Temporal Information Retrieval and Extraction: From Foundations to RAG

Bhawna Piryani
bhawna.piryani@uibk.ac.at
University of Innsbruck
Innsbruck, Tyrol, Austria

Avishek Anand
avishek.anand@tudelft.nl
Delft University of Technology
Delft, Netherlands

Omar Alonso
omralon@amazon.com
Amazon
Santa Clara, California, United States

Adam Jatowt
adam.jatowt@uibk.ac.at
University of Innsbruck
Innsbruck, Tyrol, Austria

Abstract

Information continuously evolves over time. Because of this dynamic nature, time becomes a fundamental dimension that shapes how we extract, retrieve, interpret, and reason about knowledge. As information systems are constantly updated, models must determine not only what is relevant, but also when that information is valid. This tutorial provides a structured and in-depth overview of the complete temporal information access pipeline: Temporal Information Extraction (TIE), Temporal Information Retrieval (TIR), and Temporal Question Answering (TQA). We examine the progression of temporal methods from early rule-based extraction and probabilistic retrieval to contemporary transformer-based and large language model (LLM) architectures. Participants gain a solid understanding of the core principles underlying the identification and normalization of time expressions, time-aware document ranking, and temporal reasoning in retrieval-augmented generation (RAG). The tutorial concludes with a discussion of open challenges and future research directions aimed at building AI systems that are temporally aware, robust, and adaptive. By connecting classical extraction and IR foundations with modern LLM-based reasoning, this tutorial presents a cohesive and up-to-date perspective on temporal information systems.

CCS Concepts

• **Information systems → Specialized information retrieval**; **Information retrieval**.

Keywords

Temporal Information Extraction, Temporal Information Retrieval, Temporal Question Answering, Large Language Models, Retrieval Augmented Generation

ACM Reference Format:
Bhawna Piryani, Avishek Anand, Omar Alonso, and Adam Jatowt. 2026. Temporal Information Retrieval and Extraction: From Foundations to RAG. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26), July 20–24, 2026, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3805712.3808638

ACM ISBN 979-8-4007-2599-9/2026/07
https://doi.org/10.1145/3805712.3808638

1 Motivation

Time fundamentally shapes how knowledge is created, updated, and interpreted. Facts change, roles evolve, policies are amended, scientific findings are revised, and real-world events unfold continuously. Yet most information access systems — including many LLMs — are built on largely static assumptions about knowledge. This mismatch between dynamic reality and static modeling leads to outdated answers, temporal inconsistencies, and reduced trust in AI systems.

When a user searches for "best practices for React development" or asks "What is the EU's stance on encryption?", the answer critically depends on *when* the question is asked. Recommendations shift across software versions, policy positions evolve, and factual knowledge decays over time. Temporal awareness is therefore not an auxiliary feature of retrieval systems — it is a core dimension of relevance, contextual appropriateness, and reliability.

The complete temporal information access pipeline addresses this fundamental challenge by explicitly modeling how time influences unstructured text. This begins with Temporal Information Extraction (TIE), which focuses on identifying, normalizing, and grounding time expressions and events within documents. These extracted signals then empower Temporal Information Retrieval (TIR) to perform time-aware ranking and document dating, which ultimately feeds into Temporal Question Answering (TQA) for reasoning over time-dependent facts [4, 9, 19]. Together, these tasks extend classical NLP and IR notions of relevance by incorporating temporal validity and evolution.

The rapid adoption of LLMs and Retrieval-Augmented Generation (RAG) systems has further amplified the importance of temporal modeling. While LLMs demonstrate strong pattern recognition and reasoning capabilities, they often exhibit temporal blind spots, generate outdated information, or fail under temporal perturbations [23]. Without mechanisms for handling knowledge drift and aligning answers with appropriate timeframes, generative systems risk producing plausible but temporally incorrect outputs.

Beyond modeling, temporal information access depends critically on the structure and dynamics of underlying data collections. Real-world corpora evolve: documents are revised or removed, knowledge bases are updated, and versioned repositories preserve historical states of information. Supporting retrieval over such evolving collections requires dedicated infrastructure, including temporal index sharding, time-travel search, and efficient index maintenance strategies [5, 6]. Moreover, temporal signals in practice are often noisy or ambiguous — publication time, event time, modification time, and reference time may diverge — making extraction, normalization, and evaluation particularly challenging. Benchmark stability, temporal leakage, and snapshot bias further complicate reliable system assessment.

