Chapter 9

Temporal Information Access

Masaharu Yoshioka and Hideo Joho

Abstract This chapter introduces the research background and details of temporal information access tasks in the NTCIR. The GeoTime task was the first attempt to evaluate temporal information retrieval as an extension of an information-retrieval-for-question-answering task. Temporalia was a task to investigate the role of temporal factors in a search.

9.1 Introduction

Temporal information is important to understand the document and to represent users' information needs. In the early age of Named Entity Recognition (NER), tasks such as MUC-6 (Sundheim 1995) and IREX (Sekine and Isahara 2000), date and time were selected as categories for NER. In information access technology research, there had been several studies on using such temporal information (e.g., Mani et al. 2004), but there have not been many studies on temporal information retrieval (Alonso et al. 2007).

Compared to the usage of temporal information, Geographical Information Retrieval (Geographic Information Retrieval (GIR)) had attracted more researchers, and a series of workshops on Geographic Information Retrieval (GIR) was started in 2004 (Purves and Jones 2004). In this series of workshops, temporal information was only discussed as a related topic of the task.

At NTCIR-8, *GeoTime* (geographic and temporal information retrieval) tasks (Gey et al. 2010) were launched as first attempts to construct a test collection for

M. Yoshioka (✉)

Faculty of Information Science and Technology, and Global Station for Big Data and Cybersecurity, Global Institution for Collaborative Research and Education, Hokkaido University, Sapporo, Japan
e-mail: yoshioka@ist.hokudai.ac.jp

H. Joho

Faculty of Library, Information and Media Science, University of Tsukuba, 1-2 Kasuga, Tsukuba, Ibaraki 305-8550, Japan
e-mail: hideo@slis.tsukuba.ac.jp

© The Author(s) 2021

T. Sakai et al. (eds.), *Evaluating Information Retrieval and Access Tasks*,
The Information Retrieval Series 43,
https://doi.org/10.1007/978-981-15-5554-1_9

temporal information retrieval. This task was designed as an extension of IR4QA tasks (Mitamura et al. 2008). There were two types of temporal-related queries. One query type asked for temporal information (“when” question), while the other query type used temporal information as constraints (winning team of Superbowl in 2002). Details of the information related to the task are discussed in Sect. 9.2.

Following the success of GeoTime Tasks in NTCIR-8 and 9, a new task was proposed to further investigate the role of temporal factors in the search. The task was called *Temporalia* (Temporal Information Access) and was run twice in NTCIR-11 and 12. One of the important innovations in *Temporalia* was to provide a test collection that allowed researchers to examine the performance of time-aware search applications using categories such as past, recent, future, and atemporal rather than focusing on recency queries. Details of the information related to the task are discussed in Sect. 9.3.

9.2 Temporal Information Retrieval

There are several IR applications that utilize temporal information; e.g., ad hoc retrieval, hit-list clustering based on the temporal aspect, exploratory search, and visualization of results based on the temporal relationships (Alonso et al. 2007). However, there was no IR evaluation campaign for temporal information retrieval except for some discussions related to Geographic Information Retrieval (GIR) (Purves and Jones 2004).

To utilize and incorporate the discussion related to *Geographic Information Retrieval (GIR)*, *GeoTime* (geographic and temporal information retrieval) tasks (Gey et al. 2010) were launched at NTCIR-8 as an extension of IR4QA for handling spatial and temporal-related queries (Mitamura et al. 2008).

9.2.1 NTCIR-8 GeoTime Task

The NTCIR-8 GeoTime Task was designed as an IR4QA task for the geographical and temporal-related queries.

Parts of queries were constructed using the information of notable events listed in Wikipedia,¹ and several queries were derived from the *ACLIA* collection (Sakai et al. 2010). This task used the New York Times collection for the English document database and the Mainichi Japanese newspaper collection for the Japanese document database.

For the evaluation, because most of the queries have both temporal and spatial aspects, the articles that can be used for answering questions for temporal and spatial aspects were categorized as “fully relevant” and ones that can answer only one aspect

¹For example, notable events in 2002 are listed at <https://en.wikipedia.org/wiki/2002>.

