SECTION 2 · 20 min

Core Concepts & Temporal IR Tasks

Canonical temporal tasks, the signals they rely on, and the benchmarks used to evaluate them.

Presenter: Bhawna Piryani

02

Temporal Dimensions of Information

Effective TQA aligns two interacting dimensions of temporality.

Corpus temporality

- **Diachronic:** time-stamped docs spanning years (news archives) — retrospective, event-based reasoning.
- **Synchronic:** a snapshot of the world at a particular time (e.g., a Wikipedia dump).

Question temporality

- **Explicitness:** explicit vs implicit temporal intent.
- **Orientation:** past, present or future.
- **Complexity:** simple lookup vs multi-hop reasoning.

Diachronic Vs Synchronic Collection

Different temporal views of the world require different retrieval strategies.

Diachronic Collections

Documents span months, years, or decades.

Characteristics

- Preserve how information evolves over time
- Multiple versions of facts may coexist
- Support historical and event-centric retrieval

Examples: News archives, Web archives, Historical document collections

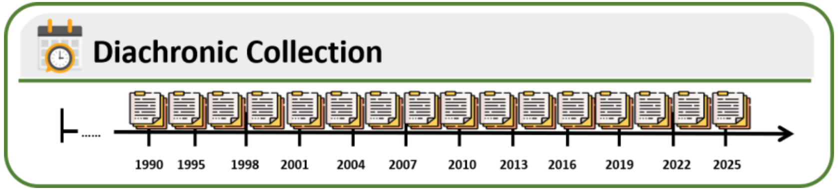
Synchronic Collections

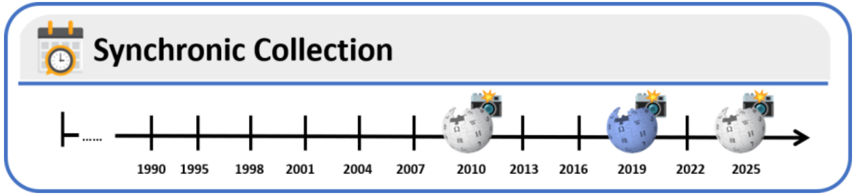
Snapshot of information at a single point in time.

Characteristics

- Represents the world "at particular point of time”
- Typically contains one current version of facts
- Easier to do temporal reasoning

Examples: Wikipedia dump, Snapshot of websites at particular time

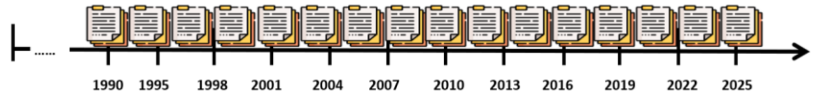




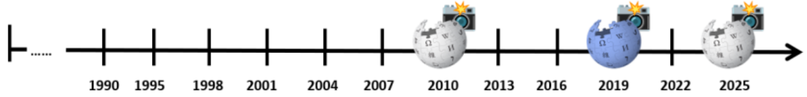
Diachronic collections support reasoning across time, while synchronic collections support reasoning within one snapshot of knowledge at a specific moment.

Diachronic Vs Synchronic Collection

Diachronic Collection



Synchronic Collection



Document from Diachronic Collection

Document Publication Date (DPD): **August 4, 2015**

Obama said, "Today, I announce that we are taking the most significant step in U.S. history to combat climate change. With the Clean Power Plan, we are setting the first-ever national limits on carbon pollution from power plants. By 2030, we will cut emissions by 32% from 2005 levels, and we'll do it by investing in cleaner energy like wind and solar not just for our health, but for the future of our planet."

And next Tuesday, the United States will join nearly 200 nations in the Paris Agreement. This global deal commits us all to limit global warming to well below 2 degrees Celsius. It's a turning point not just for our climate, but for our shared leadership on the world stage. We are showing that the U.S. does not sit on the sidelines when the future of our children is at stake.

Document from Synchronic Collection

Barack Hussein Obama (born August 4, 1961) is an American politician who was the 44th president of the United States from 2009 to 2017. .... Obama previously served as a U.S. senator representing Illinois from 2005 to 2008 and as an Illinois state senator from 1997 to 2004. ....