Although temporal IR has been studied for over a decade, early work focused primarily on timestamp-based ranking and rule-driven modeling [5, 14]. In contrast, recent advances introduce temporal language models, timestamp conditioning, continual adaptation, and temporal dense retrieval integrated within RAG architectures [1, 10, 11, 21, 26, 27]. Despite significant progress, these developments remain fragmented across subfields. Currently, no tutorial provides a unified perspective bridging classical temporal IR with modern neural paradigms.

The most recent related SIGIR tutorial [18] provided a comprehensive overview of pre-neural temporal IR methods. However, the field has since transformed due to the rise of transformer architectures, dense retrieval, temporal language modeling, and LLM-based systems. Building on our recent survey [19], which reviews over 150 papers, this tutorial synthesizes foundational and modern advances into a cohesive framework.

By bridging classical IR principles with contemporary neural, generative, and infrastructure-aware approaches, this tutorial provides a comprehensive end-to-end view of temporal information access — from data management and indexing to ranking, reasoning, and evaluation. As AI systems increasingly support decision-making in medicine, law, finance, governance, and everyday life, temporal reliability is no longer optional; it is essential for building trustworthy, adaptive, and context-aware intelligent systems.

2 Objectives

By the end of this tutorial, attendees will:

(1) Understand the foundations of TIE, including the identification of temporal expressions and events in unstructured text, and their normalization into standardized formats (e.g., TIMEX3).
(2) Understand the fundamental principles of TIR and TQA, including temporal intent detection, document dating, focus time estimation, and recency-aware ranking.
(3) Compare classical time-aware retrieval methods (e.g., temporal indexing and timestamp-based ranking) with modern neural and transformer-based approaches, and understand how temporal signals are incorporated into dense retrievers, temporal language models, and RAG pipelines.
(4) Analyze the temporal limitations of LLMs, including outdated knowledge, temporal hallucination, and sensitivity to distribution shifts, and explore paradigms for temporal reasoning such as event ordering, duration inference, multi-hop reasoning, and continual adaptation.
(5) Understand practical challenges in evolving collections, temporal indexing infrastructure, versioned corpora, and evaluation under dynamic knowledge settings.
(6) Identify open research directions in temporal robustness, uncertainty modeling, dynamic knowledge updating, and temporally-aware AI systems.

3 Relevance to SIGIR

The topic directly aligns with SIGIR's core themes of information retrieval, ranking, and evaluation. TIE, TIR, and TQA address challenges such as extracting temporal metadata, time-aware ranking, document dating, recency modeling, and retrieval over evolving collections. As search systems increasingly operate over dynamic and continuously updated corpora, the ability to accurately extract time signals and use them for temporal awareness becomes a critical dimension of relevance estimation and system reliability.

The recent integration of LLMs into retrieval pipelines further amplifies these challenges. While LLM-based systems are now widely studied within the SIGIR community, their temporal limitations — including outdated knowledge, temporal inconsistencies, and sensitivity to distribution shifts — remain insufficiently addressed. This tutorial connects long-standing research in temporal IR with contemporary advances in neural retrieval, dense indexing, and RAG, highlighting how temporal reasoning and recency awareness can enhance the robustness and trustworthiness of modern retrieval systems.

The tutorial also builds upon SIGIR's prior engagement with temporal IR, including the 2016 SIGIR tutorial on Temporal Information Retrieval [18]. Since then, the field has evolved substantially with the rise of transformer architectures, temporal language models, and dynamic evaluation frameworks. This tutorial updates and extends these foundations in light of recent developments.

A related version of this tutorial will be presented at The Web Conference 2026, where the emphasis is placed on web-scale dynamics and evolving content ecosystems. In contrast, this SIGIR 2026 edition is specifically tailored to the IR community by presenting a complete, end-to-end pipeline that explicitly introduces TIE alongside Retrieval and Answering. Furthermore, it places a stronger emphasis on core IR research settings, including ranking models, temporal indexing strategies, evaluation methodologies, and retrieval-augmented generation. The material has been substantially expanded and adapted to reflect these methodological additions, aligning closely with the priorities of the SIGIR audience.

Overall, the tutorial offers a timely and IR-centered contribution to the SIGIR 2026 audience by integrating foundational temporal IR research with modern neural and generative retrieval paradigms.

4 Format and Audience

Format: This tutorial is to be conducted in a lecture-style format as a half-day, on-site session with scheduled breaks.
Intended Audience: This tutorial is aimed at researchers, students, and practitioners in NLP and IR who are interested in temporally aware LLM and RAG systems. It is suitable for introductory to intermediate audiences in both academia and industry.
Prerequisite Knowledge: Attendees are expected to have a basic familiarity with language models and information retrieval. The tutorial is self-contained and introduces key concepts in temporal IR, time-aware RAG, temporal reasoning in LLMs, and the evaluation of temporal reliability.

