Index

1 RQ1: Additional Result Details

This section provides additional details about the results for each category and sub-category identified in RQ1. Figure F.1 illustrates the study distribution in terms of sub-categories, while Table T.1 reports the studies that fit each sub-category.

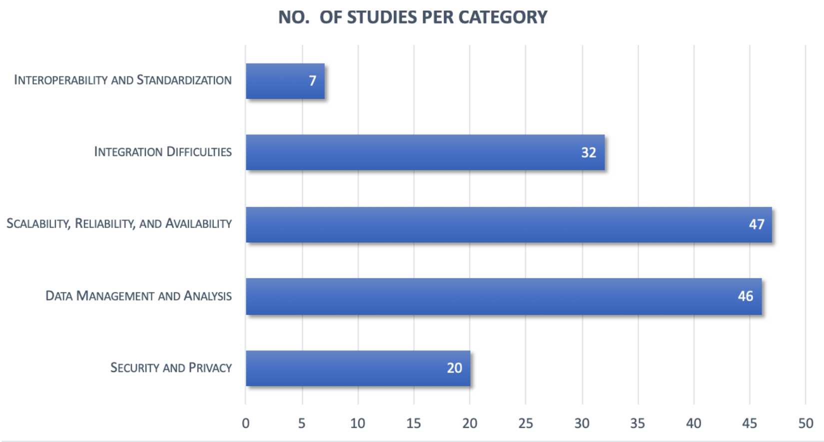

Figure F.1: RQ1: Categories distribution of the studies.

Table T.1: RQ1: Studies for categories and sub-categories

Category	Sub-Category	Studies
Security and Privacy	Cyber Threat Detection and Access Control	[SLR1, 2, 8, 10, 17, 19, 21, 28, 32, 36, 38, 43, 47, 48, 55]
	Data Privacy	[SLR14, 21, 29, 32, 34, 38, 43]
	Security in Communication	[SLR1, 28, 38, 39, 51, 55]
Data Management and Analysis	Efficient Data Processing and Analysis	[SLR 2-9, 11-18, 20-25, 27-38, 40, 42-45, 47-55]
	Semantic Data Handling	[SLR 2-4, 6, 7, 11-18, 21-25, 27-35, 37, 48-53]
	Data Classification and Categorization	[SLR 2-5, 8, 9, 11-14, 16-18, 24, 28, 31-36, 38, 42, 43, 45, 47, 49, 51-55]
	Data Integrationand Fusion	[SLR 4, 6, 7, 30, 37, 40, 54]
Scalability, Reliability, and Availability	Enhanced Scalability	[SLR 2-13, 15-20, 22-28, 30, 31, 33-39, 41-43, 46-49, 51, 53-55]
	Improved Reliability	[SLR 1, 2, 4, 8-10, 12, 15, 17-19, 22-24, 28, 30, 33-36, 38, 42, 44, 46-49, 51, 53-55]
	Increased Availability	[SLR 1, 11, 12, 34]
Integration Difficulties	Ease of System Integration and Configuration	SLR 2-4, 6-9, 11, 12, 15, 17, 19, 20, 22, 24-26, 28, 30, 33, 34, 39, 40, 41, 44-46, 50-53]
	Semantic Integration	[SLR 3, 6, 12, 18, 22, 24, 25, 28, 30, 41, 50, 51, 53]
	Technological Compatibility	[SLR 4, 26, 33, 39-41, 45]
Interoperability and Standardization	Protocol and Data Format Harmonization	[SLR 1, 40, 53]
	Service Discovery and Composition	[SLR 11-13, 31]

1.1 Security and Privacy

Figure F2 shows the distribution of the analyzed studies for the sub-categories of the "Security and Privacy" category.

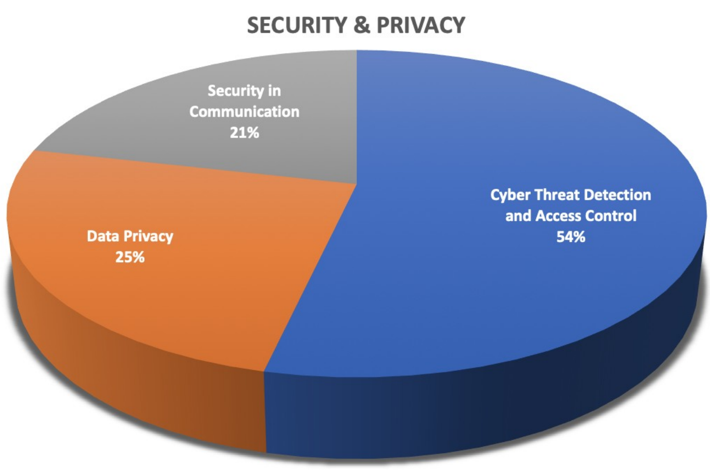

Figure F.2: RQ1: "Security and Privacy" papers' distribution.

1.2 Data Management and Analysis

Figure F.3 shows the distribution of the analyzed studies for the sub-categories of the "Data Management and Analysis" category.