Obama was awarded the 2009 Nobel Peace Prize for efforts in international diplomacy, a decision which drew both criticism and praise. During his first term, his administration responded to the 2008 financial crisis with ...... took steps to combat climate change, signing the Paris Agreement, a major international climate agreement, and an executive order to limit carbon emissions.

Obama enrolled at Harvard Law School in the fall of 1988, living in nearby..... ...graduated from Harvard Law in 1991 with a Juris Doctor magna cum laude. He then taught constitutional law at the University of Chicago Law School for twelve years, first as a lecturer from 1992 to 1996, and then as a senior lecturer from 1996 to 2004.

Temporal Expressions

Temporal expressions are the primary way users and documents refer to time. Correct detection and interpretation are essential for retrieval and reasoning.

Types of Temporal Expressions

- **Absolute Temporal Expressions**

Denotes specific calendar references
Examples: June 2026, July 2, 2002, 2020s

- **Relative Temporal Expressions**

Depend on the reference time.
Examples: yesterday, last week, next month

- **Event-based Temporal expression (temponyms)**

Defined relative to an event.
Examples: after World War II, before the olympics, during Trump‘s first presidency

The Temporal Grounding Challenge

- Relative and sometimes Event-based Temporal Expressions requires context for proper anchoring.
- They must be grounded using context, events or reference time.

"last year"	*"during Olympics"*	*"after WWII"*
Reference Time Depends on when the query is issued.	**Event Knowledge** Resolved using knowledge of which Olympics is being referred to.	**Historical Event** Resolved using knowledge of the historical event.
2024 (if query issued in 2025)	**26 Jul–11 Aug 2024** (if the query refers to the Paris 2024 Olympics)	**after 1945** (end of WWII)

Temporal grounding converts natural-language time expressions into explicit dates or intervals, enabling **accurate retrieval and reasoning**.

Temporal Granularity

Temporal information can be expressed and reasoned over at different levels of resolution, from centuries to minutes or seconds.

Why Granularity Matters

- Different questions need different resolution.
- Wrong granularity may miss relevant evidence.
- Reasoning may move between granularities.

Example:

Q: What happened during the Renaissance? → Century-Level

Q: Who won the Nobel Prize in 2020? → Year-Level

Q: What happened on Sept. 11, 2001? → Day-Level

Metadata Time vs Content Time

A document's **publication date** is not always the same as the **time period it discusses**.

Publication Time

When the document was written or published.
Used as a proxy for freshness.

Focus Time

What the document is primarily about.
Captures temporal relevance.

Example document: *"A retrospective analysis of the 2010 Academy Awards"*

2010 — **Focus Time** (Academy Awards)

≈ 3-year gap

2013 — **Publication Time** (Article published)

Query: *"Who won Best Picture in 2010?"*
A system relying only on publication dates may miss relevant documents published years later.

Effective temporal retrieval reasons over both **when a document was written** and **what time it is about**.

Temporal Proximity

Focus time or publication time of documents closer to the target time are often more relevant.

Example:

- **Latest Apple earnings**

Apple earnings (2018) — ✕ Too Old

Apple Q4 earnings (2024) — ✓ Relevant

Apple Q2 earning (2026) — ✓✓Best Match

- **Covid-19 lockdown policies**

Government Directives (2020) — ✓✓Highly Relevant

Policy Update (2021) — ✓Relevant

Retrospective Analysis (2026) — ✕ Likely Less Relevant

Temporal relevance depends not only on content, but also on proximity to the user’s search intent’s target time.

Explicit vs Implicit Temporal Intent

Temporal constraints may be directly stated or inferred from context.

Explicit Intent

Time is directly mentioned.

Examples:

- Olympics 2024
- US Elections 2016
- Nobel Prize 2020

Query → Time Mention → Temporal Constraint

vs.

Implicit Intent

Time must be inferred.

Examples:

- Latest Apple earnings
- Current Prime Minister of India
- COVID-19 lockdown policies

Query → Temporal Context → Temporal Constraint

Implicit temporal intent is often **more challenging** because the relevant time period is not explicitly stated.

Temporal Reasoning Types

Temporal IR/QA requires reasoning beyond retrieval.

Ordering

Reason about the order of events.

World War I
↓
World War II

Example Question

What war happened before World War II?