Section & Duration	**Topics and Learning Objectives**
1. Introduction and Motivation (15 min)	*Goal:* Establish why time is a fundamental dimension in information access. • Why time matters in NLP and IR: recency, evolution, factual decay. • Key challenges: knowledge volatility, temporal ambiguity, context shift. • Core definitions: introducing the Extraction → Retrieval → Answering pipeline.
2. The Temporal Pipeline: Extraction, Retrieval, & QA (25 min)	*Goal:* Introduce the core canonical tasks, formats, and benchmarks across the pipeline. • **Extraction (TIE):** Recognizing time expressions/events and normalization (e.g., TIMEX3). • **Retrieval (TIR):** Temporal intent detection, document dating, focus time estimation. • **Answering (TQA):** Temporal granularity and query understanding. • Benchmark overview: design, extraction metrics, and evaluation paradigms.
3. Foundations: Pre-LLM Extraction & Retrieval (20 min)	*Goal:* Review classical methods prior to the neural era. • Rule-based and statistical temporal extraction (e.g., HeidelTime, SUTime). • Probabilistic and temporal language modeling approaches for retrieval. • Temporal indexing, metadata annotation, and early ranking strategies.
4. Neural and Transformer-based Temporal Models (20 min)	*Goal:* Explain the evolution of temporal representation learning. • Transformer-based extraction (NER for time/events) and temporal language models. • Temporal adaptation, continual learning, and timestamp conditioning. • How neural encoders capture or forget temporal signals.
Quick Q&A (10 min)	Clarify foundational concepts (TIE and TIR) and prepare participants for the RAG and reasoning sections.
Coffee Break (15 min)	*Informal networking and discussion.*
5. Temporal RAG and Reasoning (30 min)	*Goal:* Explore answering and reasoning in LLM-based temporal systems. • Integrating extracted temporal metadata within LLM pipelines (Temporal RAG). • Temporal dense retrievers and alignment. • Temporal reasoning: event ordering, duration inference, multi-hop reasoning. • Benchmarks for evaluation and temporal reasoning limits of LLMs.
6. Data, Infrastructure, and Evolving Collections (20 min)	*Goal:* Address real-world challenges in applying the pipeline to dynamic collections. • Applying extraction at scale: metadata reliability and noisy temporal signals. • Evolving corpora: updates, revisions, and versioned collections. • Temporal index sharding, archive search, and time-travel retrieval. • Evaluation under dynamic and temporally shifting datasets.
7. Emerging Topics and Open Challenges (10 min)	*Goal:* Discuss open challenges and future directions. • Temporal uncertainty, diachronic–synchronic integration. • Continual updates, temporally aware LLM agents, dynamic evaluation.
8. Concluding Discussion and Q&A (15 min)	*Goal:* Summarize insights and engage participants. • Key takeaways unifying Extraction, Retrieval, and Answering. • Interactive discussion on open research problems.

Table 1: Tutorial structure, and learning objectives (total duration: 3 hours, including a 10-minute Q&A and 15-minute coffee break).

5 Tentative Outline & Schedule

The tutorial spans 3 hours and comprises 8 main sections. We present a detailed breakdown of the tutorial's structure in Table 1.

6 Tutorial Material

The materials and logistics are summarized below:

- **Duration:** The tutorial is planned as a 3-hour session.
- **Tutorial Website:** Our tutorial website is available at: https://datascienceuibk.github.io/temporal-ir-extraction-tutorial-sigir2026/
- **Slides:** All presentation slides will be made publicly available on the tutorial website after the event.
- **Survey:** Our preprint survey paper is available at: https://arxiv.org/abs/2505.20243
- **Related Papers:** We have compiled a repository of related research papers and resources, accessible at: https://github.com/DataScienceUIBK/TemporalQA-Survey
- **Organization details:** The tutorial will be delivered in an on-site format, with all presenters physically attending and presenting.

7 Presenter Qualifications

The tutorial is delivered by a team of three senior and one junior researcher with extensive expertise in temporal information extraction, retrieval, question answering, temporal reasoning, and large language models [1, 5, 6, 12–14, 17, 20, 24–26]. Collectively, the instructors maintain a strong record of publications and community service in these areas, covering both foundational and applied aspects of time-aware information access.

Bhawna Piryani focuses on temporal IR and QA in her PhD study, contributing to recent work on temporally robust retrievers and

LLM-based temporal reasoning. She is the lead author of the recent survey on Temporal Question Answering [19].

