGs and linked data are commonly used in dataspace to establish relationships between data sources. [7, 24] In the context of dataspace querying, [17], [8], and [36] provide insights into linked data, context-based querying, and historical data streams. Franklin et al. [16] define the logical components of a dataspace, which require "best-effort" and "approximate" answers, as well as conflict resolution strategies. TIKD [25] is a trusted dataspace that utilizes cryptography and blockchain for the secure sharing of healthcare-related data through KGs.

Multimodal Data Representation and Enrichment. Scene graph generation is widely used for extracting multimodal features from unstructured multimedia data, including images, videos and text, for scene understanding and structured representation [4]. It is used in semantic image retrieval [51], detection of civic issues in multimedia content [34], and visual question answering [33] that enable useful smart city applications. Multimodal KGs (MMKG), comprehensively reviewed in [37, 58], represent multimodal data either in attribute values or as entities. Here are some examples of MMKGs: GAIA [35] that extracts knowledge from text and video frames from multilingual news reports, RESIN [52] that is a cross-document cross-lingual, cross-media information extraction system with semantic elements extracted from text and images. Other MMKGs include IMGpedia [15], Image Graph [40], Richpedia [50], VisualSEM [2], NIEL [5].

Common sense knowledge infusion in scene graph generation is a growing trend that enhances representation quality and downstream task performance [29]. Previous approaches relied on language [38] and statistical priors [56]; the newer methods leverage knowledge graphs [27], while the state-of-the-art methods [30] are based on heterogeneous common sense knowledge sources, such as the Common Sense Knowledge Graph (CSKG) [26]. Approaches such as multimodal scene graphs [19], entity recognition [57], and CSKG-based enrichment [30] have been proposed to represent and analyze multimodal data. Multimodal event processing [9, 31] is an efficient way to handle large amounts of data, enabling real-time analytics and useful applications in smart cities.

Smart City Ontologies. Espinoza-Arias et al. [14] presented a comprehensive survey on smart city ontologies. De-Nicola et al. [13] provided ontology design patterns and requirements for seven domains, including Public Sector, Administrative Area, KPIs, City Object, Topology, Event, and Observations/Measurements. Ontologies related to smart city include Km4City [3], Fiesta-IoT [1], VITAL [28], Smart-City Ontology [32], CityPulse [43], Smart Cities and Emergency Management Ontology [12]. In addition, Prosumer-Oriented Smart Grid [20], Connected Traffic Data Ontology [49], and Semantic Sensor Ontology (SSN) [6] are also relevant ontologies. Ready4SmartCities [42] and Lov4IoT [22] are common catalogs, and CityGML [21] specifies 3D city models' interoperability and representation.

3 MULTIMODAL KNOWLEDGE GRAPHS FOR DATA SPACES

Data spaces provide a more flexible and scalable solution for handling large amounts of heterogeneous data from multiple, multimodal sources. By integrating structured sources, such as sensor data and databases, with unstructured sources, such as images, videos, and text, data spaces can provide detailed insights into complex patterns in huge amounts of data. Common sense knowledge enrichment can help reason about the entities in the data and their relationships, allowing for the identification of events that might not be immediately apparent from the data alone. Data spaces powered by multi-modal knowledge graphs can be utilized for effective traffic management and emergency response in smart cities by combining structured data from traffic sensors and GPS with unstructured data from traffic cameras and social media feeds.

For example, during a traffic accident, sensor data might indicate heavy traffic congestion. Unstructured data can provide additional details such as the exact location, number of vehicles involved, and injuries. By enriching the multi-modal representation with external knowledge, such as traffic flow patterns or the location of nearby hospitals, better decisions can be made on how to respond, like rerouting traffic, dispatching emergency services, or notifying hospitals to prepare for potential patients. Combining structured and unstructured data enriched with external knowledge can provide a more comprehensive understanding of events in smart cities, ultimately leading to better decision-making and problem-solving.