(temporal or spatial) are categorized as “partially relevant”. The submitted results were evaluated by the same schemes used for the ACLIA IR4QA collection (Sakai et al. 2010).

The following are examples of the queries.

- How old was Max Schmeling when he died, and where did he die?
- When and where did a massive earthquake occur in December 2003?

The former question asks for temporal information using a “when” question. The latter question also has the “when” question style, but it also uses temporal information to represent constraints (“in December 2003”).

There were 14 teams that participated in NTCIR-8 GeoTime (8 and 7 teams submitted runs for Japanese and English runs, respectively) using various approaches (Gey et al. 2010). The baseline system utilized ordinary ad hoc IR systems such as probabilistic IR with blind relevance feedback. This baseline system worked well for the English run but underperformed in the Japanese run. Another approach utilized a NER system and/or geographic resources to extract named entity information including geographic and temporal information from the queries and documents. The best performing NTCIR-8 Japanese run was a hybrid approach that combined the probabilistic approach and weighted Boolean query formulation based on the NER results (Yoshioka 2010). There were approaches that focused on geographic information including the hierarchical relationship among location names (e.g., Tokyo is a part of Japan) and the distance between the extracted location of the query and document, and there were several discussions about the temporal information.

Another approach emphasized the style of the query in GeoTime. Because the query was provided as a question in IR4QA style, the relevant documents should contain the information for its answer. Based on this understanding, one team counted the number of temporal or geographic mentions that can be candidates for the answer for re-ranking (Kishida 2010). Another approach decomposes the question into one for geographic information and another for temporal information. After decomposing the question, they used a factoid question answering system to determine the answer and utilize its information for constructing new queries (Mori 2010). However, those approaches did not perform well for the task.

From the analysis of the difficult queries based on the evaluation of the submitted results, two types of difficult queries were identified. One type is that the system tends to misinterpret the constraint of the query. An example of the query is “When and where were the 2010 Winter Olympics host city location announced?”. In this question, “2010” is used as a part of an event name and not as a constraint-specifying articles should be selected from those published in 2010 or after. Another type of difficult query requires a list of events to determine relevant articles. An example of this type of query is “When and where were the last three Winter Olympics held?”. It is difficult to retrieve relevant articles without generating an event list that satisfies the query constraint. Details of the discussion about the difficulties of the problem are addressed in Sect. 9.2.3.

9.2.2 NTCIR-9 GeoTime Task Round 2

By comparing the English runs and Japanese runs, there were queries that have large performance variability for the same topics. Therefore, the news article data for English runs were expanded to include newspapers from different countries. In addition to the news articles of the New York Times collection, English versions of Korean Times (Korea), Mainichi (Japan), and Xinhua (China) were used to construct a document database.

There were 12 teams that participated for NTCIR-9 GeoTime (5 and 9 teams submitted runs for the Japanese and English runs, respectively) using various approaches (Gey et al. 2011). One large difference from the previous GeoTime was the usage of external resources such as Yahoo PlaceMaker, Wikipedia, DBpedia, Geonames, Google Maps, and the Alexandria Digital Library gazetteer. Most of the teams utilized such information for improving the retrieval results related to the geographic queries. However, the query that required reverse geocoding (finding place names from a latitude/longitude information) was not appropriately handled except that the team manually extracted the related event name using Wikipedia.

The best performing team for both Japanese and English runs used manual query expansion with a related event name and/or name of the location using Wikipedia and Google Maps (Sato 2011). Because this approach was not automatic, it was difficult to compare this result with others. However, this result suggested that the extraction of such related event names or locations is crucial for improving the recall of the related articles.

9.2.3 Issues Discussed Related to Temporal IR

One of the difficult queries in NTCIR-8 GeoTime was “When and where were the 2010 Winter Olympics host city location announced?”. To discuss the difficulties of this query, it was necessary to discuss the types of temporal expression. Alonso et al. (2007) proposed the following types of temporal expression.

