Brussels, 19 May 2025

COST 030/25

DECISION

Subject: Memorandum of Understanding for the implementation of the COST Action “Knowledge Graphs in the Era of Large Language Models” (KGELL) CA24121

The COST Member Countries will find attached the Memorandum of Understanding for the COST Action Knowledge Graphs in the Era of Large Language Models approved by the Committee of Senior Officials through written procedure on 19 May 2025.

MEMORANDUM OF UNDERSTANDING

For the implementation of a COST Action designated as

COST Action CA24121
KNOWLEDGE GRAPHS IN THE ERA OF LARGE LANGUAGE MODELS (KGELL)

The COST Members through the present Memorandum of Understanding (MoU) wish to undertake joint activities of mutual interest and declare their common intention to participate in the COST Action, referred to above and described in the Technical Annex of this MoU.

The Action will be carried out in accordance with the set of COST Implementation Rules approved by the Committee of Senior Officials (CSO), or any document amending or replacing them.

The main aim and objective of the Action is to develop and evaluate an automated pipeline that uses LLMs to generate and refine Knowledge Graph content from domain-specific textual datasets, with a focus on reducing hallucinations and increasing factual accuracy. This will be achieved through the specific objectives detailed in the Technical Annex.

The present MoU enters into force on the date of the approval of the COST Action by the CSO.

OVERVIEW

Summary
Knowledge Graphs (KGs) have gained attention due to their ability to represent structured and interlinked information. KGs represent knowledge in the form of relations between entities, referred to as facts, typically grounded in formal ontological models. Such machine-readable formats enable AI systems to make decisions using clear and verifiable data. Consequently, KGs have become essential elements in web search engines, recommendation systems, etc. Large Language Models (LLMs) have revolutionized the landscape of AI and are widely utilized in various NLP tasks such as natural language understanding, question answering, etc. Despite their remarkable performance, LLMs suffer from some significant drawbacks. First, they are trained on general-purpose data and have lower performance in domain-specific tasks and low- resource languages. Secondly, they often reflect societal biases present in training data, which can result in biased outcomes. Third, LLMs sometimes produce inaccurate or made-up information, termed "hallucinations". Finally, understanding the decision-making process of LLMs is challenging and their outputs may lack consistency. A potential solution to all these problems is to integrate LLMs with KGs, since KGs can provide factual information and the ability to perform reasoning. This would boost the LLM's domain-specific reasoning, and interpretability, and mitigate biases and hallucinations. A notable challenge with KGs is their requirement for frequent updates, usually performed by processing and integrating information from vast textual datasets, LLMs can aid in generating and refining KGs. Therefore, combining LLMs and KGs offers a promising opportunity to advance both technologies and represents a pivotal challenge in the contemporary research landscape.

Areas of Expertise Relevant for the Action	**Keywords**
● Computer and Information Sciences: Artificial intelligence, intelligent systems, multi agent systems	● Large Language Models ● Knowledge Graphs ● Machine Learning ● Deep Learning

Specific Objectives
To achieve the main objective described in this MoU, the following specific objectives shall be accomplished:

Research Coordination
● Design a Comprehensive Research Agenda
● Bridge Academia and Industry
● Foster interdisciplinary collaboration to develop shared evaluation standards in the field
● Promote knowledge sharing accessibility through a collaborative and open-access infrastucture

Capacity Building
● Define and disseminate best practices for KGs and LLMs
● Strengthen the research community through inclusive, ongoing engagement initiatives
● Support capacity development of early-career researchers through structured mentoring

TECHNICAL ANNEX

1. S&T EXCELLENCE

1.1. SOUNDNESS OF THE CHALLENGE

1.1.1. DESCRIPTION OF THE STATE OF THE ART

This initiative will establish a platform for researchers and practitioners across various application domains to collaborate on the integration of Knowledge Graphs (KGs) and Large Language Models (LLMs). The aim is to further research, synchronize efforts, and spread awareness about the foundational technologies.

KGs are increasingly recognized for their ability to represent and systematize structured data semantically. They effectively encapsulate the complex relationships between entities and attributes, providing a machine-interpretable depiction of the domain, useful for numerous intelligent applications. Central to KGs is a domain ontology, a formal framework that delineates entity types and their interrelations, facilitating reasoning tasks. LLMs have become instrumental in advancing Natural Language Processing (NLP), exhibiting exemplary outcomes across various domains. They include both **masked language models** (e.g., BERT, RoBerta) as well as **generative language models** (e.g., LLaMa 2, GPT-4, Falcon). These models have exhibited proficiency in **few-shot and zero-shot learning** paradigms through prompting and in-context learning, positioning them as a leading paradigm in AI research.

The combination of KGs and LLMs offers one of the most compelling avenues for research in AI. On one hand, integrating KGs can equip LLMs with factual and verifiable data, leading to more interpretable, robust, and enhanced results. On the other hand, the expertise of LLMs in natural language understanding facilitates the extraction and organization of information from extensive textual data, which aids in the creation and refinement of KGs.

In recent years, there have been multiple efforts to **incorporate structured knowledge into deep learning models**, especially LLMs, for diverse applications. For example, Liu et al. 2019 proposed K-BERT, which extends BERT with a knowledge-enabled language representation model, where KGs triples are injected into the sentences during fine-tuning. Furthermore, Ostendorff et al. (2019) proposed a method for book genre classification by combining KG embeddings of authors with book titles and other metadata features to efficiently tackle classification tasks. Xu et al. (2023) introduced a novel approach incorporating entity-related knowledge into encoder-decoder large pre-trained language models through a generative knowledge infilling objective during continued pre-training. In contrast, Emelin et al. (2022) proposed injecting domain-specific knowledge before fine-tuning task-oriented dialogue tasks using lightweight adapters. Additionally, Moiseev et al. (2022) described a method to integrate structured knowledge into LLMs by directly training T5 models on factual triples from KGs. Finally, Wang et al. (2021) proposed a method that keeps the original parameters of the pre-trained model fixed and enables continuous knowledge infusion through a neural adapter for each type of infused knowledge, serving as a plug-in connected to the LLM. However, determining the most effective method to integrate KGs within LLMs for various tasks remains an open challenge in the academic community.

Moreover, a substantial body of literature explores the **creation and development of KGs**, across various domains and under different constraints. The overarching methodology often encompasses standard steps, such as 1) data identification, 2) KG ontology construction, 3) knowledge extraction, 4) analysis of extracted knowledge, 5) KG creation, and 6) maintenance (Tamašauskaitė and Groth, 2023). KGs typically originate from a combination of both structured and unstructured data, including text and images. Structured data can be assimilated using the RDF Mapping Language, such as RML (Dimou et al., 2014). Unstructured data usually require semi-automatic information extraction methods, yielding a set of key entities and statements, in the form *<s,p,o>*. This process is commonly achieved through a variety of approaches for named entity recognition and relationship extraction. For example, FRED (Gangemi et al., 2017) extracts frames, events, concepts, and entities, aligns them with various ontologies, and organizes the extracted text segments according to RDF. An interesting solution is based on building ensemble models that can capture a large range of entities and relationships in multiple domains (Martinez et al., 2018). Today, the most promising solution entails the adoption of large-scale language models based on the transformer architecture (Giorgi et al.,

2019), such as GPT- 4 (OpenAI, 2023), LaMDA (Thoppilan et al., 2022), and LLaMA 2 (Touvron et al., 2023). This is typically accomplished through either task-specific fine-tuning or the utilization of a few-shot learning approach (Mihindukulasooriya et al., 2023). Nonetheless, further research is paramount to attain the desired level of performance and usability, essential for reliably generating large-scale KGs.