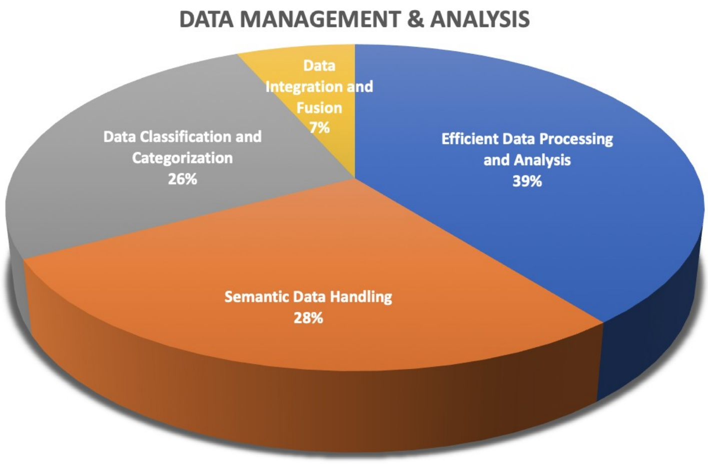

Figure F.3: RQ1: "Data Management And Analysis" papers' distribution.

1.3 Scalability, Reliability, and Availability

Figure F.4 highlights the distribution of the analyzed studies for the sub-categories of the "Scalability, Reliability, and Availability" category.

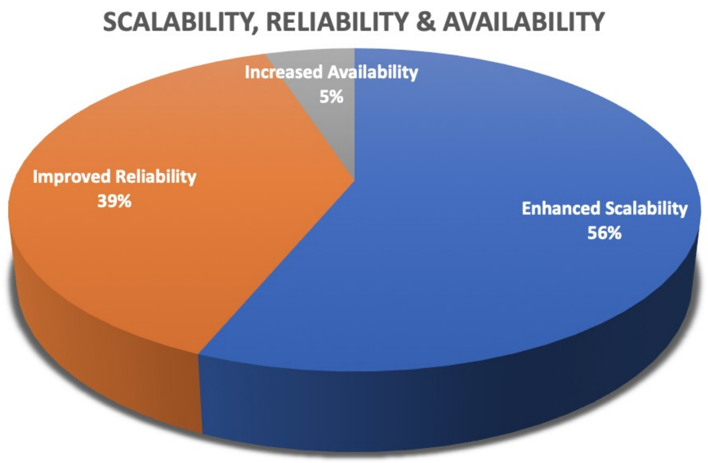

Figure F.4: RQ1: "Scalability, Reliability, Availability" papers' distribution.

1.4 Integration Difficulties

Figure F.5 illustrates the distribution of the analyzed studies for the sub-categories of the "Integration Difficulties" category.

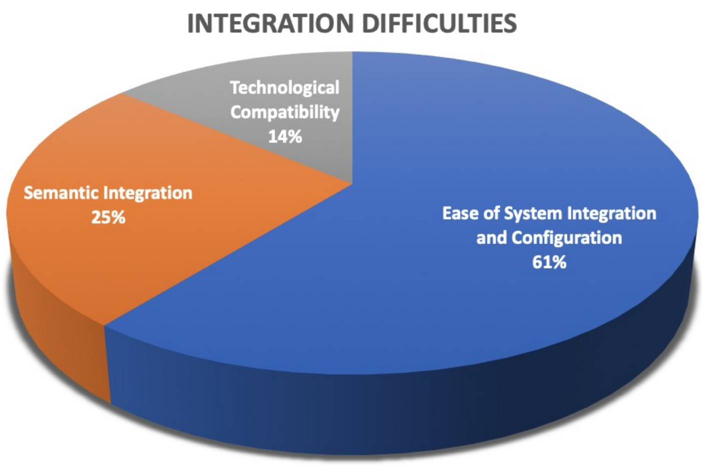

Figure F.5: RQ1: "Integration Difficulties" papers' distribution.

1.5 Interoperability and Standardization

Figure F.6 shows the distribution of the analyzed studies for the sub-categories of the “Interoperability and Standardization” category.

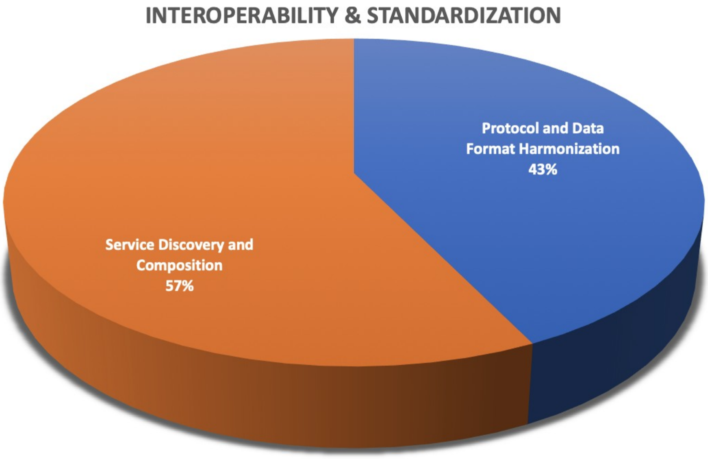

Figure F.6: RQ1: “Interoperability and Standardization” papers’ distribution.

2 RQ2: Additional Result Details

In this section, we provide additional result details for each category and sub-category identified in RQ2.

2.1 Pre-training and Fine-tuning Large Language Models

Figure F.7 shows the distribution of the studies that have used Pre-training and Fine-tuning to address IoT challenges. Tables T.2 and T.3 report the studies and the IoT domains in which Pre-training and Fine-tuning have been used to develop the LLM-based solution.