1. Explicit. Temporal expression directly describes its information (e.g., September 11, 2001).
2. Implicit. There is imprecise temporal information, such as names of holidays or events. It is possible to extract temporal information using knowledge about such holidays or events (e.g., Labor Day, 2001, can be mapped to September 1, 2001, and Vancouver Winter Olympics can be mapped to February 2010).
3. Relative. Temporal expressions represent temporal entities that refer to other temporal entities. Temporal information resolution is necessary to extract its temporal information (e.g., “yesterday” of the news article published on September 12, 2001, can be mapped to September 11, 2001).

In the query discussed above, “the 2010 Winter Olympics” is the name of the event and can be treated as an implicit temporal expression. However, it is not a

constraint for selecting relevant articles. It is necessary to have a mechanism to select which kinds of temporal expression should be used for constraints to retrieve relevant articles.

Another problem is related to handling the relationship between temporal information and event names that represent imprecise temporal information. An example of this difficult query is “When and where were the last three Winter Olympics held?”; “the last three” uses relative and imprecise temporal information to select relevant event names (three Winter Olympic event names). Because most of the relevant documents contain such event names but do not have such relative expression, it is difficult to retrieve such articles without event names. As we confirmed in the case of NTCIR-9 GeoTime, query expansion using such event names significantly improves the performance.

9.3 Temporal Query Analysis and Search Result Diversification

To facilitate research on temporal information access, Temporalia-1 in NTCIR-11 (Joho et al. 2014) focused on each of the four categories in a structured way, while Temporalia-2 in NTCIR-12 (Joho et al. 2016) was designed to encourage researchers to explore ways to combine the four categories in a meaningful way. Both were designed to address the temporal ambiguity and diversity of the search space.

9.3.1 NTCIR-11 Temporal Information Access Task

Temporalia-1 at NTCIR-11 consisted of two subtasks: Temporal query intent classification and temporal information retrieval.

9.3.1.1 Temporal Query Intent Classification

The Temporal Query Intent Classification (TQIC) subtask was used to classify a given query into one of the following classes: past, recency, future, and atemporal. Example queries are ground truth temporal classes are shown in Table 9.1. The classes were defined as follows.

Past: class characterizing queries about past entities/events whose search results are not expected to change much with the passage of time.

Recency: class characterizing queries about recent entities/events, whose search results are expected to be timely and up to date. The information contained in the search results usually changes quickly with the passage of time. Note that this type of query usually refers to events that happened in the near past or at the present time.

Table 9.1 Example queries and ground truth temporal classes for the TQIC subtask (dry run)

Query class	Query example
Past	Price hike in Bangladesh 2008
Past	Who was Martin Luther
Past	When did the Titanic sink
Past	Yuri Gagarin cause of death
Past	History of Coca-Cola
Recency	Apple stock price
Recency	Number of millionaires in USA
Recency	Time in London
Recency	Trendy plus size clothing
Recency	Did the Pirates win today
Future	2013 MLB playoff schedule
Future	Release date for iOS7
Future	College baseball regional projections
Future	Disney prices 2014
Future	Long-term weather forecast
Atemporal	Blood pressure monitor
Atemporal	Distance from earth to sun
Atemporal	How to start a conversation
Atemporal	New York Times
Atemporal	Lose weight quickly

In contrast, the “past” query category tends to refer to events in a relatively distant past.

Future: class characterizing queries about predicted or scheduled events, and the search results of which should contain future-related information.

Atemporal: class characterizing queries without any clear temporal intent (i.e., their returned search results are not expected to be related to time and should not change much over time). Navigational queries are considered to be atemporal.

Participants were handed a set of query strings and query submission dates and were asked to develop a system to classify each of the query strings to one of the four above-mentioned temporal classes. As this problem rather requires different kinds of knowledge (e.g., historical information or information on planned events), the participants were allowed to use any external resources to complete the TQIC subtask as long as the details of external resource usage were described in their reports. Each participating team was asked to submit a temporal class (past, recency, future, or atemporal) for each one of the queries. The performance of submitted runs was measured by the number of queries with correct temporal classes divided by the total number of queries.