Multilinguality is one of the major aspects in NLP where the problems related to high and low-resource languages are targeted. The field of **multilingual and cross-lingual KG generation** has evolved significantly over the years, driven by the growing need to connect and harness information from diverse linguistic and cultural contexts. Foundational projects such as DBpedia, FreeBase, and Wikidata laid the groundwork for subsequent research in multilingual and cross-lingual KGs. Multilingual entity linking, a crucial building block for cross-lingual KGs, has seen notable advancements. Works like BabelNet (Navigli and Ponzetto, 2010) have demonstrated effective approaches for aligning entities across languages and populating multilingual KGs with cross-references. Other works such as Message- Passing (Carbonell et al., 2021) or convolutional neural networks (Min et al., 2017) take advantage of advanced neural network architectures to tackle named entity recognition. Cross-lingual ontology alignment has gained attention, with studies proposing methods for harmonizing ontological structures across languages. However, in this case, deep approaches such as the one proposed by Bento et al. (2020) are still rarely used, while most methods are based on interlingual repositories such as BabelNet and Wikipedia. The development of multilingual word and entity embeddings has enhanced the ability to bridge language gaps in KGs. Tools like MUSE have enabled cross-lingual KG construction by providing a common vector space for entities across multiple languages. Finally, some EU-funded initiatives, such as the Lynx project (Rodriguez-Doncel et al., 2020), leverage KGs for specific domains. Challenges in multilingual and cross-lingual KG generation, including linguistic diversity, cultural nuances, and scalability, remain subjects of ongoing research. Emerging trends include the integration of NLP models to enhance entity linking and relation extraction in cross-lingual KGs.

As KG construction relies nowadays on NLP techniques and LLMs, which are on turn prone to include bias that exist in data, it is important to ensure that the construction process **addresses bias and other ethical concerns**. For instance, Fisher et al., 2020 highlighted that, as with word embeddings, harmful social biases related to professions are encoded in KG embeddings with respect to gender, religion, ethnicity and nationality. Moreover, KGs are often considered as “gold standard” data sources that safeguard the correctness of other systems. Because possible biases inherent to KGs may become magnified and spread through such systems, it is paramount to acknowledge and address various types of bias in KG construction. However, no work has addressed ethical KG construction until now, delegating the bias mitigation tasks to ethical LLMs and other deep methods.

1.1.2 DESCRIPTION OF THE CHALLENGE (MAIN AIM)

Based on the existing efforts and the problems that are identified so far in the literature regarding LLMs and the use of KGs, this Action will target the following key problems:

(1) **Hallucination and confabulations in LLMs.** Hallucinations and confabulations are one of the major challenges that LLMs face. While recent advancements in generative models have mitigated these issues to some extent, there is still no proper way to handle such problems for specific domains.
(2) **Lack of high-quality domain-specific applications.** Currently, LLMs are excelling at generic NLP tasks, however, their performance has not been widely measured on domain specific tasks. These domains include the field of biomedicine, scholarly data, finance, etc.
(3) **The development of KGs is complex and time-consuming.** Constructing and maintaining an extensive KG is a resource-intensive and complex task. LLMs offer the capability to mine information from vast textual datasets, thereby enhancing KGs with additional data.
(4) **Multilinguality.** LLMs need huge quantities of data to perform well. This constitutes a potential issue regarding low-resource languages. Some solutions have tried to address this problem by creating embeddings for language families instead of individual languages, but these approaches lag in accuracy compared to monolingual approaches.
(5) **Ethical Concerns and Bias.** Bias in LLMs refers to the presence of systematic and unfair prejudices in the outputs generated by these models. These biases can emerge from the training data, the design of the model, and the fine-tuning process.

Following these lines, this Action will focus on five main research questions:

- **RQ1:** Can the use of KG alleviate the hallucination and confabulation problems of LLMs?
- **RQ2:** Can LLMs be used in combination with KGs to develop efficient, reliable, and explainable domain-specific applications?
- **RQ3:** Are LLMs capable of augmenting KGs with comprehensive information, which can in turn provide support for addressing the concerns raised in RQ1 and RQ2?
- **RQ4:** How can the Action improve the coverage of multilingual KGs for low-resource languages to enhance cross-lingual knowledge accessibility for linguistically diverse communities?
- **RQ5:** How can the biases leading to the ethical concerns in LLMS and KGs (e.g., harmful stereotypes, misinformation, or discrimination) be targeted as a community?

1.2. PROGRESS BEYOND THE STATE OF THE ART

1.2.1. APPROACH TO THE CHALLENGE AND PROGRESS BEYOND THE STATE OF THE ART

Six distinct Working Groups (WGs) will address the primary challenges outlined previously:

- **WG1: Augmenting LLMs with KGs** - This WG will explore various methods to incorporate KGs into LLMs for producing more accurate and reliable LLMs.
- **WG2: Domain-Specific Applications Using LLMs and KGs** - It will produce resources (data and tools), to address domain-specific issues in areas like health, finance, and science.
- **WG3: KG Construction Aided by LLMs** - This WG will delve into the techniques for creating KGs, emphasizing LLM-based information extraction methods.
- **WG4: Multilingual Aspects of KGs and LLMs** - The focus here will be on understanding how KGs can effectively encapsulate and interlink information across languages.
- **WG5: Bias and Ethics in KGs** - It will advance research on methods that ensure fairness, transparency, and accountability in the development, upkeep, and utilization of KGs.
- **WG6: Evaluation and Validation Frameworks -** It will establish comprehensive evaluation and validation frameworks to systematically identify, quantify, and address the weaknesses and needs associated with LLMs augmented with KGs, supporting the efforts of all other working groups.

Building on this foundation, the Action anticipates the following advancements beyond state of the art:

Augmenting LLMs with KGs. Existing methods for augmenting LLMs with structured knowledge are typically developed for very specific use cases, employing domain-specific solutions and features that are not easily adaptable to other domains and tasks. Thus, the Action will investigate a range of general solutions able to integrate information from KGs. Integration will also be approached as post-hoc operation where KGs are used as external resource that the LLM can consult during inference. Key strategies under consideration include fine-tuning LLMs with KG data, granting LLMs direct access to KGs via plugins, and verifying LLM outputs against KGs. Benchmarks will be used for this process.

Domain-Specific Applications. In the current research landscape, many applications utilizing LLMs and KGs for significant domains (e.g., health, finance, scientific research) are still in their early stages. Therefore, the Action will pursue two primary objectives: 1) to develop a detailed set of domain-specific resources and 2) to improve LLM training methods for these areas. As such, WG2 will systematically assess potential tasks in these key domains and match them with appropriate datasets, KGs, and LLM techniques. Moreover, to encourage innovation, the Action will host competitions, inviting specialists to expand technological limits and foster cross-disciplinary cooperation.

KG Construction. Numerous fundamental challenges persist in the field of KG construction (as mentioned in (Groth et al., 2023)). The primary challenge, central to our Action, pertains to the role of LLMs. Although initial experiments in this domain exhibit significant potential, the Action still requires a comprehensive study exploring optimal methods for utilizing this technology to produce large-scale KGs. The Action aims to address this by tackling issues such as producing a unique representation of entities, integrating LLMs and ontological models, and minimizing hallucination and computational costs.