Duration

Reason about the length of time.

2009 — 2017

Obama Presidency

Example Question

How long did Obama serve as president?

Arithmetic

Perform calculations with time.

1961 → 2009

2009 – 1961 = 48

Example Question

How old was Obama when he won the Nobel Prize?

Comparison

Compare events or time points.

FIFA WC Russia vs. FIFA WC Qatar

Example Question

Which FIFA World cup was more popular: the one in Qatar or the one in Russia.?

SUTime: A library for recognizing and normalizing time expressions

Core idea

- Deterministic rule-based temporal tagger for English.
- Implemented as a Stanford CoreNLP annotator and Java library.
- Uses TokensRegex to express patterns over tokens, POS tags, and existing annotations.
- Maps text to temporal objects, then outputs TIMEX3 annotations.

Recognition pipeline

Text regex rules → **Compositional rules** → **Filtering rules**

Rules are applied in stages: first simple expressions, then composite intervals, then filters for ambiguous candidates.

Supported temporal objects

Object	Example	Normalized Form
Time	Last Friday	2011-09-16
Duration	3days	P3D
Interval	From July to August	Begin/end interval
Set	Every third Sunday	XXXX-WXX-7

Example normalization

Document date: 2011-09-19
“last Friday” → DATE 2011-09-16

Angel X. Chang and Christopher Manning. 2012. **SUTime: A library for recognizing and normalizing time expressions**. In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), pages 3735–3740, Istanbul, Turkey. European Language Resources Association (ELRA).

HeidelTime: High Quality Rule-Based Extraction and Normalization of Temporal Expressions

Core idea

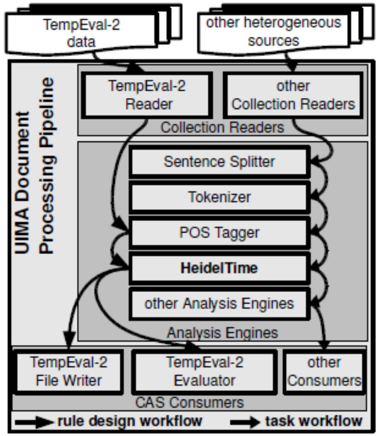
- Hand-crafted rules organized by TIMEX3 type: Date, Time, Duration, Set.
- Extraction mainly uses regular-expression patterns plus POS constraints.
- Normalization uses knowledge resources for months, seasons, holidays, etc.
- Post-processing resolves underspecified values using document creation time or a previously mentioned date.

Rule structure

expression pattern + normalization function + type

Example: “Independence Day 2010” → DATE 2010-07-04

UIMA-based pipeline



- Sentence splitting, tokenization and POS tagging feed the temporal tagger.
- Post-processing resolves underspecified values and removes invalid overlaps.

Jannik Strötgen and Michael Gertz. 2010. **HeidelTime: High Quality Rule-Based Extraction and Normalization of Temporal Expressions**. In Proceedings of the 5th International Workshop on Semantic Evaluation, pages 321–324, Uppsala, Sweden. Association for Computational Linguistics.

Temponym Tagging: Temporal Scopes for Textual Phrases

Core idea

- Standard temporal taggers identify explicit TIMEX3 expressions: dates, times, durations, and sets.
- Temponyms are textual phrases that are not temporal expressions themselves, but still denote a temporal scope.
- Extends HeidelTime to recognize explicit temponyms and normalize them to dates or intervals.

Examples: from phrase to temporal scope

Phrase	Type	Temporal Scope
WWW Conference 2016	Event-style	2016-04-11 to 2016-04-15
FIFA World Cup Final 1998	Event-style	1998-07-12
Hillary Clinton's term as Secretary of State	Event-style	2009-01-21 to 2013-02-01

Why it matters

Plain phrase
"WWW Conference 2016"

→

Temporal interval
[2016-04-11, 2016-04-15]

- Useful for temporal IR and exploration.
- Captures event-style and fact-style scopes.
- Bridges tagging and knowledge-base grounding.

not every temporal clue looks like a date.