Table 9.2 Example topics for the TIR subtask (dry run)

	Girl with the Dragon Tattoo
Description	I have recently watched a film called Girl with the Dragon Tattoo and I really liked it. Therefore, I would like to gather information about the movie
Past question	How did the casting of the film develop?
Recency question	What did the recent reviews say about the film?
Future question	Is there any plan about a sequel?
Atemporal question	What are the names of the main actors and actresses of the film?
Search date	28 Feb 2013 GMT+0:00

9.3.1.2 Temporal Information Retrieval

The Temporal Information Retrieval (TIR) subtask was used to retrieve a set of documents in response to a search topic that incorporates a time factor. In addition to a typical search topic description (i.e., title, description, and subtopics), the TIR search topic description also contains a query submission date (see Table 9.2). This subtask required indexing of the document collection with any standard information retrieval toolkit. Participants were asked to submit the top 100 documents for each temporal question per topic (e.g., top 100 documents for a past question and another 100 for a recency question). The retrieval effectiveness was evaluated by the precision at 20 for each of the temporal questions. Similar to the TQIC subtask, the results section presents an analysis of the performance across temporal questions.

9.3.2 NTCIR-12 Temporal Information Access Task Round 2

Temporalia-2 at NTCIR-12 also consisted of two subtasks: temporal intent disambiguation and temporally diversified retrieval.

9.3.2.1 Temporal Intent Disambiguation

The Temporal Intent Disambiguation (TID) subtask determined a probability distribution of a query over four classes denoting the types of temporal intent: past, recency, future, and atemporal. The definitions of the four classes were based on TQIC in Temporalia-1. An example of the probability distribution of temporal intents is shown in Tables 9.3.

Table 9.3 Example queries for the TID subtask (dry run) with query submission date of May 1, 2013. Ground truth probability of temporal intents was determined by votes from crowd workers

Query	Past	Recency	Future	Atem.
Australian open	0.091	0.0	0.455	0.455
Motorcycle accident June	0.7	0.0	0.3	0.0
NBA finals	0.1	0.0	0.4	0.5
NBA playoff schedule	0.0	0.2	0.6	0.2
Price of oil	0.0	0.9	0.0	0.1
How to lose weight	0.0	0.1	0.0	0.9
Time in India	0.0	1.0	0.0	0.0
History of volleyball	1.0	0.0	0.0	0.0

Table 9.4 Example topics for the TDR subtask

	Junk food health effect
Description	I am concerned about the health effects of junk food in general. I need to know more about their ingredients, impact on health, history, current scientific discoveries, and any prognoses
Past question	When did junk foods become popular?
Recency question	What are the latest studies on the effect of junk foods on our health?
Future question	Will junk food continue to be popular in the future?
Atemporal question	How junk foods are defined?
Search date	29 May 2013 GMT+0:00

9.3.2.2 Temporally Diversified Retrieval

The Temporally Diversified Retrieval (TDR) subtask required participants to retrieve a set of documents relevant to each of four temporal intent classes for a given topic description (see Table 9.4). Participants were also asked to return a set of documents that is temporally diversified for the same topic. They received a set of topic descriptions, query issuing times, and indicative search questions for each temporal class (past, recency, future, and atemporal). The objective of the indicative search questions was to show one possible subtopic under a particular temporal class. Participants were asked to develop systems that can produce a total of five search results per topic (past, recency, future, atemporal, and diversified).

9.3.3 *Implications from Temporalia*

This section discusses the implications of Temporalia tasks on system development and test collection, respectively.

9.3.3.1 **Implications on System Development**

From the meta-analysis of 17 runs submitted to the TQIC subtask, the classification of recency queries was found to be the most difficult with 56% accuracy, and past queries were the easiest with 73%. Another overall trend was that no single approach was effective across the four temporal classes. A confusion matrix showed that: (1) atemporal queries are likely to be confused as either recency or past queries (16.7% and 9.6%, respectively), (2) past queries are likely to be confused as atemporal queries (13.1%), (3) recency queries are likely to be confused as future or atemporal queries (28.2% and 13.5%, respectively), and (4) future queries tend to be confused as recency queries (25.9%). Correlation analysis suggested that it was difficult to apply the same technique to predict recency queries and atemporal queries with high accuracy.