3.1 Challenges

Multimodal KGs and knowledge enrichment techniques can help address several challenges associated with data spaces:

(1) **Data Integration:** One of the primary challenges of data spaces is integrating data from multiple sources [25]. Multimodal KGs can help integrate information from various structured and unstructured data sources and external knowledge sources, providing a unified representation of and semantically rich insights into the data.
(2) **Data Governance:** It is also a major challenge to ensure all the data is managed to comply with regulations, policies, and standards [48]. Knowledge enrichment techniques can be used to annotate data in a multimodal KG with metadata that describes the provenance, ownership, and usage of data to make sure data is managed in compliance with regulations, policies, and standards.
(3) **Scalability:** As data volumes continue to grow, data spaces need to be able to handle increasingly large and complex datasets [41]. Multimodal KGs can represent large and complex datasets by efficiently organizing data into a network of interlinked entities for seamless scalability and efficient data retrieval and analysis.
(4) **Interoperability:** Data spaces involve integrating data from various systems and platforms, which requires data to be interoperable and exchangeable seamlessly between systems [47]. Multimodal KGs can provide a common data model representing data consistently across different applications and platforms to ensure data interoperability.
(5) **Semantic Interoperability:** Semantic interoperability involves ensuring that data is represented in a way that is understandable and interpretable by humans and machines [46]. Multimodal KGs can help achieve semantic interoperability by providing a standardized data representation using ontologies and vocabularies to represent data to ensure that data is understood and interpreted consistently across different systems and applications.
(6) **Data Security, Privacy and Ethics:** To establish a trusted network for data exchange and sharing, it is crucial to ensure secure data access and restrictions, including confidentiality, digital rights management, and secure access control, even within a decentralized peer-to-peer network [25, 53]. Multimodal KGs can incorporate digital rights management frameworks and access control mechanisms to secure access to sensitive data while ensuring confidentiality. Knowledge enrichment techniques can verify user identity and authorization, providing an additional layer of security. Standardized security solutions and exchange protocols can be implemented across all nodes and participants in the data-sharing space. Additionally, ethical and privacy-related external knowledge can be infused into multimodal KGs, such as information about data sensitivity, usage restrictions, and user consent, to ensure data is collected and used ethically and responsibly.
(7) **Data Lifecycle Management:** Data lifecycle management in data spaces is not designed around sharing. Existing models need improvement to prepare data for sharing and address the complexity of different data types. [10] Multimodal KGs and knowledge enrichment techniques can provide a standardized way of representing and annotating data throughout its lifecycle. By incorporating metadata, improving data quality, and enabling cross-domain interoperability, multimodal KGs can help prepare data for sharing and improve data management in data spaces.
(8) **Data Usage Rights:** Data producers need to retain their ownership rights to control who can use their data, and under which terms and conditions [39]. Multimodal KGs can standardize data ownership, access rights, and usage restrictions in the representation and knowledge base. By identifying and classifying sensitive data, these techniques can ensure that usage rights are enforced in compliance with

legal and regulatory requirements, enabling data owners to manage and control the usage of their data in data spaces.

(9) **Data Decentralization:** Decentralized data storage architectures require standard data exchange protocols to support data sharing and processing [55]. By incorporating metadata, multimodal KGs can provide a shared understanding of data across multiple platforms and enable interoperability between them. Knowledge enrichment techniques can help ensure the accuracy and completeness of data, enabling reliable data processing in decentralized environments. Standard data exchange protocols can also be defined based on the KG representation, enabling seamless data exchange across different platforms.

(10) **Data Veracity:** Data veracity is crucial for the sustainability of data-sharing ecosystems, but weak verification and provenance support hinder trust and transparency [10]. To address this challenge, multimodal KGs can enable traceability and transparency in data sharing ecosystems by incorporating metadata about data origins, processing steps, and algorithms. Knowledge enrichment/infusion techniques can also verify data accuracy and integrate provenance tracking into the KG, creating a comprehensive audit trail of data usage.