Avishek Anand has organized and delivered tutorials at top IR venues, including *Explainable Information Retrieval* [7, 8] (SIGIR 2023, FIRE 2023) *Question Answering on the Curated Web* [22] (SIGIR 2020, ICTIR 2021) and *Temporal Information Retrieval* [18] (SIGIR 2016).

Omar Alonso has extensive experience in temporal information retrieval, recency-aware ranking, and search evaluation under dynamic conditions, with research examining how temporal signals influence ranking models, freshness modeling, query intent, and large-scale search architectures [2–4]. Bridging foundational research and real-world deployment, he brings both academic and industry perspectives on designing, evaluating, and deploying temporally-aware retrieval systems at scale under evolving data conditions.

Adam Jatowt has been an active contributor to the **Temporalia Task Series** [15, 16] and has multiple publications on temporal ranking, intent detection, development of various temporal tasks and evaluation of time-aware systems. He co-authored comprehensive surveys on temporal IR [9, 19].

Together, the team brings complementary expertise in temporal extraction, retrieval, QA, explainability, and LLM-based reasoning, ensuring a comprehensive and high-quality introduction to **Temporal Information Extraction, Retrieval, and Question Answering**.

References

[1] Abdelrahman Abdallah, Bhawna Piryani, Jonas Wallat, Avishek Anand, and Adam Jatowt. 2026. TempRetriever: Fusion-based Temporal Dense Passage Retrieval for Time-Sensitive Questions. In *Proceedings of the Nineteenth ACM International Conference on Web Search and Data Mining* (USA) *(WSDM '26)*. Association for Computing Machinery, New York, NY, USA, 5–15. doi:10.1145/3773966.3777938

[2] Omar Alonso, Michael Gertz, and Ricardo Baeza-Yates. 2007. On the value of temporal information in information retrieval. *SIGIR Forum* 41, 2 (Dec. 2007), 35–41. doi:10.1145/1328964.1328968

[3] Omar Alonso, Marc Spaniol, and Ricardo Baeza-Yates. 2025. Introduction to the Special Issue on Temporal Web: Studying Time and the Temporal Dimension. *ACM Trans. Web* 19, 3, Article 25 (Aug. 2025), 3 pages. doi:10.1145/3743140

[4] Omar Alonso, Jannik Strötgen, Ricardo Baeza-Yates, and Michael Gertz. 2011. Temporal Information Retrieval: Challenges and Opportunities. In *Temporal Web Analytics Workshop TWAW 2011*. 1.

[5] Avishek Anand, Srikanta Bedathur, Klaus Berberich, and Ralf Schenkel. 2011. Temporal index sharding for space-time efficiency in archive search. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. 545–554.

[6] Avishek Anand, Srikanta Bedathur, Klaus Berberich, and Ralf Schenkel. 2012. Index maintenance for time-travel text search. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in Information Retrieval*. 235–244.

[7] Avishek Anand, Sourav Saha, Procheta Sen, and Mandar Mitra. 2023. Explainability of Text Processing and Retrieval Methods. In *Proceedings of the 15th Annual Meeting of the Forum for Information Retrieval Evaluation, FIRE 2023, Panjim, India, December 15-18, 2023*. ACM, 153–157. doi:10.1145/3632754.3632944

[8] Avishek Anand, Procheta Sen, Sourav Saha, Manisha Verma, and Mandar Mitra. 2023. Explainable information retrieval. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 3448–3451.

[9] Ricardo Campos, Gaël Dias, Alípio M. Jorge, and Adam Jatowt. 2014. Survey of Temporal Information Retrieval and Related Applications. *ACM Comput. Surv.* 47, 2, Article 15 (Aug. 2014), 41 pages. doi:10.1145/2619088

[10] Bhuwan Dhingra, Jeremy R. Cole, Julian Martin Eisenschlos, Daniel Gillick, Jacob Eisenstein, and William W. Cohen. 2022. Time-Aware Language Models as Temporal Knowledge Bases. *Transactions of the Association for Computational Linguistics* 10 (2022), 257–273. doi:10.1162/tacl_a_00459

[11] Anoushka Gade and Jorjeta G Jetcheva. 2024. It's About Time: Incorporating Temporality in Retrieval Augmented Language Models. *CoRR* (2024).