Multilinguality. The current state of KGs in multiple languages presents several challenges that the Action intends to tackle. Firstly, the Action will explore cross-lingual knowledge alignment by investigating methods that combine KGs from different languages using LLMs for multilingual connections. The project will mark a progress for LLMs for low-resource languages. The Action will

work on models such as NLLB (No Language Left Behind), improving their accuracy on these languages. Secondly, the Action will study techniques for multilingual entity linking to map entities from various linguistic backgrounds to a KG.

Bias and Ethics. The current discussion on ethics and bias in KGs is moving towards more systematic research methods emphasizing fairness, transparency, and accountability. First, the Action aims to explore methods for actively detecting and correcting unintentional biases in KGs, particularly those related to sensitive aspects such as ethnicity, gender, and race. Second, the Action aims to further research on transparent KG construction processes. This involves clearly outlining the methodology, from data sourcing to reasoning, enabling stakeholders to evaluate the KG's ethical foundations. Our goal is to identify biases, and promote ethical growth in knowledge systems.

Evaluation and Validation Frameworks. The Action will progress beyond the current state of the art by developing new benchmarks to improve LLM-KG integration in terms of accuracy, knowledge retention, and real-time inference. It will introduce gold standards for evaluating LLM outputs before and after KG augmentation across various domains. Novel benchmarks will also be created to reduce hallucinations and enhance consistency in entity representation with ontological models. The project will also advance multilinguality by improving models like NLLB and setting higher standards for LLM performance in multilingual and cross-lingual KG alignment. Additionally, the Action will develop advanced metrics to ensure transparency, ethical integrity, and bias correction in KG construction.

1.2.2. OBJECTIVES

1.2.2.1. Research Coordination Objectives

RO 1: Design a Comprehensive Research Agenda:

- Develop a research agenda detailing challenges and objectives.
- Publish the agenda within the first 6 months.
- Collaborate with experts to refine it.
- Address current gaps and emerging trends in KGs and LLMs.
- Ensure publication and dissemination within 6 months.

RO 2: Bridge Academia and Industry:

- Organize at least four workshops, two hackathons, and two datathons.
- Track participation and outcomes of each event.
- Engage industry and academic institutions to sponsor them.
- Focus on practical applications and challenges in KGs and LLMs.
- Schedule events at regular intervals.

RO 3: Foster interdisciplinary collaboration to develop shared evaluation standards in the field:

- 3 working groups of at least 5 experts from academia and industry.
- They produce >= 3 benchmark datasets and 2 sets of evaluation metrics.
- They meet bi-monthly.
- The benchmarks set standards for future research
- To be published within 12 months.

RO 4: Promote knowledge sharing accessibility through a collaborative and open-access infrastucture:

- Develop an open-access repository with tools and annotated datasets.
- Include 10 resources, 5 contributors, and 100 users in the first year.
- KGELL team has the expertise.
- Aligned with KGELL goals.
- Launch in month 2, grow from month 7

1.2.2.2. Capacity-building Objectives

CBO 1: Define and disseminate best practices for KGs and LLMs.

- Distribute scholarly publications and educational materials within 12 months.
- Develop educational resources and training sessions.
- Provide practical guidance and support for researchers and practitioners.
- Launch educational initiatives within 6 months and maintain them through the project.

CBO 2: Strengthen the research community through inclusive, ongoing engagement initiatives:

- Organize conferences, seminars, and webinars to foster community engagement.
- Host at least four major events and track participant engagement.
- Collaborate with organizations to promote these events.
- Include underrepresented genders and researchers from ITCs.
- Execute events throughout the project, ensuring ongoing community involvement.

CBO 3: Support capacity development of early-career researchers through structured mentoring:

- Pair senior experts with early-career researchers.
- 10 mentor-mentee pairs, 3 sessions per year.
- Utilize existing KGELL network.
- Builds capacity and supports researcher growth.
- Launch by month 4, review after 12 months.

2. NETWORKING EXCELLENCE

2.1. ADDED VALUE OF NETWORKING IN S&T EXCELLENCE

2.1.1. ADDED VALUE IN RELATION TO EXISTING EFFORTS AT EUROPEAN AND/OR INTERNATIONAL LEVEL

Several European initiatives focus on KGs and LLMs, though rarely in combination. Many projects address KGs indirectly, focusing on topics like Big Data, Semantic Web, Linked Data, or linguistics, often involving resources like dictionaries or thesauri. A search on the EU research portal CORDIS identified 32 KG-related projects active after October 2023. However, after filtering, only 7 projects were found to focus on KGs as a core topic, excluding those that merely use KGs for data storage without advancing research on their integration with LLMs.

- The **ENEXA** project funded under the Digital, Industry, and Space Horizon cluster works on "Efficient Explainable Learning on KGs", relevant to this COST Action. The main goals are to scale machine learning techniques to large graphs, exploit explicit semantics, and improve their explainability. The project does not focus on LLMs, but since one of the secondary proposers of the proposed Action is directly involved in the project a direct synergy is expected on the KG aspects.

- For the same reason, the **Graph-Massivize**r project funded under the Horizon Digital, Industry, and Space cluster is relevant to the current Action. For this project, the leaders of the work package responsible for scaling machine learning and approximate reasoning over KGs are secondary proposers for this COST Action.

- The **HumAIne** project, funded under the Horizon Digital, Industry and Space cluster works, among other things, on Neuro-Symbolic Learning capabilities, which "combines Deep Learning with semantics and rules to complete highly complex tasks with high accuracy while requiring considerably less training data than current AI models". Differently from them, the Action specifically focus on integrating LLMs and KGs to enhance solutions in both NLP and knowledge representation. The Action expects interaction with this project since one of our second proposers is a partner in the project.

- **GLACIATION** is a project focused on developing a "novel distributed KG that stretches across the edge-core-cloud architecture" and optimizing "the location where analytics are performed to significantly reduce power consumption". Its focus is quite orthogonal to this COST Action.

- **SciLake,** funded under the Research infrastructures cluster, is a clear support for our COST Action, specifically for the scholarly data applications. They aim to "empower the creation of scientific/scholarly KGs and the implementation of technologies to support data science and graph mining queries", which is one of the aspects clearly relevant for this COST Action.

- The **AIDAVA** project funded under the Horizon Health cluster is focused on a personal health KG. The goal is to "maximise automation of data curation and publish unstructured and structured, heterogeneous data using a virtual assistant powered by AI". Hence, while there is some relation, the goals are clearly complementary, and the Action aims to investigate whether collaboration is possible.

- The **TRIFECTA** project, funded by an ERC grant, aims to create a database that describes complex entities and concepts and their contexts by combining language and semantic web technology to extract and relate information from different texts over time. However, this project does not focus on machine learning techniques, which are central for our Action.

- The **SCENE** project funded under the Culture, creativity, and Inclusive Society Horizon cluster has a rather narrow focus on the life cycle of filmmaking.

Besides these Horizon projects, there are also two active COST Actions CA19134 on **Distributed KGs (DKG)** and CA23147 on **Global Network on Large-Scale, Cross-domain and Multilingual Open Knowledge Graphs** focused on KGs, their scale, decentralization, quality, and privacy. They therefore concentrate specifically on the management of KGs, rather than on combining them with LLMs to achieve a more comprehensive solution. Also the bias and ethics aspects are not part of those Actions.

An additional relevant initiative is the **LeMuR** (Learning with Multiple Representations) project. This Marie Skłodowska-Curie Actions Doctoral Network focuses on building "theoretical foundations and a first set of algorithms for the new Learning with Multiple Representations (LMR) paradigm". There is a strong synergy between LeMuR and KGELL, and the Action plans to collaborate to amplify each other's impact. KGELL will offer LeMuR access to our extensive network of researchers and relevant events, while LeMuR will contribute its expertise in LMR and related areas.