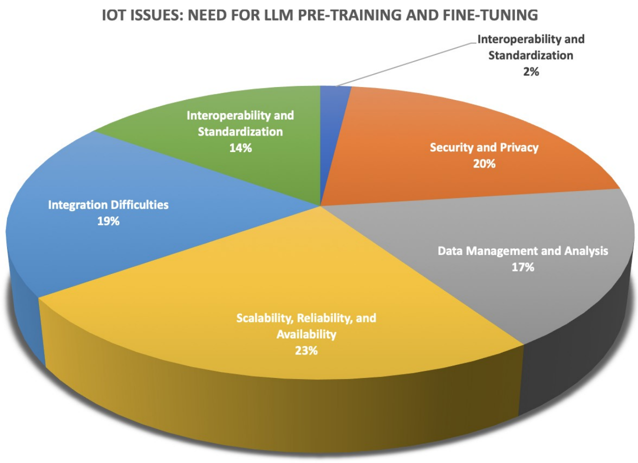

Figure F.7: RQ2: “IoT challenges where pre-training and fine-tuning have beenused”.

Table T.2: Studies in which Pre-training and Fine-tuning have been used

Strategyy	Studies
Pre-training	[SLR 1, 6, 10, 11, 13, 15, 20, 22, 24, 38, 39, 54]
Fine-tuning	[SLR 1-19, 22-25, 27-30, 32-34, 36, 38-40, 43, 46-48, 51, 53, 54]

Table T.3: IoT domains where LLMs have been pre-trained and/or fine-tuned.

IoT domain	Studies
Healthcare	[SLR 11, 14, 22-24, 29, 33, 34, 51, 54]
Smart Grid	[SLR 2, 9, 11, 22]
Transportation	[SLR 11, 12, 22, 28]
Smart Home and Building	[SLR 4, 7, 9-11, 14, 15, 18, 22, 30, 32, 33, 43, 53, 54]
Smart Cities	[SLR 7, 10-12, 14, 22, 36]
Agriculture	[SLR 11, 22]
Industry	[SLR 1, 4, 6, 8, 10, 11, 13, 22, 25, 27, 38, 40, 48]
Military	[SLR 11, 14, 17, 22, 46]
Other	[SLR 3, 5, 11, 14, 16, 19, 20, 22, 39, 47]

2.1 Software and Hardware Configurations

In the following sub-sections, we present further details for Software and Hardware configurations used for LLM-based solutions to address IoT challenges.

2.1.1 Software Configurations

This sub-section focuses on the Software Configurations of LLM-based solutions, in terms of LLM, Programming Languages, ML Frameworks, and Libraries used to develop the solutions.

Specifically, Table T.4 and Figure F.8 illustrate the LLM used in IoT applications.

Table T.4: LLMs used in IoT applications.

LLM	Studies
BERT or BERT variants	[SLR 1-3, 5-15, 17-19, 22-32, 34, 37, 38, 43, 47, 51, 53-55]
GPT-2	[SLR 1, 48]
GPT-3	[SLR 4, 21, 33, 36, 39-41, 50, 52
GPT-4	[SLR 4, 33, 34, 36, 40, 42, 44-46, 52]
Visual ChatGPT	[SLR 16]
DeviceGPT	[SLR 20]
ChatGLM	[SLR 35]
LLAMA	[SLR 40]
Generic	[SLR 49]

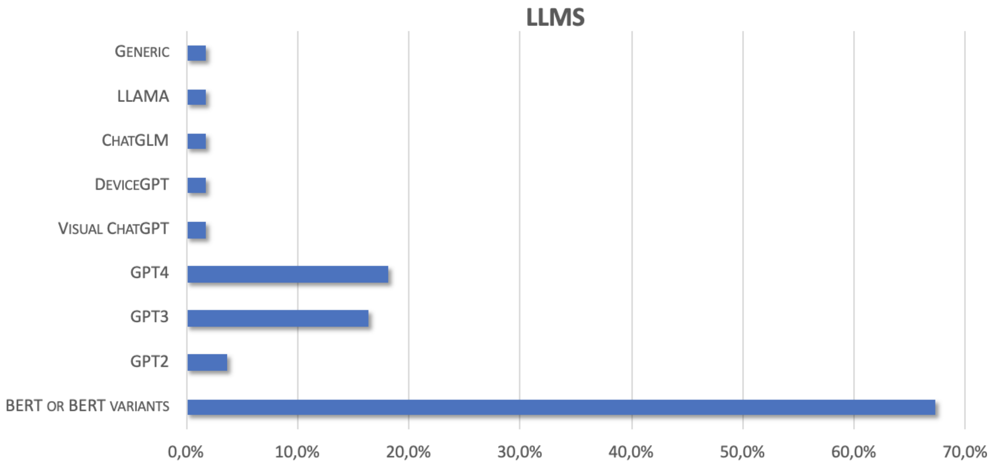

Figure F.8: RQ2: LLMs used to develop LLM-based solutions.

Table T.5 and Figure F.9 show the Programming Languages used to develop LMM-based solutions.

Table T.5: Programming Languages used in IoT applications.

Programming Language	Studies
Python	[SLR 1-11, 13-19, 21, 24-28, 32, 34-36, 38, 40, 43-45, 47, 50-52, 54, 55]
Java	[SLR 44]
C/C++	[SLR 45]