The TIR subtask showed a similar pattern with varied performance across the four classes. No single system was able to perform the best for all classes. The learning-to-rank approach was effective for atemporal and past queries, while BM25 performed well for recency and future topics.

The meta-analysis of 37 runs submitted to the TID subtask suggested that when a query was temporally ambiguous and multiple temporal classes can be inferred, detecting atemporal features was the most difficult. Also, some techniques were good at modeling temporally less diverse queries (i.e., a fewer number of nonzero probability classes), while other methods were good at modeling temporally more diverse queries.

The results of the TDR subtask suggested that a learning-to-rank approach was effective in retrieving relevant documents for all classes compared to BM25. However, the best performance on temporal search result diversification was obtained by a round-robin of BM25 rankings of four temporal classes, suggesting that there is still room for improvement in this area.

9.3.3.2 **Implications for Test Collection**

Document Collections

One of the challenges in building a test collection for temporal-aware technologies was to obtain access to document collections that have rich temporal features. Temporalia was fortunate to have support to use the “LivingKnowledge news and blogs annotated sub-collection” constructed by the LivingKnowledge project and distributed by the Internet Memory Foundation. The collection was approximately

20 GB large when uncompressed and over 5 GB large when zipped. The collection spanned from May 2011 to March 2013 and contains around 3.8 million documents collected from about 1,500 different blogs and news sources. The data were split into 970 files based on the date and sources (there might be more than one file per day). Texts in the collection were annotated by entities and by temporal expressions that were resolved to a specific day, month, or year (Matthews et al. 2010). The relative expressions such as “next month” was resolved based on the publication date of the articles.

In Temporalia-2, we also made efforts to diversify the target language of document collections to Chinese using SogouCA-2012² and SogouT-2012.³ Similar to the English collection, SogouCA-2012 was based on news articles from major publishers in China. For annotating temporal expressions, a variant of the standard format TIMEX3 used in TempEval task was applied.⁴

Relevance Assessments

Another challenge we faced during the construction of Temporalia test collections was relevance assessment. The temporality of topics and relevance can be subjective and not always deterministic. Therefore, we used a mixture of methods to ensure that both queries and documents were temporally annotated for evaluation.

We had a combination of workshops and crowdsourcing in formal runs. In another series of workshops, participants (not necessarily the same people as topic creators) were asked to read the formal run topic descriptions carefully and assess the relevance of the retrieved documents.

The documents were then evaluated using crowdsourcing as for their relevance to each of the temporal subclasses. For each assigned subtopic, CrowdFlower workers were asked to identify at least one highly relevant and one irrelevant document. They were also asked to note the relevant text from original documents in the case of highly relevant documents. The relevance of these documents was verified by a third person during the workshop to improve their reliability.

The documents initially identified by the workshop participants were then used as “test questions” of crowdsourcing jobs. Test questions were questions that crowdsourcing workers had to pass to participate in our relevance assessment jobs. We used CrowdFlower to run relevance judgments. Our configuration of crowdsourcing is based on common settings used by various IR evaluations (e.g., Kazai et al. 2013).

- Each task had five documents to judge
- Ten cents were paid for one task
- Each task had 120 s of minimum work time
- Each document had at least three judgments

²<http://www.sogou.com/labs/dl/ca.html>.

³<http://www.sogou.com/labs/dl/t-e.html>.

⁴<http://www.timeml.org/tempeval2/tempeval2-trial/guidelines/timex3guidelines-072009.pdf>.

We had several iterations of revising job instructions and relevance criteria before running all formal run subtopics. We tested both detailed instructions and simple instructions, but we received mixed responses from workers. Also, detailed instructions caused the time required for relevance assessment to increase too much. After several iterations, we decided to use the following three levels of relevance criteria.

Not Relevant The web page does not contain any information to answer the search question.