There are, of course, also older projects which the Action sees as an input to this activity and also broader activities on Big Data, Semantic Web, Analytics and Linked Data are potentially relevant to this proposed Action. A good example is the COST Action **Keystone** (2013-2017, IC1302). This COST Action was successful at bringing together the semantic web, information retrieval, big data, AI, machine learning and NLP research communities. There are also several international and European initiatives relevant for this Action:

- The Big Data Value Association (**BDVA)**. This organization facilitates requirements gathering and analysis for next generation data analytics and processing. Its scope is, however, very wide, and includes visualization, cloud platforms, semantics, etc. The main focus is on how to extract value from large volumes of data, which is also relevant for the proposed Action.
- The European Laboratory for Learning and Intelligent Systems (**ELLIS**) is a network of researchers actively researching machine learning in Europe. Among our secondary proposers are members of this network. The Action will actively advertise this Action within the network, and work together with researchers in the network to reach the goals of this Action.
- The **DL4KG** workshop. This workshop is set to be held for the 9th time in 2025, will be co-located with leading semantic web conferences and is expected to attract a significant audience. This workshop serves as a primary platform for showcasing the latest advancements in the field.
- The **BigScience** research workshop, which was co-organized by Huggingface and various French research organizations, with the purpose to develop a multilingual LLM. This model is intended for subsequent analysis by the multitude of researchers participating in the project.

There are several initiatives focused on KGs and LLMs, but none specifically address the foundation for combining these technologies, which is the goal of this Action. Some projects focus on LLM applications or NLP for tasks like sentiment analysis and linguistic studies. As interest in LLMs grows, more related projects are expected to emerge. The Action aims to bridge the gap between KGs and LLMs to tackle diverse tasks across various domains, while addressing challenges like bias, ethics, and multilingualism. Use cases will span finance, biomedicine, research, news, and more.

2.2. ADDED VALUE OF NETWORKING IN IMPACT

2.2.1. SECURING THE CRITICAL MASS, EXPERTISE, AND GEOGRAPHICAL BALANCE WITHIN THE COST MEMBERS AND BEYOND

The Action will secure a critical mass of expertise from academic scholars and industry practitioners by i) starting with a robust network of contributors from both the KG and NLP fields and ii) leveraging the existing relationships of our proposers with industry stakeholders to bolster support for the Action.

The network of proposers includes top researchers from 22 European countries, including 11 ITCs. This diverse group brings expertise in both foundational and applied research, as well as industrial applications, vital to the initiative's success. The team includes specialists in integrating KGs with LLMs, addressing multilingual challenges, and tackling bias and ethics in AI. They work in domains like finance, research, biomedicine, and news analysis, and maintain strong ties with businesses ranging from SMEs to multinational corporations that apply KG and NLP technologies. To further expand the

network and reach the critical mass of expertise in KG technologies, the Action will organize **tutorials, workshops, and hackathons**. Additionally, it will maintain a strong **online presence** through a dedicated website and active engagement on social media platforms. Participants will also be encouraged to share their involvement in the Action through their personal social media channels.

2.2.2. INVOLVEMENT OF STAKEHOLDERS

The network of proposers includes both **core researchers** in KG and LLM technologies and researchers who apply these technologies across a **wide variety of domains**, including news analysis, finance, education, healthcare, AI for scientific discovery, legal technology, sustainability, ethics, social sciences, biomedicine, and digital humanities. This diversity brings a strong **interdisciplinary dimension** to the Action. To support both core researchers and domain experts, the Action will organize a series of events (e.g., tutorials, hackathons, workshops, training schools) that will pursue two primary objectives: 1) fostering collaboration and knowledge exchange among KG and LLM researchers to enhance resource sharing and amplify their impact, and 2) providing targeted training and support to researchers in applied fields, equipping them to fully exploit the latest advancements in KG and LLM technologies. The network of proposers, though primarily composed of academics, has established strong connections with **industry stakeholders**, including several SMEs and academics with industry affiliations. These relationships have evolved through past collaborations, ongoing projects, and former students from the research groups now working in industry. Companies across a wide range of industries, including healthcare, finance, education, retail, manufacturing, telecommunications, entertainment, legal services, and public administration, are increasingly adopting LLMs and KGs to drive AI-powered solutions. For instance, in **healthcare**, companies are integrating biomedical knowledge graphs with LLMs to deliver transparent, reliable services that improve patient care and support clinical decision-making. In **finance**, these technologies enhance fraud detection, risk management, and customer service, while also enabling more accurate predictions of market trends by uncovering patterns in unstructured data. In the **legal sector**, LLMs and KGs are revolutionizing legal research and contract analysis, providing lawyers with advanced tools to navigate vast amounts of case law, statutes, and legal documents. In **retail**, AI-powered tools such as intelligent chatbots, recommendation engines, and personalized marketing strategies are transforming customer engagement and driving sales. Given these opportunities, a key focus will be on **fostering engagement with industry partners**, particularly through WG2. The Action will actively involve its industry network by organizing hackathons, datathons, industry-focused workshops, training programs, and short-term scientific missions. A major objective is to train and **support SMEs** that may lack in-house expertise in these emerging technologies. Ongoing support will be provided to industry partners in implementing AI solutions, with continuous feedback loops helping to refine and advance these technologies. During these initiatives, industry stakeholders will be encouraged to contribute use cases, share sample data, relay best practices, and even host events. These efforts aim to facilitate a mutual **exchange of knowledge between academia and industry**, promoting the adoption of state-of-the-art AI technologies across European businesses.

3. IMPACT

3.1 IMPACT TO SCIENCE, SOCIETY AND COMPETITIVENESS, AND POTENTIAL FOR INNOVATION/BREAKTHROUGHS

3.1.1 SCIENTIFIC, TECHNOLOGICAL, AND/OR SOCIOECONOMIC IMPACTS (INCLUDING POTENTIAL INNOVATIONS AND/OR BREAKTHROUGHS)

Scientific Impact. Our primary goal is to lead community efforts in developing innovative solutions for the integration of LLMs and KGs, which emerged as a paramount scientific challenge in AI. This integration is twofold: it involves embedding knowledge within LLMs (WG1) as well as utilizing LLMs to construct and enhance KGs (WG3). The integration of LLMs and KGs is crucial to address several critical limitations of LLMs, especially in domains where factual accuracy, consistency, and transparency are paramount. Key challenges faced by LLMs include: 1) hallucinations, where LLMs generate plausible but inaccurate or non-factual information; 2) the "long tail" problem, where LLMs struggle to deal with rare or specialized domains entities; 3) difficulties in handling numerical data, complex calculations, and logical reasoning; 4) a lack of transparency, as LLMs do not inherently provide sources or explanations for their responses; and 5) inability to stay up-to-date with real-time changes, as they are trained on static datasets. These challenges are particularly pronounced in high-