3.2 Requirements

Key requirements for designing and constructing a multimodal KG for data spaces are as follows:

(1) **Structured Representation:** To integrate unstructured visual data with structured sensor and location data in the unified KG, a structured representation of visual data is necessary, which requires accurate detection and effective linking of semantic elements in visual data.
(2) **Formal Ontology:** A formal ontology is required to represent multimodal smart city data effectively in the data space.
(3) **Expressiveness:** Incorporating an external knowledge base is necessary to improve the KG's expressiveness and align its ontology with the external knowledge base for consistency and interoperability.
(4) **Optimized Storage and Management:** Multimodal KG needs to be stored and managed using a graph database optimized for handling graph-structured data and unstructured data files like images and videos need to be stored in an object store.
(5) **Efficient and Expressive Querying and Matching:** Multimodal KG needs to support efficient querying and matching through a graph query language, allowing for complex pattern matching based on various data types and external knowledge base.
(6) **Real-time Multimodal Stream Processing:** Real-time processing of multimodal data streams requires complex event processing to analyze large volumes of incoming data, detect patterns, and trigger actions based on the detected patterns.
(7) **Visualization and Reporting:** Tools for data visualization and reporting are necessary to explore and analyze data in an intuitive and informative way.

3.3 Proposed Ontology

The proposed multimodal KG integrates visual, sensor and location data from sensors and cameras in a smart city data space, enriched with contextual knowledge from an external knowledge base. The ontology, shown in Figure 1(d), serves as a foundation for representing, querying, and analyzing multimodal data in the data space. Existing ontologies, including SSN[1], OM [44], OWL-Time[2], Geo[3], GeoSPARQL[4], Dublin Core[5], and RDFS[6], were reused, following best practices for ontology design [23]. The ontology defines the main classes and relationships, as well as their properties, for entities in the multimodal KG as listed below:

3.3.1 *Classes and Properties.*

(1) Related to visual data
   - ***mmkg:object*** represents the objects detected by object detectors on video data captured by the cameras, for example, a person, a car, a building, a tree etc. They are spatial objects as defined in `geosparql:SpacialObject` and have similar properties, hence we link them with the help of a "sameAs" relation (`owl:sameAs`). Furthermore, the boundaries of an `mmkg:Object` are defined with the help of a bounding box (`mmkg:bounding_box`). A bounding box is represented by four coordinates forming a rectangle or `geosparql:coordinateDimension` or the number of measurements or axes needed to describe the position of this geometry in a coordinate system. Other properties of `mmkg:object` are `rdfs:label` or the label assigned by object detectors to the object and `mmkg:confidence` which is the classification score between 0 and 1.
   - ***mmkg:scene*** represents a scene captured by the camera. A scene contains multiple objects inside it (`mmkg:in`). Other properties include `dc:identifier` and `time:instant` to represent the timestamp when the scene was captured.
   - ***mmkg:Location*** represents the geographical location of different classes of mmkg like `mmkg:object, mmkg:scene` and `ssn:SensingDevice` using the relation `mmkg:locatedAt`. The location is represented using properties `geo:latitude` and `geo:longitude` to represent the lattitude and logitude of a geographical point (`geo:point`).
   - ***mmkg:source*** represents the unstructured data files received from cameras, such as images. Properties used to describe the source include `dc:sizeOrduration, dc:title, dc:format, dc:created` (represents the date the file was created), `time:instant`, and `mmkg:filepath` that represents the path of the file.
   - ***mmkg:visualRelationship*** represents a visual relationship linking two `mmkg:object` entities via an `mmkg:predicate` property along with `mmkg:confidence` and `mmkg:bounding_box` of the visual relationship.