Highly Relevant The web page discusses the answer to the search question exhaustively. In the case of a multifaceted search question, all or most subthemes or viewpoints are covered. Typical extent: several text paragraphs, at least four sentences or facts.

Relevant The web page contains some information to answer the question, but the presentation is not exhaustive. In the case of a multifaceted search question, only some of the subthemes or viewpoints are covered. Typical extent: one text paragraph, or one to three sentences or facts.

9.4 Related Work and Broad Impacts

After introducing temporal information retrieval task as a part of GeoTime task at NTCIR 8, there were several lines of research emerged as a variation of temporal information retrieval. Kanhabua et al. (2015) is a comprehensive textbook that introduces such research results. Moulahi et al. (2016) also summarizes past efforts in temporal information retrieval evaluation and discuss future directions. From these results, we would like to introduce some research that is highly related to the tasks discussed above.

Strötgen and Gertz (2013), Daoud and Huang (2013) both proposed proximity methods for the Geotemporal Information Retrieval task. In this method, proximity of the geographic and temporal information are considered for ranking documents in addition to the standard information retrieval ranking such as BM25. Another interesting example is event-centric search and exploration (Strötgen and Gertz 2012). This framework was proposed for analyzing historic documents using geographic and temporal constraints constructed from event information. In the discussion of GeoTime, there was a consideration of using the name of an event for time constraints. This event-centric approach utilizes these characteristics to find documents relevant to the event for exploration.

There have been related efforts to construct test collections for Information Access technologies with temporal awareness, such as the TREC Temporal Summarization Track (2012–2015) (Aslam et al. 2015; Guo et al. 2013) and TREC Knowledge Base Acceleration Track (2012–2014) (Frank et al. 2014). The TREC Temporal Summarization Track had two subtasks: Sequential Update Summarization and Value Tracking. Sequential Update Summarization sought to find timely, sentence-level, reliable, relevant, and nonredundant updates about developing event, while Value Tracking

aimed at tracking values of event-related attributes that were of high importance to the event. TREC Knowledge Base Acceleration Track was a challenge for filtering a large stream of text to find documents that can help update knowledge bases like Wikipedia, Facebook, or Crunchbase. Both efforts either explicitly or implicitly had a focus on recency information about entities. NTCIR Temporalia was, on the other hand, designed to facilitate research on diverse temporal attributes in a systematic manner.

There have been several extensions of the original work. For example, Hasanuz-zaman et al. (2016) applied temporal query intent classification techniques to stock market analysis. Rizzo and Montesi (2017) used the LivingKnowledge collection to conduct a temporal analysis of a digital library collection. Finally, Joho et al. (2013, 2015) used the Temporalia test collection to study temporal information-seeking behavior in a controlled user study and a questionnaire-based study. The studies identified the difference in resource selection and relevant content types across temporal attributes of information needs. These are some of the ways in which the test collection for temporal information access can have broader impacts than the original objectives of the resources.

See the citation of the overview papers (Gey et al. 2010, 2011; Joho et al. 2014, 2016) for more details of broader impacts from GeoTime and Temporalia.

9.5 Conclusion

We have introduced two tasks related to temporal information access in the NTCIR workshop. GeoTime was the first attempt to place more emphasis on temporal search, and Temporalia provided a framework to examine the performance of time-aware search application using a test collection. The review of the literature suggests that these resources have been useful for researchers to advance temporal information access technologies and to better understanding temporal information-seeking behavior.

Acknowledgements The authors would like to thank Prof. Fred Gey, Prof. Ray Larson, Dr. Jorge Machado, Dr. Roi Blanco, Dr. Adam Jatowt, Dr. Haitao Yu, Dr. Shuhei Yamamoto, and Mr. Hajime Naka for their efforts to organize the NTCIR GeoTime and NTCIR Temporalia Tasks. The authors also thank all participants in the tasks for their interest and contributions to the technological advance in temporal information access. LivingKnowledge project, Internet Memory Foundation, and Sogou have provided valuable document collections that enabled us to build the test collections. Finally, the authors would like to thank Doug Oard, Tetsuya Sakai, and Noriko Kando for their editorial work on this manuscript.