stakes fields such as biomedicine, law, finance, and scientific research, where outputs must be accurate, explainable, and verifiable. Integrating LLMs with KGs can mitigate these issues by enabling LLMs to rely on structured, up-to-date, and verifiable knowledge. The collaboration between computer scientists in the Action will also allow us to support the development of a multitude of novel applications spanning diverse domains and communities (WG2). Moreover, the datathons and hackathons shall ground the directions of research in practice. In addition to these pursuits, our commitment extends to addressing the development of LLMs and KGs for languages with limited resources. To this purpose, the Action will support the development and adoption of language-agnostic approaches, facilitating the mutual reinforcement of information across multiple languages (WG4). Finally, it is essential to recognize that both LLMs and KGs inherently manifest the biases present during their creation. For instance, LLMs trained on skewed data inevitably inherit the bias, similar to KGs constructed based on a prejudiced corpus or structured data. Beyond these circumstances, biases might also creep into KGs through schema design decisions made by their creators. In light of these challenges, the Action has established WG5, dedicated to rigorously examining issues of bias and ethical considerations throughout the duration of the Action. To support these comprehensive efforts, the Action will also focus on developing robust evaluation frameworks and validation protocols within WG6. Specifically, the Action will establish standardized benchmarks to assess the performance of LLM-KG integrations, define quantitative metrics for evaluating their effectiveness, and implement validation processes to test these integrations in real-world applications. Our activities will ensure that each working group's contributions are rigorously tested and aligned with best practices, ultimately leading to more reliable, effective, and ethically sound AI systems. This constellation of working groups will lead to several scientific publications with a significant impact on science and society at large.

Technological Impact. KGs and LLMs are two prominent technologies with the transformative capability to reshape our engagement with information. KGs provide a structured representation of the world, while LLMs can generate and understand text in a human-like way. By combining them the Action can create systems that can answer our questions, generate creative content, and help us to make better decisions. Furthermore, facts generated by the model can be traced back to the structured data sources they were derived from. Such traceability can be instrumental in detecting and addressing hallucination and inherent biases in the models or, at the very least, pinpointing their origins. The convergence of LLMs and KGs holds significant potential to enhance AI systems across various sectors, including healthcare, finance, education, retail, manufacturing, telecommunications, entertainment, legal services, and public administration. Specific systems that can benefit from this convergence include conversational agents, automated document generation and summarization, recommendation engines, compliance and risk management tools, systems for exploring and understanding news articles and media content, personalized learning platforms, virtual assistants, knowledge discovery systems, and advanced search engines. This powerful combination has the potential to dramatically reshape how the Action interacts with information and address complex problems.

Socioeconomic impact. KGs and LLMs have the potential to strengthen European competitiveness, improve transparency of government, and even raise the efficiency of journalists and other writers of factual information. By connecting software companies to the latest research in these technologies, and by improving the data management of companies that use this software, the Action targets European companies and aims to strengthen European competitiveness. The technologies developed in this Action can also help open government data for a wider audience and make legal and other formal documents accessible, understandable, and actionable for the less literate. Consequently, this Action enhances governmental transparency and bolsters data-driven decision-making processes. Such endeavors resonate with European directives that emphasize the importance of making governmental data more accessible to the citizenry. After evaluating the risks and benefits, the trustworthiness and accuracy of the combination of KGs and LLMs has the potential to benefit European society in a number of ways. By carefully managing the risks associated with these technologies, the Action can ensure that they are used to create a more informed and equitable future.

3.2 MEASURES TO MAXIMISE IMPACT

3.2.1 KNOWLEDGE CREATION, TRANSFER OF KNOWLEDGE AND CAREER DEVELOPMENT

The Action will create innovative courses and organize meetings, workshops, and hackathons. These

events will be focused on educating both consumers and prosumers about LLMs and KGs and how these techniques can be beneficial in their work. To support this educational initiative, the Action will develop course curricula for postgraduate students, record instructional materials as videos, and make them available online. Moreover, the network of proposers has strong connections with the individuals who run major conferences and workshops in this field. This facilitates the Action's ability to introduce the subject of LLMs and KGs at these influential events. Also, the Action intends to engage with industry professionals and other practitioners through standardization bodies, industry gatherings, and hackathons. These activities will serve as avenues for transferring knowledge to European businesses. Finally, the Action aims to advance the current state of knowledge through publication in prestigious scientific journals and conferences, as well as by participating in standardization efforts with organizations like **W3C**. The results of these endeavors will be documented in a comprehensive book providing an in-depth introduction to LLMs and KGs. Furthermore, the Action places a high priority on fostering the career development of early-career researchers by involving them in Action-related activities, including short-term scientific missions and leadership roles in the Working Groups and event organization.

3.2.2 PLAN FOR DISSEMINATION AND/OR EXPLOITATION AND DIALOGUE WITH THE GENERAL PUBLIC OR POLICY

The Action intends to disseminate its findings across scientific, industrial, and societal platforms to ensure broad-reaching impact. In the realm of **scientific dissemination**, the focus will be on sharing results at esteemed conferences, journals, and relevant workshops. In the context of the Semantic Web and KG domain, **conferences** of interest include the Web Conference, ISWC, ESWC, K-CAP, and LDK. For the LLM field, relevant conferences are ACL, EMNLP, LREC, COLING, Neurips, ICML, ECML/PKDD, etc. There are also venues where we already see works which combine these two areas like ICLR and LOG. There will also be a concerted effort to engage in conferences known for significant **industry engagement**, such as SEMANTiCS and the KG conference. The action will also support researchers presenting a tutorial at these conferences. Additionally, the Action will explore adjacent areas like the Internet of Things, Big Data, and multi-agent systems to showcase the wide-reaching significance of its work. As for **journal publications**, the intention is to target top-tier journals addressing both KGs and LLMs technologies such as like Neurocomputing, Artificial Intelligence Review, IEEE TNNLS, the Semantic Web Journal, Journal of Web Semantics, KBS, IPM, JMLR, ESWA, Nature Machine Intelligence, and TACL. The Action will also support a range of **workshops** of particular relevance, including Linked Data on the Web, Decentralising the Semantic Web, Linked Data in Linguistics, Workshop on Ontology Modularity, Context and Evolution, and industry forums like LT- Innovate, GALA, and TAUS. The Action also expects the involvement of our network in DL4KG and LM-KBC. Finally, the Action will organize **training schools** by leveraging its connections and potentially collaborating with institutions like the International Semantic Web Summer School (ISWS) and the ReasoningWeb Summer School.

Industrial dissemination will occur through various channels. Firstly, the Action will directly engage with industry via the W3C's Community and Industry Groups. These encompass groups like Ontology-Lexicon (OntoLex), Linked Data for Language Technologies (LD4LT), RDF Stream Processing, Web of Things, AI and Knowledge Representation, and JSON-LD community groups. The network also plans to extend its collaborations with industry, spanning from European multinationals to SMEs. Moreover, the Action will participate in a range of industry events, including those organized by the Big Data Value Association, the Semantic Technology Conference, and SemWeb.Pro.

Finally, the Action aims to **engage the public** in its research by adopting an engaged research methodology. Collaborative engagement with the community will drive the investigation of public interest, societal concerns, and challenges. The Action will work closely with community partners to ensure that its research benefits them. Open consultations on the Action and its methodologies, as well as open science events, will invite the general public to view and comment on ongoing research. The Action will leverage the widespread public familiarity with LLM-based applications, such as conversational agents, to design a workshop aimed at educating users on the limitations and potential pitfalls of these models. This workshop will feature scientifically sound demonstrations that highlight common issues like bias, misinformation, and contextual misunderstandings in LLM outputs. A strong focus will be placed on engaging open-source developers, making the results available on major open-source platforms like GitHub, making awesome lists of related papers and available code repositories, as well as engaging with the platforms supporting open source models such as Hugging Face. The Action will also connect with open-source practitioners through a series of technology-focused webinars.