Erdal Kuzey, Jannik Strötgen, Vinay Setty, and Gerhard Weikum. 2016. **Temponym Tagging: Temporal Scopes for Textual Phrases.** In Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16 Companion). https://doi.org/10.1145/2872518.2889289

tieval: An Evaluation Framework for Temporal Information Extraction Systems

tieval is an open-source Python library for the development and evaluation of Temporal Information Extraction (TIE) systems — identifying events, time expressions, and the temporal links between them.

An open-source library

A pip-installable Python package — MIT-licensed and available on PyPI & GitHub.

Built for temporal IE

Focused on events, temporal expressions (timex), and the temporal relations (tlinks) that connect them.

One unified interface

Replaces per-corpus parsers and ad-hoc metrics with a single, consistent API.

`pip install tieval` — its goal: make TIE benchmarking fair, comparable, and reproducible.

Hugo Sousa, Ricardo Campos, and Alípio Mário Jorge. 2023. **Tieval: An Evaluation Framework for Temporal Information Extraction Systems**. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23). https://doi.org/10.1145/3539618.3591892

tieval: An Evaluation Framework for Temporal Information Extraction Systems

It supports the entire TIE pipeline through three modules — datasets, models, and evaluation.

datasets

- Reads 18+ corpora across 11+ languages
- Unifies them in one object model
- Loads a full corpus in a single line

models

- Baselines for timex & event ID
- HeidelTime and CogCompTime2
- Train from scratch or just predict

evaluation

- Scores four TIE subtasks
- P / R / F1, macro and micro
- Temporal awareness via closure

Hugo Sousa, Ricardo Campos, and Alípio Mário Jorge. 2023. **Tieval: An Evaluation Framework for Temporal Information Extraction Systems**. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23). https://doi.org/10.1145/3539618.3591892

Temporal Prediction Tasks

A family of tasks that infer implicit or missing temporal information directly from text — anchoring documents, queries, and events in time.

- **Bridge the gaps**
Critical when explicit temporal metadata is sparse, noisy, or unavailable.

- **Align the pieces**
Improve alignment between queries, documents, and the events they describe.

- **Enable applications**
Power historical search, timeline construction, and time-sensitive retrieval.

HOW THE METHODS HAVE GROWN

1. **Rule-based**
Handcrafted normalization & decay functions
2. **Statistical LMs**
Time-period language models, ranking features
3. **Neural**
Transformer encoders, temporal embeddings, graph-based reasoning

Document Dating

DEFINITION

Estimate a document's creation time (e.g., publication date) from its textual content — especially when metadata is missing, unreliable, or absent.

INPUT
Full document text

→ **OUTPUT**
Timestamp (year / month)

EXAMPLE

A historical newspaper page with no reliable header date is placed on the timeline using the vocabulary and style of its era.

KEY APPROACHES

de Jong et al. (2005)

Unigram language models over distinct time periods locate when a document's vocabulary was most prevalent.

Kanhabua & Nørvåg (2008)

Add POS tags, tf-idf scores, and collocations to capture temporal patterns.

Dalli (2006) · Kumar et al. (2012)

Unsupervised periodic word usage; LMs over "chronons" from Wikipedia biographies.

Niculae (2014) · Vashishth (2018)

Pairwise ranking via logistic regression; neural GCNs over syntactic & temporal relations.

Jong, F. M., Henning Rode, and Djoerd Hiemstra. "**Temporal language models for the disclosure of historical text**." (2005).
Kanhabua, Nattiya, and Kjetil Nørvåg. "**Improving temporal language models for determining time of non-timestamped documents**." In International conference on theory and practice of digital libraries,2008.
Dalli, Angelo. "**Temporal classification of text and automatic document dating**." In Proceedings of the Human Language Technology Conference of the NAACL, 2006.
Kumar, Abhimanu, Jason Baldridge, Matthew Lease, and Joydeep Ghosh. "**Dating texts without explicit temporal cues.**" arXiv preprint arXiv:1211.2290 (2012).
Niculae, Vlad, et al. "**Temporal text ranking and automatic dating of texts**." Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics. 2014.
Vashishth, Shikhar, et al. "**Dating documents using graph convolution networks**." Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018.

Document Focus Time Estimation

DEFINITION

Identify the historical period(s) a document discusses — its narrative scope — which may differ from when it was written.