(2) Related to text data:

[1] https://www.w3.org/2005/Incubator/ssn/ssnx/ssn#
[2] https://www.w3.org/2006/time
[3] https://www.w3.org/2003/01/geo/wgs84_pos#
[4] http://schemas.opengis.net/geosparql/1.0/geosparql_vocab_all.rdf
[5] https://www.dublincore.org/specifications/dublin-core/dcmi-terms/#
[6] https://www.w3.org/2001/sw/wiki/RDFS

  - ***ssn:SensingDevice*** is a device that implements sensing and represents the sensors and cameras in the smart city. It is modelled by reusing the SensingDevice class of ssn ontology (`ssn:SensingDevice`). The main properties used for this class are `ssn:MeasurementRange`, `dc:identifier`, `ssn:frequency` and `rdfs:label`.
  - ***mmkg:Data*** represents the data that the sensor outputs. This class is modelled using properties `ssn:precision`, `ssn:observationValue`, `ssn:Accuracy`, and `om:units` to represent the unit of measurement of the observation value.
  - ***ssn:MeasurementProperty*** represents the property being measured like air quality, water, length etc.

(3) Related to knowledge extracted from CSKG:
  - ***mmkg:CSKGentity*** represents the nodes extracted from CSKG that are semantically related to the visual concepts in a scene, for example, 'vehicle' and 'driving' nodes extracted from CSKG based on their relevance to a car object in a scene. It is defined by rdfs:label.
  - ***mmkg:CSKGrelation*** represents the edges extracted from CSKG that link mmkg:object and mmkg:CSKGentity based on their semantic similarity, for example, the relation 'isA' in the link (car, isA, vehicle). It has two properties: `mmkg:cskg_edge` which is based on `rdfs:label`, and `mmkg:similarity_score` which is the cosine similarity between the linked mmkg:object and mmkg:CSKGentity (ranging between 0 and 1).

3.3.2 Relationships.

(1) **mmkg:*outputs*** links a sensing device (ssn:SensingDevice) with its corresponding observations (mmkg:data).
(2) **mmkg:*hasMeasuringProperty*** is a relationship between a sensing device (ssn:SensingDevice) and the property it measures (mmkg:MeasurementProperty).
(3) **mmkg:*capturedBy*** is a relationship between `mmkg:scene` and `mmkg:sensor` .
(4) **mmkg:*in*** is a relation between classes `mmkg:object` and `mmkg:scene` to define the objects present inside a scene.
(5) **mmkg:*storedIn*** is a relationship between `mmkg:scene` and `mmkg:source`.
(6) **mmkg:*locatedAt*** is a relation between `mmkg:location` and other classes like `mmkg:scene`, `ssn:SensingDevice`, `mmkg:object`.
(7) **mmkg:*hasSubject*** links the predicate of a `mmkg:visualRelationship` to its subject `mmkg:object`.
(8) **mmkg:*hasObject*** links the predicate of a `mmkg:visualRelationship` to its object `mmkg:object`.
(9) **mmkg:*hasNode*** links an object (`mmkg:object`) and an entity (`mmkg:CSKGentity`), or interlinks two entities.

4 CONCLUSION

Multimodal KGs enriched with common sense knowledge have promising potential to enhance the effectiveness of modern data spaces by enabling semantically-rich integration and analysis of diverse data types and providing a formal representation of data across multiple domains, enabling automated reasoning, and facilitating enhanced data interoperability, management and governance. This paper provides a comprehensive overview of the literature on this line of work and presents the major challenges, key requirements and a proposed ontology for building effective multimodal KGs with knowledge enrichment for data spaces.

ACKNOWLEDGMENTS

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number 18/CRT/6223 and 12/RC/2289_P2. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

REFERENCES

[1] Rachit Agarwal, David Gomez Fernandez, Tarek Elsaleh, Amelie Gyrard, Jorge Lanza, Luis Sanchez, Nikolaos Georgantas, and Valerie Issarny. 2016. PUnified IoT Ontology To Enable Interoperability and Federation of Testbeds. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*. IEEE, 70–75.

[2] Houda Alberts, Teresa Huang, Yash Deshpande, Yibo Liu, Kyunghyun Cho, Clara Vania, and Iacer Calixto. 2020. Visualsem: A High-Quality Knowledge Graph for Vision and Language. *arXiv preprint arXiv:2008.09150* (2020).

[3] Pierfrancesco Bellini, Monica Benigni, Riccardo Billero, Paolo Nesi, and Nadia Rauch. 2014. Km4City Ontology Building vs Data Harvesting and Cleaning for Smart-City Services. *Journal of Visual Languages & Computing* 25, 6 (2014), 827–839.