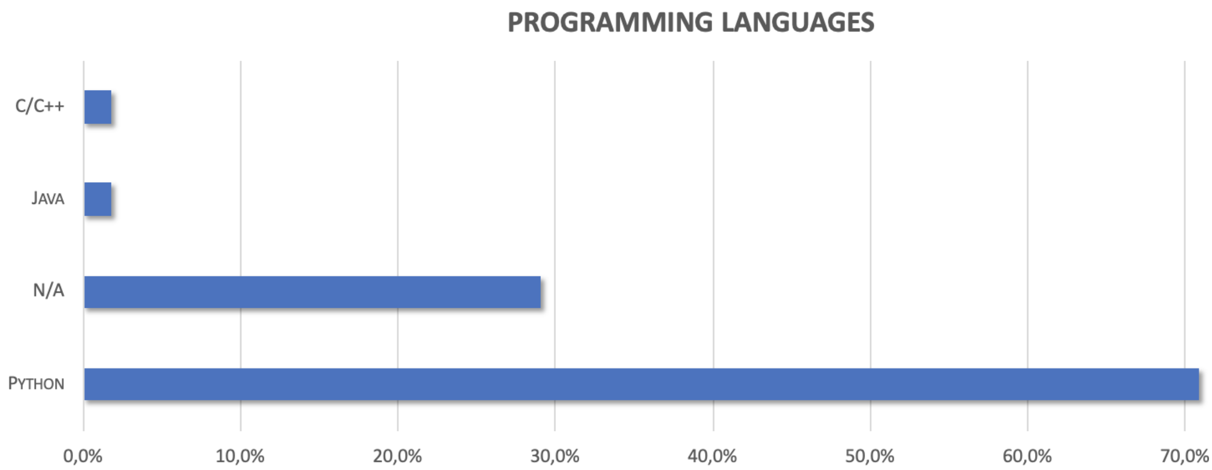

Figure F.9: RQ2: Programming Languages used to develop LLM-based solutions.

Table T.6 and Figure F.10 report the ML Framework used to develop LMM-based solutions.

Table T.6: Frameworks used in IoT applications.

Framework	Studies
PyTorch	[SLR 6, 19, 32, 51, 52, 54, 55]
TensorFlow	[SLR 3, 4, 8-10, 14, 15, 19, 25, 51]
HuggingFace Transformer	[SLR 1, 5, 10, 16, 17, 27, 28, 32, 38, 43]
Scikit-learn	[SLR 11, 13, 14, 16, 18, 25, 27, 32, 51]

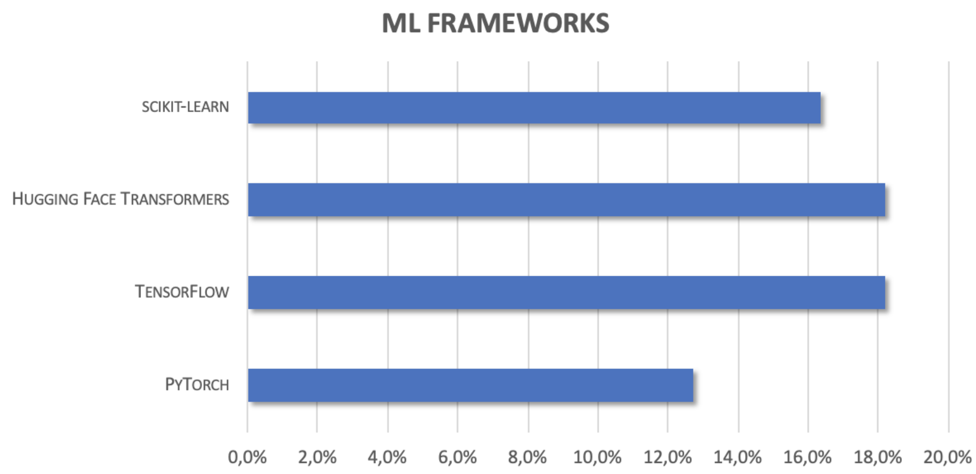

Figure F.10: RQ2: ML Frameworks used to develop LLM-based solutions.

Table T.7 and Figure F.11 illustrate the libraries used to develop LMM-based solutions.

Table T.7: Libraries used in IoT applications.

Library	Studies
Keras	[SLR 10, 14, 19, 28, 51]
NLTK	[SLR 1, 14, 27]
Jieba	[SLR 3, 15]
spaCy	[SLR 14, 18, 19, 21, 52]
OpeAI API	[SLR 4, 16, 21, 33, 34, 36, 40, 44, 45, 50]
Pandas	[SLR 14, 43, 44]
Gensim	[SLR 14, 27]
WikiExtractor	[SLR 15]
Opencc	[SLR 15]
Flask	[SLR 16, 50]
OpenCV	[SLR 16]
FastText	[SLR 27]
PineCone	[SLR 36]
WeightWatcher	[SLR 38]
LangChain	[SLR 40]
OpenPrompt	[SLR 52]

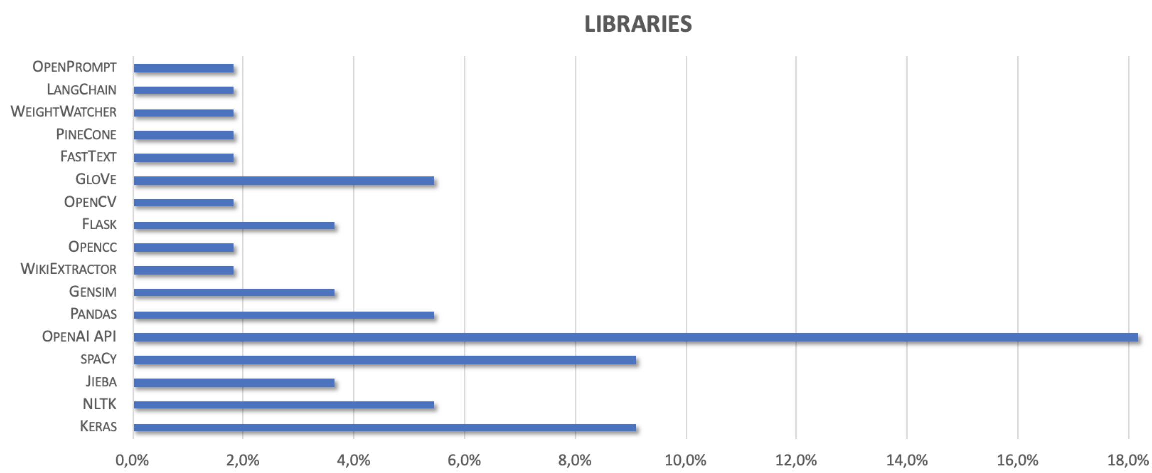

Figure F.11: RQ2: Libraries used to develop LLM-based solutions.