INPUT
Full document text

→

OUTPUT
Temporal interval(s)

EXAMPLE

An article published in 2021 that analyzes the 9/11 attacks has a focus time centered on September 2001, not 2021.

KEY APPROACHES

Jatowt et al. (2013)

Graph-based method modeling co-occurrences between terms and dates to surface salient temporal associations.

Jatowt et al. (2015)

Estimates focus time from statistical evidence in external corpora, even when explicit time expressions are sparse.

Shrivastava et al. (2017)

Links documents to Wikipedia concepts and leverages their temporal relations to infer focus times.

Jatowt, Adam, Ching-Man Au Yeung, and Katsumi Tanaka. "**Estimating document focus time.**" Proceedings of the 22nd ACM international conference on Information & Knowledge Management. 2013.
Jatowt, Adam, Ching Man Au Yeung, and Katsumi Tanaka. **"Generic method for detecting focus time of documents."** Information Processing & Management 51.6 (2015): 851-868.
Shrivastava, Shashank, Mitesh Khapra, and Sutanu Chakraborti. **"A concept driven graph based approach for estimating the focus time of a document."** *International Conference on Mining Intelligence and Knowledge Exploration*. Cham: Springer International Publishing, 2017.

Temporal Query Profiling

DEFINITION

Determine a query's temporal intent and time of interest — whether it refers to the past, the future, or is atemporal.

INPUT
Short keyword query → **OUTPUT** Inferred time / distribution

EXAMPLE

The query "Ukraine–Russia war" is profiled into an inferred time of interest or a distribution over candidate periods.

KEY APPROACHES

Kanhabua & Nørvåg (2010)

Estimate query time from the timestamps of the top-k retrieved documents.

Dakka (2008) · Jones & Diaz (2007)

Model the temporal distributions of relevant documents to characterize intent.

Kanhabua & Nørvåg (2011)

Compare five temporal ranking methods (LMT, LMTU, TS, TSU, FuzzySet).

Gupta & Berberich (2014)

Combine timestamp metadata with in-content temporal expressions for precise intervals.

Kanhabua, Nattiya, and Kjetil Nørvåg. **"Determining time of queries for re-ranking search results."** International conference on theory and practice of digital libraries. Berlin, 2010.
Dakka, Wisam, Luis Gravano, and Panagiotis G. Ipeirotis. **"Answering general time sensitive queries."** Proceedings of the 17th ACM conference on Information and knowledge management. 2008.
Jones, Rosie, and Fernando Diaz. **"Temporal profiles of queries."** ACM Transactions on Information Systems (TOIS) 25.3 (2007): 14-es.
Kanhabua, Nattiya, and Kjetil Nørvåg. **"A comparison of time-aware ranking methods."** Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. 2011.
Gupta, Dhruv, and Klaus Berberich. **"Identifying time intervals of interest to queries."** Proceedings of the 23rd ACM international conference on conference on information and knowledge management. 2014.

Event Occurrence Time Estimation

DEFINITION

Predict the specific date an event occurred from a short textual description — focusing on the event mention itself, at high granularity.

INPUT
Short event description → **OUTPUT** Fine-grained date (day / month)

EXAMPLE

Given "Plane crash in Armenia kills 36," the system infers the day- or month-level date of the event.

KEY APPROACHES

Das et al. (2017)

Time vectors combining word and global temporal embeddings; estimate dates via cosine similarity.

Morbidoni et al. (2018)

Use structured KBs (DBpedia, Wikipedia) to link descriptions to grounded entities.

Honovich et al. (2020)

Sentence extraction + LSTM with attention + an MLP classifier for date prediction.

Wang et al. (2021b) — TEP-Trans

Transformer framing event-time prediction as multivariate time-series forecasting.

Das, Supratim, et al. **"Estimating event focus time using neural word embeddings."** Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017.
Morbidoni, Christian, Alessandro Cucchiarelli, and Domenico Ursino. **"Leveraging linked entities to estimate focus time of short texts."** Proceedings of the 22nd International Database Engineering & Applications Symposium. 2018.
Honovich, Or, et al. **"Machine reading of historical events."** Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020.
Wang, Jiexin, Adam Jatowt, and Masatoshi Yoshikawa. **"Event occurrence date estimation based on multivariate time series analysis over temporal document collections."** Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval. 2021.