4. IMPLEMENTATION

4.1 COHERENCE AND EFFECTIVENESS OF THE WORK PLAN

4.1.1 DESCRIPTION OF WORKING GROUPS, TASKS, AND ACTIVITIES

WG1: Augmenting LLMs with KGs

Objectives: LLMs are trained on text corpora and inherit biases from that data. The objective of this WG is to explore strategies for using KGs to improve LLMs, making them more accurate and trustworthy. This includes injecting knowledge from KGs into LLMs.

Tasks:

***Benchmarks and Measurements*:** There is much anecdotal evidence about factfulness, hallucinations, bias, and skew in LLMs, but little formal measurement. This work aims to create experimental setups and benchmarks to measure these issues, including developing new quality metrics for LLMs. While this task focuses on general metrics (WG1), WG2 will concentrate on evaluating LLM performance in specific downstream tasks.

Fine-Tuning LLMs with KG: One method to inject factual knowledge into LLMs is through fine-tuning. This can be done by creating a Q&A corpus using facts from KGs or by evaluating the accuracy of facts using both real and synthesized negative instances from KGs. The goal is to explore these approaches while also studying how the content of the fine-tuning data impacts the bias and skew in LLMs.

Using KGs as Plugins: Rather than modifying the core model, utilizing plugins can facilitate the execution of specific pieces of code within the language model. One use case would involve querying KGs for factual data via a Retrieval Augmented Generation (RAG) method. By integrating explicit knowledge in this manner, the Action anticipates enhanced factual accuracy in LLMs. Additionally, the Action intends to study the effects on other quality indicators, including bias and skew.

Post-processing outputs: The Action will investigate methods that utilize a KG to verify the accuracy of LLM outputs. If discrepancies are identified, potential responses include direct rectification or initiating an internal dialogue with the LLM, prompting it to amend its error.

WG2: Domain-specific tasks of LLMs based on KGs

Objectives: This WG focuses on analyzing various tasks that can be effectively tackled using LLMs integrated with KGs across different domains. It aims to create data and algorithmic resources to assist the community in leveraging LLMs based on KGs to address specific challenges in their fields.

Tasks:

***Task Exploration*:** The Action aims to systematically explore the manifold downstream tasks that can be effectively tackled within the aforementioned domains. By identifying these tasks, the WG aims to shed light on the transformative capabilities of this technology.

***Data Collection*:** To empower the LLMs with rich, domain-specific knowledge, this task will rigorously scrutinize and mine extensive datasets. These datasets will comprise a wealth of scientific papers, patents, news, finance and medical reports sourced from both publicly available repositories and past research projects (e.g. https://aida.kmi.open.ac.uk/). The meticulous curation of this data is paramount to ensuring the efficacy of subsequent training and application phases.

LLM Training: As a crucial pillar of the WG's endeavors, this task delves into the intricacies of training LLMs for optimal performance within the specified domains. It will involve the various training techniques analysed in WP1 to ensure that LLMs can excel in the specific contexts they are deployed. The Action will establish a roadmap for future research developments in this area, to help researchers and practitioners in navigating the complexities of domain-specific LLM+KG applications.

Challenges: This task aims to promote innovation and excellence by organizing challenges in specific domains, recognizing that competition and collaboration can drive progress. Researchers and industry experts will be invited to participate in these friendly competitions to develop novel solutions, expanding the potential of LLMs and KGs in industrial applications. The challenges will provide a platform to showcase advanced technologies and encourage interdisciplinary collaboration.

WG3: KG Construction assisted by LLMs

Objectives: The primary objective of this working group is to analyze and develop new methodologies for generating KGs. The focus will include semi-automated information extraction methods from various sources and best practices for creating ontological representations of specific domains. Emphasis will be placed on information extraction methodologies using LLMs, which are currently state-of-the-art for identifying entities and relationships in extensive text repositories across different domains and languages. Ultimately, the goal is to create an open library of methods and resources to improve the community's ability to construct new KGs and integrate them with existing ones.

Tasks:

***Ontology Engineering*:** Constructing a new KG necessitates the development of a formal ontological model. This task will concentrate on analysing and making available best practices for creating new ontologies and reusing existing schemas.

***Information Extraction*:** It will analyze recent LLM-based techniques for extracting information from various textual resources, including academic articles, news stories, technical documents, web pages, and social media. The focus will be on methods for named entity recognition and relationship extraction that produce statements in triple form (i.e., *<s,p,o>*), allowing for easy integration into a KG.

***Publishing*:** This task will focus on the challenges related to the publication, management, and querying of KGs using various techniques and tools. It will analyze different technologies for storing KGs, highlighting their advantages and costs, and explore approaches to support federated queries.

***Interlinking*:** This task will examine techniques and tools used for the interlinking of KGs both at the schema and data level. The objective is to streamline the process for those introducing new KGs, enabling them to link with a range of publicly accessible KGs, including both generic ones (e.g., Wikidata, DBpedia) and domain-specific ones. The long-term vision is to foster a robust network of intricately interlinked open KGs, which could bolster the evolution of knowledge-enhanced AI services.

***Enrichment and Refinement*:** This task will scrutinize both automated and semi-automated strategies utilized for the enrichment of KGs from diverse data reservoirs. Notably, the analysis will encompass modern techniques for link prediction, which have been empirically shown to enhance data quality and scope.

WG4: Multilinguality for KGs and LLMs

Objectives: This WG aims to improve the understanding of how KGs can effectively capture, connect, and utilize information in various languages, particularly low-resource languages, to facilitate cross-cultural and cross-linguistic knowledge sharing. By developing new methodologies, tools, and approaches, the goal is to overcome language barriers and enhance the accessibility and usability of KGs in diverse linguistic contexts, ultimately promoting global information exchange and collaboration.

Tasks:

***Cross-lingual knowledge Alignment/Language-agnostic representation of knowledge*:** Investigate methods for aligning and integrating KGs across multiple languages, ensuring that equivalent concepts and entities in different languages are represented and connected appropriately. LLMs can be used to derive multilingual embeddings for this task.

***Multilingual Entity Linking*:** Develop techniques for entity linking that work effectively in multilingual contexts, allowing for the identification and disambiguation of entities in various languages and mapping them to a common KG.

***Multilingual Triple Generation*:** Explore approaches for generating triples (*<s,p,o>*) from textual sources in different languages, focusing on the extraction of structured information that can populate a multilingual KG.

***Language-Aware Querying*:** Develop query languages and interfaces that allow users to get information from multilingual KGs while considering the linguistic context of their queries.

***Multilingual Entity Disambiguation*:** Work on improving entity disambiguation techniques in multilingual KGs to ensure that entities with similar names but different meanings are correctly disambiguated.

WG5: Bias and Ethics

Objectives: This WG will focus on advancing the research and development of methods that promote fairness, transparency, and accountability in the construction, maintenance and use of KGs.

Tasks:

***Bias Mitigation*:** it aims at mitigating any unfair biases that may inadvertently emerge during the creation and maintenance of KGs. The Action will explore techniques to detect and rectify biases related to ethnicity, gender, race, and other sensitive attributes, ensuring that our KGs remain neutral and unbiased representations of information.

***Interpretability and Explainability*:** It focuses on improving the interpretability and explainability of KG construction processes. The aim is to provide clear explanations of how KGs are built, including the sources of information, data selection criteria, and reasoning processes involved. This transparency fosters trust and allows for critical assessments of the KGs' integrity and ethical standards. By achieving these goals, the WG seeks to create KGs that uphold high ethical standards while empowering users and stakeholders to make informed decisions, identify potential biases, and contribute to the responsible development of interconnected knowledge systems.