2.1.2 Hardware Configurations

This sub-section focuses on the Software Configurations of LLM-based solutions, in terms of LLM, Programming Languages, ML Frameworks, and Libraries used to develop the solutions.

Specifically, Table T.8 lists the hardware characteristics reported in the analyzed studies. At the same time, Figures F.12 and F.13 illustrate the distribution of the operating systems/environments and how the LLM-based solutions, which used GPUs and high-performant servers, contributed to addressing the different IoT issues.

Table T.8: Hardware used in IoT applications

Characteristic		Studies
Operating System	Windows 10	[SLR 2, 3, 8]
	Ubuntu	[SLR 1, 9, 10, 17]
	Cloud based	[SLR 4, 5, 7, 14, 50]
	Embedded	[SLR 24, 26, 33-35, 43, 45, 50, 54]
GPU		[SLR 1-3, 6, 8, 9, 11, 13-15, 17, 23, 25, 27, 28, 31, 38, 40, 47, 52, 54, 55]
Multi-core CPUs		[SLR 2, 6, 10, 17, 18, 23, 25, 27, 38, 47, 51, 52, 54]
RAM	Up to 32GB	[SLR 6, 10, 17, 18, 25, 27, 51]
	Up to 64GB	[SLR 40]
	Up to 128GB	[SLR 54]
	Up to 1TB	[SLR 40]

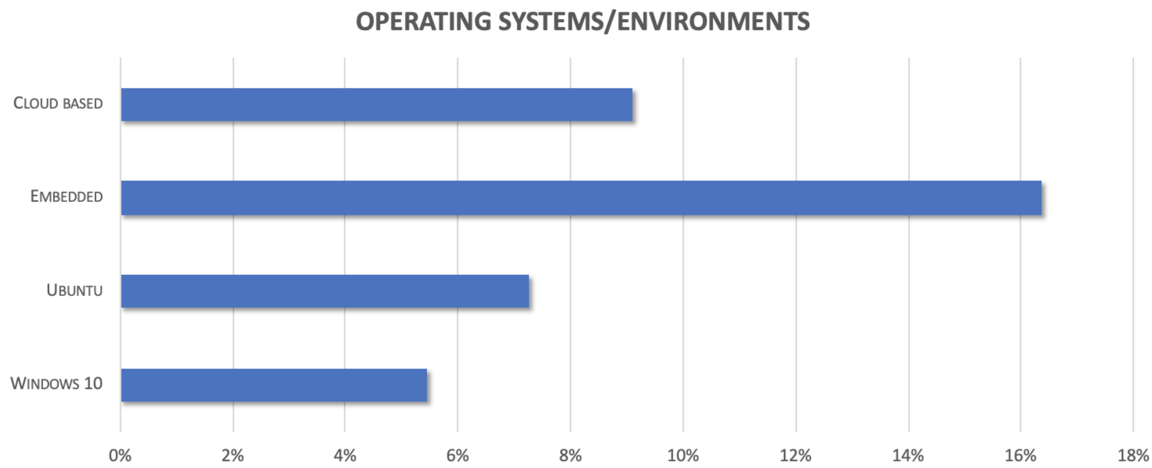

Figure F.12: RQ2: Operating Systems/Environments used to develop LLM- based solutions.

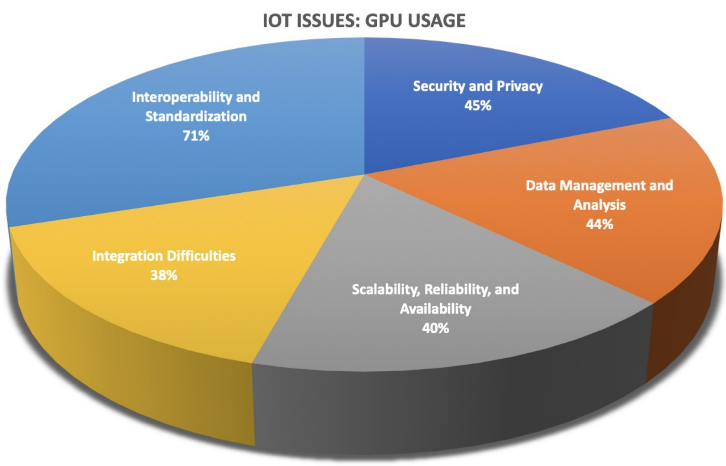

Figure F.13: RQ2: IoT Issues - GPU usage.

2.2 Use of Specific Techniques

In the following sub-sections, we present further details for Specific Techniques exploited to develop LLM-based solutions and address IoT challenges.

2.2.1 *Integration with Other Machine Learning Techniques*

Table T.9 reports the studies in which LLM-based solutions have been integrated with other ML techniques, while Figure F.14 shows how the combination of LLM with other ML techniques contributed to address IoT issues.