[4] Xiaojun Chang, Pengzhen Ren, Pengfei Xu, Zhihui Li, Xiaojiang Chen, and Alex Hauptmann. 2023. A Comprehensive Survey of Scene Graphs: Generation and Application. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 1 (2023), 1–26. https://doi.org/10.1109/TPAMI.2021.3137605

[5] Xinlei Chen, Abhinav Shrivastava, and Abhinav Gupta. 2013. Neil: Extracting Visual Knowledge From Web Data. In *Proceedings of the IEEE international conference on computer vision*. 1409–1416.

[6] Michael Compton, Payam Barnaghi, Luis Bermudez, Raul Garcia-Castro, Oscar Corcho, Simon Cox, John Graybeal, Manfred Hauswirth, Cory Henson, Arthur Herzog, et al. 2012. The SSN Ontology of the W3C Semantic Sensor Network Incubator Group. *Journal of Web Semantics* 17 (2012), 25–32.

[7] Edward Curry, Wassim Derguech, Souleiman Hasan, Christos Kouroupetroglou, and Umair ul Hassan. 2019. A Real-Time Linked Dataspace for the Internet of Things: Enabling "Pay-As-You-Go" Data Management in Smart Environments. *Future Generation Computer Systems* 90 (2019), 405–422.

[8] Edward Curry, André Freitas, and Seán O'Riáin. 2020. Querying and Searching Heterogeneous Knowledge Graphs in Real-time Linked Dataspaces. *Real-time Linked Dataspaces: Enabling Data Ecosystems for Intelligent Systems* (2020), 105–124.

[9] Edward Curry, Dhaval Salwala, Praneet Dhingra, Felipe Arruda Pontes, and Piyush Yadav. 2022. Multimodal Event Processing: A Neural-Symbolic Paradigm for the Internet of Multimedia Things. *IEEE Internet of Things Journal* 9, 15 (2022), 13705–13724.

[10] Edward Curry, Simon Scerri, and Tuomo Tuikka. 2022. *Data Spaces. (Design, Deployment and Future Directions)*. Springer.

[11] Edward Curry, Tuomo Tuikka, Andreas Metzger, Sonja Zillner, Natalie Bertels, Charlotte Ducuing, Davide Dalle Carbonare, Sergio Gusmeroli, Simon Scerri, Irene López de Vallejo, et al. 2022. Data Sharing Spaces: The BDVA Perspective. In *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage*. Springer International Publishing Cham, 365–382.

[12] Antonio De Nicola, Michele Melchiori, and Maria Luisa Villani. 2019. Creative Design of Emergency Management Scenarios Driven by Semantics: An Application to Smart Cities. *Information Systems* 81 (2019), 21–48.

[13] Antonio De Nicola and Maria Luisa Villani. 2021. Smart City Ontologies and Their Applications: A Systematic Literature Review. *Sustainability* 13, 10 (2021), 5578.

[14] Paola Espinoza-Arias, María Poveda-Villalón, Raúl García-Castro, and Oscar Corcho. 2018. Ontological Representation of Smart City Data: From Devices to Cities. *Applied Sciences* 9, 1 (2018), 32.

[15] Sebastián Ferrada, Benjamin Bustos, and Aidan Hogan. 2017. IMGpedia: A Linked Dataset With Content-Based Analysis of Wikimedia Images. In *The Semantic Web–ISWC 2017: 16th International Semantic Web Conference, Vienna, Austria, October 21-25, 2017, Proceedings, Part II 16*. Springer, 84–93.

[16] Michael Franklin, Alon Halevy, and David Maier. 2005. From Databases to Dataspaces: A New Abstraction for Information Management. *ACM Sigmod Record* 34, 4 (2005), 27–33.

[17] André Freitas, Edward Curry, João Gabriel Oliveira, and Sean O'Riain. 2011. Querying Heterogeneous Datasets on the Linked Data Web: Challenges, Approaches, and Trends. *IEEE Internet Computing* 16, 1 (2011), 24–33.