***Ethical requirement planning*:** It aims at collecting clear, measurable definitions of concrete ethical requirements and defining ways to automate this process for all steps related

***Stakeholder Engagement*:** It fosters engagement with diverse stakeholders, including experts, communities, and user groups, to solicit feedback and perspectives on bias and ethical considerations.

***Ethics Training*:** It provides training and awareness programs for KG contributors to raise awareness of ethical concerns and ensure ethical data handling practices.

***Bias Assessment*:** It aims at developing methodologies to systematically assess and quantify potential biases in existing KGs, identifying areas where bias may be present.

***General-purpose bias mitigation*:** This will be done at the LLM level, to reduce the risk of importing into KGs biases that are implicit in the language models. For instance, the Action would like to avoid the acquisition into the KG of biased terminology (slurs, homophobic terms, etc.).

***Domain-specific (i.e., scientific) debiasing*:** Some bias can be related to the domain model by a KG. Therefore, the Action needs to avoid biased generalization or in any case relations that can lead to biased KGs. For instance, if the KG is modeling hospital and health roles, the Action needs to avoid the creation of triples such as *(<nurse, is a, woman>)*.

***Interpretable and explainable KGs*:** It is necessary to present the source of the information acquired during the construction of a KG in a way that users can inspect the reliability of the information and identify the source of possible biased information.

***Diversity Enhancement*:** It aims at developing strategies to enhance the diversity of data sources and contributors, reducing the risk of systemic bias in KGs.

WG6: Evaluation and Validation Frameworks

Objectives: WG6 will focus on creating standardized benchmarks and evaluation methodologies to systematically assess the performance of LLMs when integrated with KGs across various tasks and domains. WG6 will also develop quantitative metrics to measure the strengths and weaknesses of these integrations, ensuring detailed analysis and validation. Furthermore, WG6 will test the practical effectiveness, scalability, and ethical implications of LLM-KG integrations in real-world applications. Additionally, WG6 will provide tailored evaluation tools to support the work of other working groups (WG1 through WG5), ensuring their research aligns with best practices in AI and KG integration.

Tasks:

Benchmark Development: This task aims at designing specific benchmarks for evaluating the integration of KGs with LLMs. WG6 will work with WG1 on general integration benchmarks, with WG2 on domain-specific benchmarks for areas like healthcare and finance, and with WG3 on benchmarks assessing the quality of KG construction using LLMs. Additionally, WG6 will partner with WG4 to

create benchmarks for cross-lingual alignment and multilingual performance, and with WG5 to establish benchmarks for ethical considerations, focusing on fairness, transparency, and bias detection. These benchmarks will ensure both technical effectiveness and adherence to ethical standards in LLM-KG integrations.

Metric Creation: It aims at establishing standardized metrics to evaluate the performance of LLMs after integrating with KGs, focusing on accuracy, consistency, knowledge retention, and inference capabilities. It will further develop cost-benefit metrics to assess the computational efficiency and sustainability of KG construction methods, offering insights into resource use and viability. Also, this task will create metrics to evaluate the ethical performance of LLMs, including their ability to detect and mitigate biases, ensuring alignment with fairness, transparency, and accountability standards.

Validation in Real-World Scenarios: It aims at creating validation protocols to test LLM-KG integrations in real-world, high-stakes environments, ensuring they perform effectively outside controlled settings. These integrations will be validated in domain-specific areas like healthcare and finance, confirming that theoretical improvements result in tangible benefits. Finally, the ethical robustness of LLMs will be assessed by simulating scenarios where bias and ethical considerations are critical.

4.1.2 DESCRIPTION OF DELIVERABLES AND TIMEFRAME

The dissemination of the project will include best practices and surveys in the form of research articles. These articles are anticipated to constitute the following chapters in a comprehensive book focused on the fusion of KGs and LLMs. Next, the Action will show the milestones and the related activities.

Working Group 1:
- D1.1 - Chapter on metrics and benchmarks for LLM factfulness, bias, skew, and other metrics (M7)
- D1.2 - Chapter on fine-tuning LLMs with KGs (M10)
- D1.3 - Chapter on combining LLMs and KGs via plugin mechanisms (M20)
- D1.4 - Chapter on KG driven trust layers for LLMs (M36)

Working Group 2:
- D2.1 - Chapter on the main application fields of LLMs and KGs and relevant resources (M10)
- D2.2 - Chapter on the future of domain-specific LLM training and KG integration (M22)
- D2.3 - Chapter on innovation and collaboration through challenges (M48)

Working Group 3:
- D3.1 - Chapter on guidelines and best practices for KG Generation (M10)
- D3.2 - Chapter on the role of LLM for KG Generation (M20)
- D3.3 - Chapter about tools and resources for KG generation, refinement, and interlinking (M30)

Working Group 4:
- D4.1 - Chapter on multilingual KG generation (M20)
- D4.2 - Chapter on multilingual embeddings for the alignment of KGs (M36)

Working Group 5:
- D5.1 - Chapter on ethical KG generation guidelines (M12)
- D5.2 - Chapter on bias mitigation in KG generation (M24)
- D5.3 - Chapter on explainable methods for KG generation (M48)

Working Group 6:
- D6.1 - Chapter on Benchmark Development and Evaluation Frameworks (M48)
- D6.2 - Chapter on Metric Definition and Quantitative Analysis (M48)
- D6.3 - Chapter on Real-World Validation and Application Testing (M48)

Dissemination & Communication Deliverables (D7)

Del	Month	Activities
D7.1	6	Science Communication Plan. A comprehensive plan detailing how the Action will communicate its objectives and outcomes to different audiences.
D7.2	6	Action Website. Launch of the official website, logo, templates to share project progress, outputs, and events. It will be maintained and updated regularly.
D7.3	6	Social Media Setup and Strategy. Creation and launch of Action's social media channels (e.g., X, LinkedIn) with a content and engagement strategy.
D7.4	6	Open Repository. Online repository (via Zenodo, GitHub, or institutional servers) collecting the Action's public deliverables, datasets, software, and best practices.

Milestone	Month	Activities
MS1	6	Kick off meeting, WG set-up, web page published, Management Committee Meeting
MS2	12	First workshop, training schools (incl. hackathons), WG meetings
MS3	18	First Training school
MS4	24	Second workshop, training schools (incl. datathons), WG meetings
MS5	30	Guidelines and best practices
MS6	36	Third workshop, training schools (incl. hackathons), WG meetings
MS7	42	Research agenda, academic curriculum
MS8	48	Fourth workshop, training schools (incl. datathons), WG meetings, Action book

4.1.3 RISK ANALYSIS AND CONTINGENCY PLANS

Risk	Prob	Crit	Solution
Difficulty in obtaining high-quality, domain-specific datasets and KGs.	med	low	Forge partnerships with industry and academic institutions to gain access to proprietary and high-quality datasets.
Challenges in ensuring the accuracy and completeness of KGs, potentially leading to flawed or biased KGs.	med	med	Implement multiple layers of validation, including manual reviews and automated cross-referencing with existing KGs.
Remote locations of the members of the network to prohibit attendance to Action events.	med	low	Make increased use of teleconferencing technology.
Inadequate performance of LLMs in low-resource languages due to limited training data and lack of effective multilingual models.	med	med	Employ techniques like data augmentation and transfer learning to improve performance in low-resource languages.
Low commitment of participants or bad results.	low	high	Utilize SOTA project management by holding individual meetings to realign expectations, ensuring continuous communication with partners, and efficiently communicating specific benefits to each stakeholder.
New technological paradigms come up during the run time of the Action that would cause a fundamental shift in the direction.	low	high	Recent advances in technologies combining KGs and LLMs will be listed on the website which will constantly be updated. Hence, the problem can be identified and addressed early.
Failure to create effective and widely accepted evaluation metrics, leading to inconsistent or unreliable results.	med	high	Engage with the broader research community to review and refine evaluation metrics, ensuring they are broadly accepted and validated. Conduct pilot tests of evaluation tools and metrics with selected projects to refine and validate their effectiveness before full-scale deployment.
Incomplete identification and mitigation of biases in LLMs and KGs, leading to persistent ethical issues.	med	low	Assemble diverse teams to review and test for biases from multiple perspectives. Implement continuous monitoring and updating protocols to address emerging biases and ethical concerns as they arise.