Table T.9: Integration with Other Machine Learning Techniques

Technique	Studies
Deep Learning	[SLR 3, 7, 9, 10, 14, 16, 19, 20, 22, 23, 25, 32, 34, 35, 36, 39, 41, 43, 48-51, 54
Traditional Machine Learning	[SLR 3, 4, 8, 13, 18, 24, 27, 31, 37, 44, 46, 47, 49, 52]

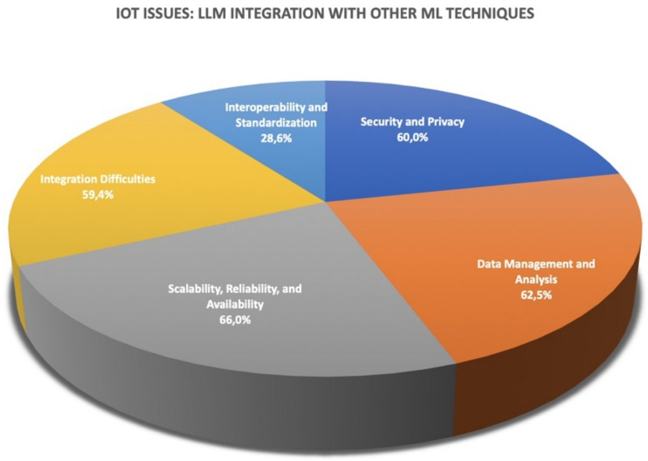

Figure F.14: RQ2: IoT Issues - LLM integration with other ML techniques.

2.2.2 *Optimization Techniques*

Table T.10 reports the studies in which Optimization techniques have been used to enhance LLM-based solutions, while Figure F.15 illustrates how the technique contributed to address IoT issues.

Table T.10: Studies in which Optimization Techniques have been used

Studies
[SLR 1, 5, 6, 9, 11, 12, 14, 17, 18, 25-32, 34, 46, 47, 51, 53-55]

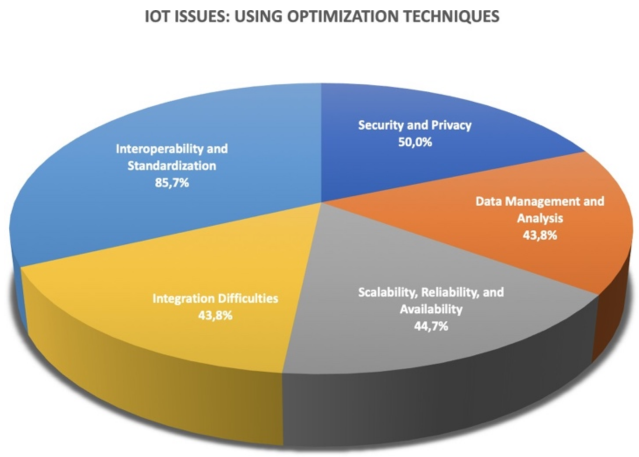

Figure F.15: RQ2: IoT Issues - Using Optimization techniques in LLM-based solutions.

2.2.3 *Extraction Techniques*

Table T.11 reports the extraction techniques used in LLM-based solutions, while Figure F.16 shows in which percentage the extraction techniques have been used to address the different types of IoT issues.

Table T.11: Extraction Techniques

Technique	Studies
Relation Extraction	[SLR 1, 7, 8, 33, 50, 52]
Name Entity Recognition	[SLR 1, 2, 7, 19, 21, 33, 41, 42, 46, 50, 51]
Feature Extraction for Sequential Data Analysis	[SLR 3-6, 9-18, 20, 22-24, 31, 33, 35, 36, 38, 45, 48, 49, 54, 55]
Feature Extraction for Textual Data Analysis	[SLR 24-30, 32-34, 37, 40-44, 46, 47, 50, 53]
Cross-modal feature extraction and alignment for data retrieval	[SLR 16, 39, 46, 50]

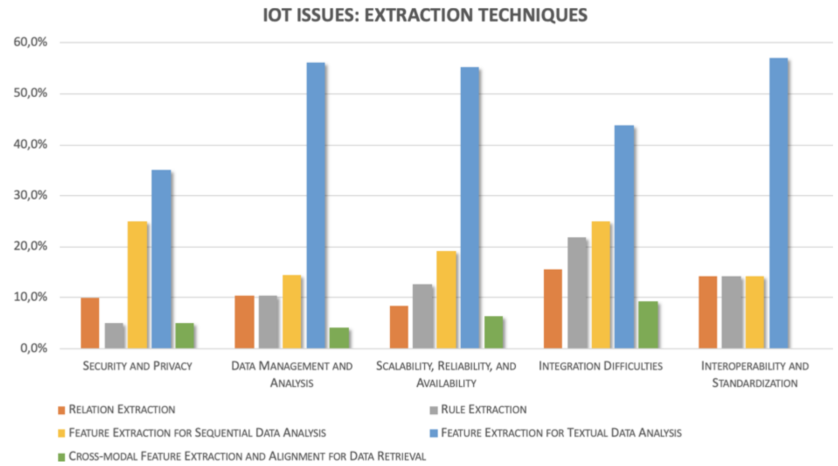

Figure F.16: RQ2: IoT Issues - Extraction Techniques used in LLM-based solutions.

2.3 Continuous Monitoring and Adjustment

Table T.12 reports the Continuous Monitoring and Adjustment techniques used in LLM-based solutions, while Figure F.17 shows the contribution of these approaches to address IoT challenges.