[18] André Freitas, Seán O'Riáin, and Edward Curry. 2020. *Querying and Searching Heterogeneous Knowledge Graphs in Real-Time Linked Dataspaces*. Springer International Publishing, Cham.

[19] Ze Fu, Changmeng Zheng, Junhao Feng, Yi Cai, Xiao-Yong Wei, Yaowei Wang, and Qing Li. 2022. DRAKE: Deep Pair-Wise Relation Alignment for Knowledge-Enhanced Multimodal Scene Graph Generation in Social Media Posts. *IEEE Transactions on Circuits and Systems for Video Technology* (2022).

[20] Syed Gillani, Frederique Laforest, Gauthier Picard, et al. 2014. A Generic Ontology for Prosumer-Oriented Smart Grid.. In *International Conference on Extending Database Technology (EDBT)/International Conference on Database Theory (ICDT) Workshops*, Vol. 1133. 134–139.

[21] Gerhard Gröger and Lutz Plümer. 2012. CityGML–Interoperable Semantic 3D City Models. *ISPRS Journal of Photogrammetry and Remote Sensing* 71 (2012), 12–33.

[22] Amelie Gyrard, Christian Bonnet, Karima Boudaoud, and Martin Serrano. 2016. LOV4IoT: A Second Life for Ontology-Based Domain Knowledge To Build Semantic Web of Things Applications. In *2016 IEEE 4th international conference on future internet of things and cloud (FiCloud)*. IEEE, 254–261.

[23] Amelie Gyrard, Antoine Zimmermann, and Amit Sheth. 2018. Building IoT-Based Applications for Smart Cities: How Can Ontology Catalogs Help? *IEEE Internet of Things Journal* 5, 5 (2018), 3978–3990.

[24] Alon Halevy, Michael Franklin, and David Maier. 2006. Principles of Dataspace Systems. In *Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. 1–9.

[25] Julio Hernandez, Lucy McKenna, and Rob Brennan. 2022. TIKD: A Trusted Integrated Knowledge Dataspace for Sensitive Data Sharing and Collaboration. In *Data Spaces: Design, Deployment and Future Directions*. Springer International Publishing Cham, 265–291.

[26] Filip Ilievski, Pedro Szekely, and Bin Zhang. 2021. CSKG: The Commonsense Knowledge Graph. In *The Semantic Web: 18th International Conference, ESWC 2021, Virtual Event, June 6–10, 2021, Proceedings 18*. Springer, 680–696.

[27] Xuan Kan, Hejie Cui, and Carl Yang. 2021. Zero-Shot Scene Graph Relation Prediction Through Commonsense Knowledge Integration. In *Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part II 21*. Springer, 466–482.

[28] Aqeel Kazmi, Zeeshan Jan, Achille Zappa, and Martin Serrano. 2017. Overcoming the Heterogeneity in the Internet of Things for Smart Cities. In *Interoperability and Open-Source Solutions for the Internet of Things: Second International Workshop, InterOSS-IoT 2016, Held in Conjunction with IoT 2016, Stuttgart, Germany, November 7, 2016, Invited Papers 2*. Springer, 20–35.

[29] Muhammad Jaleed Khan, John G Breslin, and Edward Curry. 2022. Common Sense Knowledge Infusion for Visual Understanding and Reasoning: Approaches, Challenges, and Applications. *IEEE Internet Computing* 26, 4 (2022), 21–27.

[30] Muhammad Jaleed Khan, John G Breslin, and Edward Curry. 2022. Expressive Scene Graph Generation Using Commonsense Knowledge Infusion for Visual Understanding and Reasoning. In *The Semantic Web: 19th International Conference, ESWC 2022, Hersonissos, Crete, Greece, May 29–June 2, 2022, Proceedings*. Springer, 93–112.

[31] Muhammad Jaleed Khan and Edward Curry. 2020. Neuro-Symbolic Visual Reasoning for Multimedia Event Processing: Overview, Prospects and Challenges.. In *CIKM (Workshops)*.