4.1.4 GANTT DIAGRAM

Activities	Year 1				Year 2				Year 3				Year 4			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Milestones		MS1		MS2		MS3		MS4		MS5		MS6		MS7		MS8
Meetings																
Kickoff																
MC Meetings																
WG Meetings																
Events																
Workshops																
Datathons/Training Schools																
Industrial Scientific Events																
Working Groups																
WG1			D1.1	D1.2			D1.3					D1.4				
WG2				D2.1				D2.2								D2.3
WG3				D3.1			D3.2			D3.3						
WG4							D4.1					D4.2				
WG5				D5.1				D5.2								D5.3
WG6																D6.*
Dissemination & Communic		D7.*														

REFERENCES

(Groth et al. 2023) Groth, P. et al.Knowledge Graphs and their Role in the Knowledge Engineering of the 21st Century (Dagstuhl Seminar 22372)

(Tamašauskaitė and Groth, 2023) Tamašauskaitė, G. and Groth, P. Defining a knowledge graph development process through a systematic review. ACM Transactions on Software Engineering and Methodology, 32(1), pp.1-40 (2023)

(Dimou et al., 2014) Dimou, A., Vander Sande, M., Colpaert, P., Verborgh, R., Mannens, E. and Van de Walle, R. RML: A generic language for integrated RDF mappings of heterogeneous data. LDOW (2014)

(Gangemi et al., 2017) Gangemi A., Presutti V., Recupero D. R., Nuzzolese A. G., Draicchio F., Mongiovì M., Semantic Web Machine Reading with FRED. Semantic Web 8(6): 873-893 (2017)

(OpenAI, 2023) OpenAI. Gpt-4 technical report. arXiv:2303.08774 (2023)

(Thoppilan et al., 2022) Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.T., Jin, A., Bos, T., Baker, L., Du, Y. and Li, Y. Lamda: Language models for dialog applications. arXiv:2201.08239 (2022)

(Touvron et al., 2023) Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S. and Bikel, D. Llama 2: Open foundation and fine-tuned chat models. arXiv:2307.09288 (2023)

(Mihindukulasooriya et al., 2023) Mihindukulasooriya, N., Tiwari, S., Enguix, C.F. and Lata, K. Text2KGBench: A Benchmark for Ontology-Driven Knowledge Graph Generation from Text. arXiv:2308.02357 (2023)

(Giorgi et al., 2019) Giorgi, J., Wang, X., Sahar, N., Shin, W.Y., Bader, G.D. and Wang, B. End-to-end named entity recognition and relation extraction using pre-trained language models. arXiv:1912.13415 (2019)

(Martinez et al., 2018) Martinez-Rodriguez, J.L. et al. OpenIE-based approach for knowledge graph construction from text. Expert Systems with Applications 113, pp. 339–355 (2018).

(Liu et al. 2019) Liu, W. et al. K-BERT: enabling language representation with knowledge graph. arXiv:1909.07606 (2019)

(Ostendorff et al. 2019) Ostendorff, M., Bourgonje, P., Berger, M., Schneider, J. M., Rehm, G., Gipp, B. Enriching BERT with knowledge graph embeddings for document classification. arXiv:1909.08402 (2019)

(Xu et al. 2023) Xu, Y., Namazifar, M., Hazarika, D., Padmakumar, A., Liu, Y., HakkaniTur, D. Kilm: Knowledge injection into encoder-decoder language models. arXiv:2302.09170 (2023)

(Emelin et al. 2022) Emelin, D., Bonadiman, D., Alqahtani, S., Zhang, Y., Mansour, S. Injecting domain knowledge in language models for task-oriented dialogue systems. arXiv:2212.08120 (2022)

(Moiseev et al. 2022) Moiseev, F., Dong, Z., Alfonseca, E., Jaggi, M. SKILL: Structured knowledge infusion for large language models. In: NAACL-HLT, 2022.

(Wang et al. 2021) Wang, R., Tang, D., Duan, N., zhongyu wei, Huang, X., Ji, J., Cao, G., Jiang, D., Zhou, M. K-adapter: Infusing knowledge into pre-trained models with adapters. arXiv:2002.01808 (2021)

(Navigli and Ponzetto, 2010) Navigli, R., Ponzetto, S. P. BabelNet: Building a very large multilingual semantic network. In: ACL (2010)

(Carbonell et al., 2021) Carbonell, M., Riba, P., Villegas, M., Fornés, A., Lladós, J. Named entity recognition and relation extraction with graph neural networks in semi-structured documents. In: ICPR (2021).

(Min et al., 2017) Min, B., Jiang, Z., Freedman, M., Weischedel, R. Learning transferable representation for bilingual relation extraction via convolutional neural networks. In: IJCNLP (2017).

(Bento et al. 2020) Bento, A., Zouaq, A., Gagnon, M. Ontology matching using convolutional neural networks. In: LREC (2020).

(Rodriguez-Doncel et al., 2020) Rodríguez-Doncel, V., Montiel-Ponsoda, E. Lynx: Towards a legal knowledge graph for multilingual Europe. Law Context: A Socio-Legal J., 37, 175 (2020).

(Bollacker et al., 2008), Kurt Bollacker, Colins Evans, Praveen Paritosh, Tim Sturge, Jamie Taylor. Freebase: a collaboratively created graph database for structuring human knowledge. SIGMOD '08: Proceedings of the 2008 ACM SIGMOD international conference on Management of data, Pages 1247 – 1250, https://doi.org/10.1145/1376616.1376746

(Fisher et al., 2020), https://arxiv.org/pdf/1912.02761.pdf

DBPedia, https://www.dbpedia.org/

Wikidata, https://www.wikidata.org/wiki/Wikidata:Main_Page

MUSE, https://ai.meta.com/tools/muse/

Neurocomputing, https://www.sciencedirect.com/journal/neurocomputing

Artificial Intelligence Review, https://link.springer.com/journal/10462

IEEE TNNLS, https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5962385

The Semantic Web Journal, https://www.semantic-web-journal.net/

Journal of Semantic Web, https://www.sciencedirect.com/journal/journal-of-web-semantics

KBS, https://www.sciencedirect.com/journal/knowledge-based-systems

IPM, https://www.sciencedirect.com/journal/information-processing-and-management

JMLR, https://www.jmlr.org/

ESWA, https://www.sciencedirect.com/journal/expert-systems-with-applications

Nature Machine Intelligence, https://www.nature.com/natmachintell/

TACL, https://transacl.org/index.php/tacl

DL4KG, https://genetasefa.github.io/dl4kg2024/

LM-KBC, https://lm-kbc.github.io/