Table T.12: Studies in which Continuous Monitoring and Adjustment have been used

Technique	Studies
Real-time Monitoring and Adjustment	[SLR 2, 4, 5, 11, 18, 24, 35, 46, 51]
Adaptive Learning and Feedback Incorporation	[SLR 18, 29, 33, 35, 41, 44, 46, 49]

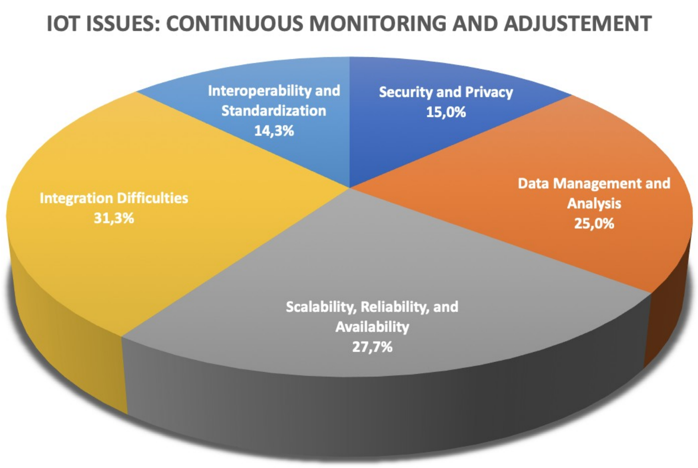

Figure F.17: RQ2: IoT Issues - Continuous Monitoring and Adjustment usage.

2.4 Data Collection and Preprocessing

In the following sub-sections, we present further details for Data Collection and PreprocessingTechniques exploited to develop LLM-based solutions and address IoT challenges.

2.4.1 Data Collection

Table T.13 lists the studies in which Data Collection has been discussed and the key aspects reported, while Figure F.18 shows the contribution of Data Collection in addressing IoT challenges.

Table T.13: Studies in which Data Collection have been used

Key aspect	Studies
Data Collection Reported	[SLR 1, 2, 4-25, 27-43, 47, 49, 50, 52-55]
Public Dataset Used	[SLR 1, 5, 7, 10, 13-18, 21-23, 27, 30-34, 36, 38, 39, 41, 47, 49, 50, 52, 55]
Automatic Data Collection	[SLR 2, 4, 6-9, 11, 12, 20, 24, 25, 28, 29, 35, 37, 40, 42, 43, 53, 54]
Manually Created Dataset	[SLR 19]

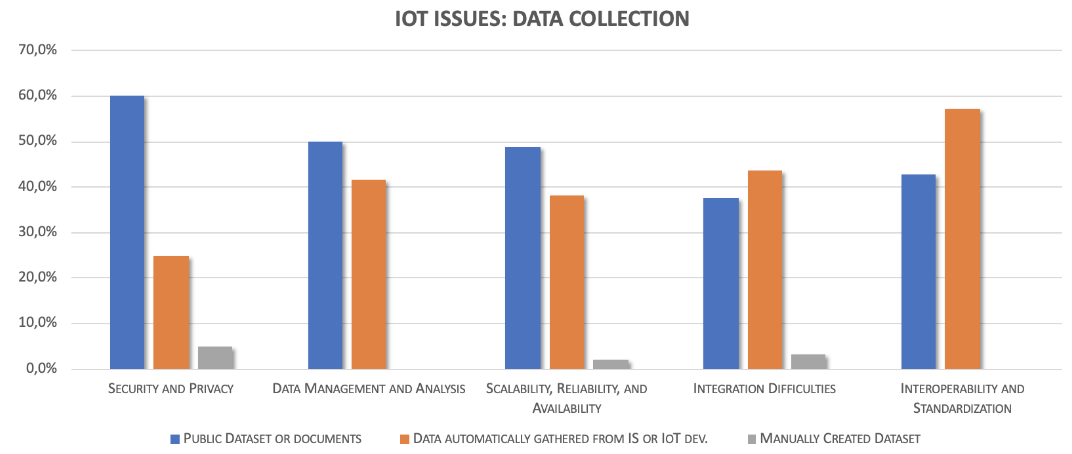

Figure F.18: RQ2: IoT issues - Data Collection approaches.

2.4.2 *Data Preprocessing*

Table T.14 lists the studies in which Data Preprocessing techniques has been used to develop LLM-based solutions, while Figure F.19 shows the contribution of Data Preprocessing in addressing IoT challenges.

Table T.14: Studies in which Data Preprocessing have been used

Technique	Studies
Data Transformation	[SLR 1, 3-7, 8, 10-21, 23, 24, 25, 27-29, 31-41, 43, 47, 49, 50-55]
Data Cleaning	[SLR 2, 6, 8, 14, 15, 24, 25, 27, 28, 33, 34, 40, 43, 47]

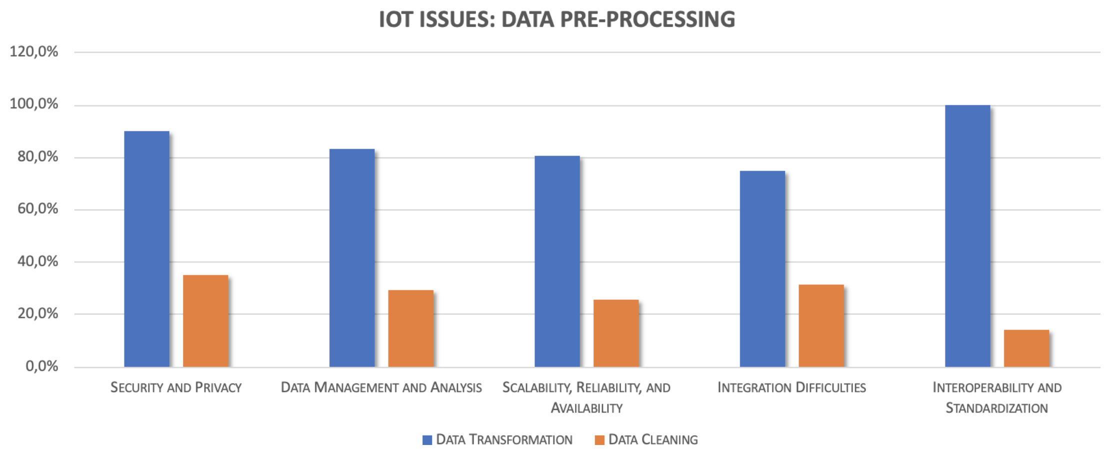

Figure F.19: RQ2: IoT issues - Data Preprocessing approaches.