[32] Nicos Komninos, Charalampos Bratsas, Christina Kakderi, and Panagiotis Tsarchopoulos. 2016. Smart City Ontologies: Improving the Effectiveness of Smart City Applications. *Journal of Smart Cities (Transferred)* 1, 1 (2016).

[33] Rajat Koner, Hang Li, Marcel Hildebrandt, Deepan Das, Volker Tresp, and Stephan Günnemann. 2021. Graphhopper: Multi-Hop Scene Graph Reasoning for Visual Question Answering. In *The Semantic Web–ISWC 2021: 20th International Semantic Web Conference, ISWC 2021, Virtual Event, October 24–28, 2021, Proceedings 20*. Springer, 111–127.

[34] Shanu Kumar, Shubham Atreja, Anjali Singh, and Mohit Jain. 2019. Adversarial Adaptation of Scene Graph Models for Understanding Civic Issues. In *The World Wide Web Conference*. 2943–2949.

[35] Manling Li, Alireza Zareian, Ying Lin, Xiaoman Pan, Spencer Whitehead, Brian Chen, Bo Wu, Heng Ji, Shih-Fu Chang, Clare Voss, et al. 2020. Gaia: A Fine-Grained Multimedia Knowledge Extraction System. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 77–86.

[36] Yukun Li and Xiaofeng Meng. 2009. Supporting Context-Based Query in Personal Data Space. In *Proceedings of the 18th ACM conference on Information and knowledge management*. 1437–1440.

[37] Ye Liu, Hui Li, Alberto Garcia-Duran, Mathias Niepert, Daniel Onoro-Rubio, and David S Rosenblum. 2019. MMKG: Multi-Modal Knowledge Graphs. In *The Semantic Web: 16th International Conference, ESWC 2019, Portorož, Slovenia, June 2–6, 2019, Proceedings 16*. Springer, 459–474.

[38] Cewu Lu, Ranjay Krishna, Michael Bernstein, and Li Fei-Fei. 2016. Visual Relationship Detection With Language Priors. In *European Conference on Computer Vision*. Springer, 852–869.

[39] Christoph Mertens, Jesús Alonso, Oscar Lázaro, Charaka Palansuriya, Gernot Böge, Alexandros Nizamis, Vaia Rousopoulou, Dimosthenis Ioannidis, Dimitrios Tzovaras, Rizkallah Touma, et al. 2022. A Framework for Big Data Sovereignty: The European Industrial Data Space (EIDS). In *Data Spaces: Design, Deployment and Future Directions*. Springer International Publishing Cham, 201–226.

[40] Daniel Oñoro-Rubio, Mathias Niepert, Alberto García-Durán, Roberto Gonzalez-Sanchez, and Roberto Javier López-Sastre. 2019. Answering Visual-Relational Queries in Web-Extracted Knowledge Graphs. In *AKBC*.

[41] Mihalis Papakonstantinou, Manos Karvounis, Giannis Stoitsis, and Nikos Manouselis. 2022. Deploying a Scalable Big Data Platform to Enable a Food Safety Data Space. In *Data Spaces: Design, Deployment and Future Directions*. Springer International Publishing Cham, 227–248.

[42] María Poveda Villalón, Raúl García Castro, and A Gómez-Pérez. 2014. Building an Ontology Catalogue for Smart Cities. In *In Proceedings of 10th European Conference on Product & Process Modelling (ECPPM 2014), Vienna, Austria*. CRC Press, 829–839.

[43] Dan Puiu, Payam Barnaghi, Ralf Tönjes, Daniel Kümper, Muhammad Intizar Ali, Alessandra Mileo, Josiane Xavier Parreira, Marten Fischer, Sefki Kolozali, Nazli Farajidavar, et al. 2016. Citypulse: Large Scale Data Analytics Framework for Smart Cities. *IEEE Access* 4 (2016), 1086–1108.

[44] Hajo Rijgersberg, Mark Van Assem, and Jan Top. 2013. Ontology of Units of Measure and Related Concepts. *Semantic Web* 4, 1 (2013), 3–13.