[12] Raghav Jain, Daivik Sojitra, Arkadeep Acharya, Sriparna Saha, Adam Jatowt, and Sandipan Dandapat. 2023. Do language models have a common sense regarding time? revisiting temporal commonsense reasoning in the era of large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. 6750–6774. doi:10.18653/v1/2023.emnlp-main.418

[13] Adam Jatowt, Ching Man Au Yeung, and Katsumi Tanaka. 2015. Generic method for detecting focus time of documents. *Information Processing & Management* 51, 6 (2015), 851–868. doi:10.1016/j.ipm.2015.05.001

[14] Adam Jatowt, Yukiko Kawai, and Katsumi Tanaka. 2005. Temporal ranking of search engine results. In *Web Information Systems Engineering–WISE 2005: 6th International Conference on Web Information Systems Engineering, New York, NY, USA, November 20-22, 2005. Proceedings 6*. Springer, 43–52. https://link.springer.com/chapter/10.1007/11581062_4#citeas

[15] Hideo Joho, Adam Jatowt, Roi Blanco, Hajime Naka, and Shuhei Yamamoto. 2014. Overview of NTCIR-11 Temporal Information Access (Temporalia) Task. In *Proceedings of the 11th NTCIR Conference on Evaluation of Information Access Technologies, NTCIR-11, National Center of Sciences, Tokyo, Japan, December 9-12, 2014*. National Institute of Informatics (NII).

[16] Hideo Joho, Adam Jatowt, Roi Blanco, Haitao Yu, and Shuhei Yamamoto. 2016. Overview of NTCIR-12 Temporal Information Access (Temporalia-2) Task. In *Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies, National Center of Sciences, Tokyo, Japan, June 7-10, 2016*. National Institute of Informatics (NII).

[17] Hideo Joho, Adam Jatowt, and Blanco Roi. 2013. A survey of temporal web search experience. In *Proceedings of the 22nd International Conference on World Wide Web (WWW '13 Companion)*. 1101–1108. doi:10.1145/2487788.2488126

[18] Nattiya Kanhabua and Avishek Anand. 2016. Temporal Information Retrieval. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 17-21, 2016*. ACM, 1235–1238. doi:10.1145/2911451.2914805

[19] Bhawna Piryani, Abdelrahman Abdallah, Jamshid Mozafari, Avishek Anand, and Adam Jatowt. 2025. It's High Time: A Survey of Temporal Question Answering. *To appear at ACL 2026; available as arXiv preprint arXiv:2505.20243* (2025).

[20] Bhawna Piryani, Abdelrahman Abdallah, Jamshid Mozafari, and Adam Jatowt. 2024. Detecting Temporal Ambiguity in Questions. In *Findings of the Association for Computational Linguistics: EMNLP 2024*. Association for Computational Linguistics, 9620–9634. doi:10.18653/v1/2024.findings-emnlp.562

[21] Guy D Rosin, Ido Guy, and Kira Radinsky. 2022. Time masking for temporal language models. In *Proceedings of the fifteenth ACM international conference on Web search and data mining*. 833–841. doi:10.1145/3488560.3498529

[22] Rishiraj Saha Roy and Avishek Anand. 2020. Question Answering over Curated and Open Web Sources. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020*. 2432–2435. doi:10.1145/3397271.3401421

[23] Tu Vu, Mohit Iyyer, Xuezhi Wang, Noah Constant, Jerry Wei, Jason Wei, Chris Tar, Yun-Hsuan Sung, Denny Zhou, Quoc Le, et al. 2024. FreshLLMs: Refreshing Large Language Models with Search Engine Augmentation. In *Findings of the Association for Computational Linguistics ACL 2024*. 13697–13720.

[24] Jonas Wallat, Abdelrahman Abdallah, Adam Jatowt, and Avishek Anand. 2025. A Study into Investigating Temporal Robustness of LLMs. In *Findings of the Association for Computational Linguistics: ACL 2025*. Association for Computational Linguistics, Vienna, Austria, 15685–15705. doi:10.18653/v1/2025.findings-acl.810

[25] Jonas Wallat, Adam Jatowt, and Avishek Anand. 2024. Temporal Blind Spots in Large Language Models. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining* (Merida, Mexico) *(WSDM '24)*. 683–692. doi:10.1145/3616855.3635818

[26] Jiexin Wang, Adam Jatowt, Masatoshi Yoshikawa, and Yi Cai. 2023. BiTimeBERT: Extending Pre-Trained Language Representations with Bi-Temporal Information. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Taipei, Taiwan) *(SIGIR '23)*. 812–821. doi:10.1145/3539618.3591686

[27] Feifan Wu, Lingyuan Liu, Wentao He, Ziqi Liu, Zhiqiang Zhang, Haofen Wang, and Meng Wang. 2024. Time-Sensitve Retrieval-Augmented Generation for Question Answering. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management* (Boise, ID, USA) *(CIKM '24)*. 2544–2553. doi:10.1145/3627673.3679800