2.4.3 *Data Labelling*

Table T.15 lists the studies in which Data Labelling techniques has been used to develop LLM-based solutions, while Figure F.20 shows the contribution of Data Labelling in addressing IoT challenges.

Table T.15: Studies in which Data Preprocessing have been used

Technique	Studies
Manual labelling	[SLR 1, 6, 8, 10-12, 16-22, 25, 29, 32-34, 42, 43, 45, 46, 51-54]
Automated or semi-automated labelling	[SLR 6, 13, 14, 27, 32, 35, 49]
Pre-labelled datasets	[SLR 3, 5, 7, 9, 14-16, 23, 24, 30, 36, 38, 39, 41, 47, 50, 55

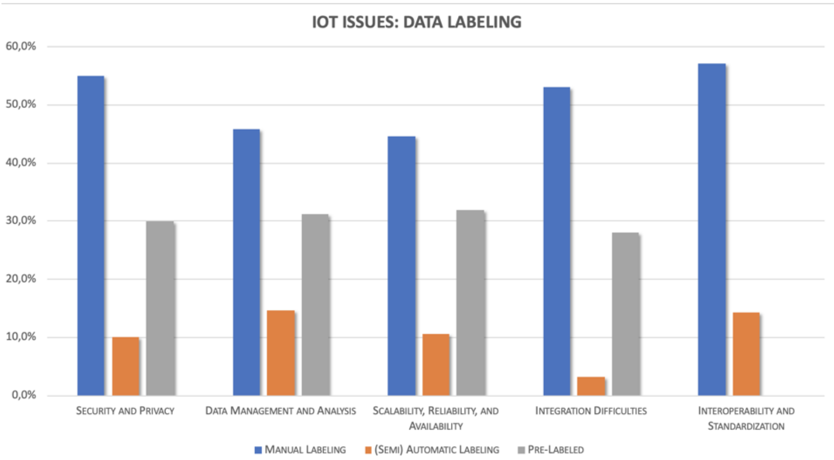

Figure F.20: RQ2: IoT issues - Data Labeling approaches.

2.5 *Implementation of LLM-based solutions*

Figure F.21 illustrates the general Workflow to develop LLM-based solutions to address IoT challenges.

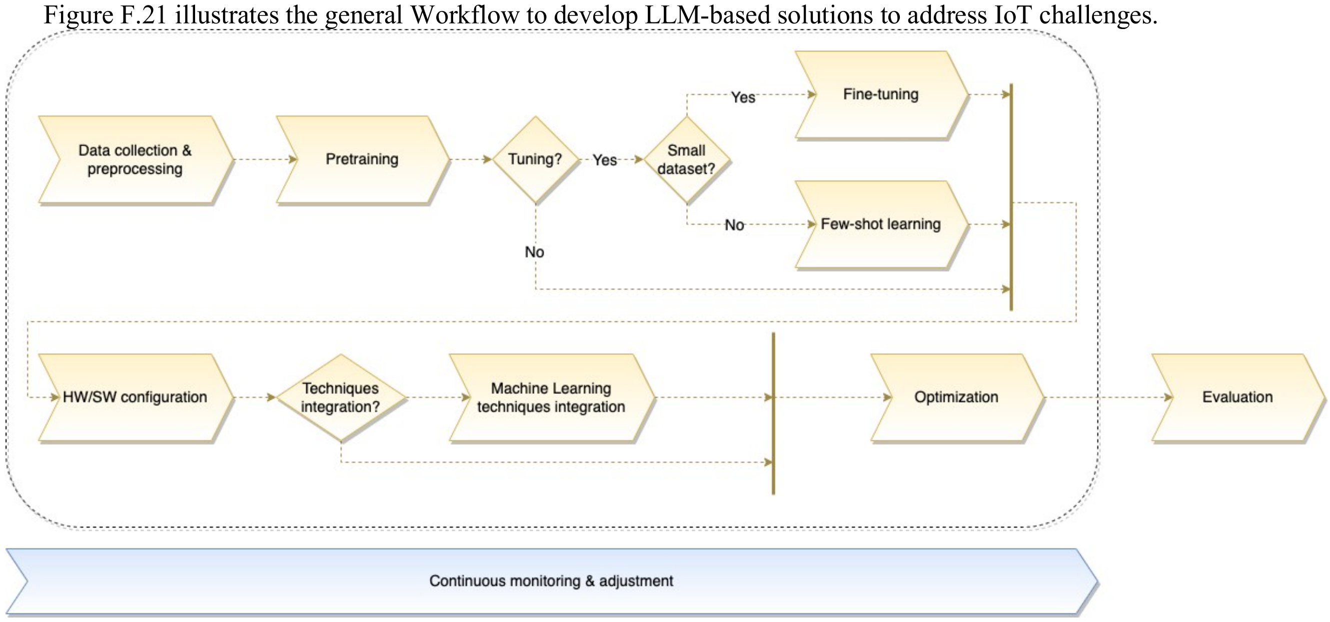

Figure F.21: Workflow to implement LLM-based solution in IoT applications.