[45] Iqbal H Sarker. 2022. Smart City Data Science: Towards Data-Driven Smart Cities With Open Research Issues. *Internet of Things* 19 (2022), 100528.

[46] Martín Serrano, Edward Curry, Richards Walsh, Gavin Purtill, John Soldatos, Maurizio Ferraris, and Ernesto Troiano. 2022. Data Space Best Practices for Data Interoperability in FinTechs. In *Data Spaces: Design, Deployment and Future Directions*. Springer International Publishing Cham, 249–264.

[47] Gürkan Solmaz, Flavio Cirillo, Jonathan Fürst, Tobias Jacobs, Martin Bauer, Ernö Kovacs, Juan Ramón Santana, and Luis Sánchez. 2022. Enabling Data Spaces: Existing Developments and Challenges. In *Proceedings of the 1st International Workshop on Data Economy*. 42–48.

[48] Ana I. Torre-Bastida, Guillermo Gil, Raúl Miñón, and Josu Díaz-de Arcaya. 2022. *Technological Perspective of Data Governance in Data Space Ecosystems*. Springer International Publishing, Cham, 65–87.

[49] Miloš Viktorović, Dujuan Yang, and Bauke de Vries. 2020. Connected Traffic Data Ontology (CTDO) for Intelligent Urban Traffic Systems Focused on Connected (Semi) Autonomous Vehicles. *Sensors* 20, 10 (2020), 2961.

[50] Meng Wang, Haofen Wang, Guilin Qi, and Qiushuo Zheng. 2020. Richpedia: A Large-Scale, Comprehensive Multi-Modal Knowledge Graph. *Big Data Research* 22 (2020), 100159.

[51] Sijin Wang, Ruiping Wang, Ziwei Yao, Shiguang Shan, and Xilin Chen. 2020. Cross-Modal Scene Graph Matching for Relationship-Aware Image-Text Retrieval. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 1508–1517.

[52] Haoyang Wen, Ying Lin, Tuan Lai, Xiaoman Pan, Sha Li, Xudong Lin, Ben Zhou, Manling Li, Haoyu Wang, Hongming Zhang, et al. 2021. Resin: A Dockerized Schema-Guided Cross-Document Cross-Lingual Cross-Media Information Extraction and Event Tracking System. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations*. 133–143.

[53] Akshika Wijesundara. 2022. *Privacy-Aware Smart Home Interface Framework. PhD thesis*. Ph. D. Dissertation. https://doi.org/doi:10.21954/ou.ro.0001412f

[54] Guang Yang, Qinghao Ye, and Jun Xia. 2022. Unbox the Black-Box for the Medical Explainable AI via Multi-Modal and Multi-Centre Data Fusion: A Mini-Review, Two Showcases and Beyond. *Information Fusion* 77 (2022), 29–52.

[55] Achille Zappa, Chi-Hung Le, Martín Serrano, and Edward Curry. 2022. Connecting Data Spaces and Data Marketplaces and the Progress Toward the European Single Digital Market With Open-Source Software. In *Data Spaces: Design, Deployment and Future Directions*. Springer International Publishing Cham, 131–146.

[56] Rowan Zellers, Mark Yatskar, Sam Thomson, and Yejin Choi. 2018. Neural Motifs: Scene Graph Parsing With Global Context. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5831–5840.

[57] Fu J. Liu X. & Huang X. Zhang, Q. (2018). Adaptive Co-attention Network for Named Entity Recognition in Tweets.. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32 (1). https://doi.org/10.1609/aaai.v32i1.11962

[58] Xiangru Zhu, Zhixu Li, Xiaodan Wang, Xueyao Jiang, Penglei Sun, Xuwu Wang, Yanghua Xiao, and Nicholas Jing Yuan. 2022. Multi-Modal Knowledge Graph Construction and Application: A Survey. *IEEE Transactions on Knowledge and Data Engineering* (2022), 1–20. https://doi.org/10.1109/TKDE.2022.3224228