References

- Alonso O, Gertz M, Baeza-Yates R (2007) On the value of temporal information in information retrieval. *SIGIR Forum* 41(2):35–41. <https://doi.org/10.1145/1328964.1328968>
- Aslam JA, Diaz F, Ekstrand-Abueg M, McCreadie R, Pavlu V, Sakai T (2015) TREC 2015 temporal summarization track overview. In: Voorhees EM, Ellis A (eds) *Proceedings of the 24th text retrieval conference, TREC 2015*, Gaithersburg, Maryland, USA, 17–20 November 2015, National Institute of Standards and Technology (NIST), vol Special Publication, pp 500–319. <http://trec.nist.gov/pubs/trec24/papers/Overview-TS.pdf>
- Daoud M, Huang JX (2013) Modeling geographic, temporal, and proximity contexts for improving geotemporal search. *J Am Soc Inf Sci Technol* 64(1):190–212. <https://doi.org/10.1002/asi.22648>
- Frank JR, Kleiman-Weiner M, Roberts DA, Voorhees EM, Soboroff I (2014) Evaluating stream filtering for entity profile updates in TREC 2012, 2013, and 2014. In: Voorhees EM, Ellis A (eds) *Proceedings of the 23rd text retrieval conference, TREC 2014*, Gaithersburg, Maryland, USA, 19–21 November 2014, National Institute of Standards and Technology (NIST), vol Special Publication, pp 500–308. <http://trec.nist.gov/pubs/trec23/papers/overview-kba.pdf>
- Gey FC, Larson RR, Kando N, Machado J, Sakai T (2010) NTCIR-GeoTime overview: evaluating geographic and temporal search. In: Kando et al 2010, pp 147–153. <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings8/NTCIR/01-NTCIR8-OV-GeoTime-GeyF.pdf>
- Gey FC, Larson RR, Machado J, Yoshioka M (2011) NTCIR9-GeoTime overview - evaluating geographic and temporal search: round 2. In: Kando et al 2011. <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings9/NTCIR/01-NTCIR9-OV-GEOTIME-GeyF.pdf>
- Guo Q, Diaz F, Yom-Tov E (2013) Updating users about time critical events. In: Serdyukov P, Braslavski P, Kuznetsov S, Kamps J, Rüger S, Agichtein E, Segalovich I, Yilmaz E (eds) *Advances in information retrieval. Lecture notes in computer science*, vol 7814. Springer, Berlin, pp 483–494. <https://doi.org/10.1007/978-3-642-36973-541>
- Hasanuzzaman M, Sze WL, Salim MP, Dias G (2016) Collective future orientation and stock markets. In: *Proceedings of the 22nd European conference on artificial intelligence, ECAI'16*, pp 1616–1617. IOS Press, Amsterdam. <https://doi.org/10.3233/978-1-61499-672-9-1616>
- Joho H, Jatowt A, Blanco R (2013) A survey of temporal web search experience. In: Carr L, Laender AHF, Lóscio BF, King I, Fontoura M, Vrandečić D, Aroyo L, de Oliveira JPM, Lima F, Wilde E (eds) *22nd international world wide web conference, WWW'13*, Rio de Janeiro, Brazil, 13–17 May 2013, Companion volume. International World Wide Web Conferences Steering Committee/ACM, pp 1101–1108. <https://doi.org/10.1145/2487788.2488126>
- Joho H, Jatowt A, Blanco R, Naka H, Yamamoto S (2014) Overview of NTCIR-11 temporal information access (temporalia) task. In: Kando N, Joho H, Kishida K (eds) *Proceedings of the 11th NTCIR conference on evaluation of information access technologies, NTCIR-11*, National Center of Sciences, Tokyo, Japan, 9–12 December 2014, National Institute of Informatics (NII). <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings11/pdf/NTCIR/OVERVIEW/01-NTCIR11-OV-TEMPORALIA-JohoH.pdf>
- Joho H, Jatowt A, Blanco R (2015) Temporal information searching behaviour and strategies. *Inf Process Manag* 51(6):834–850. <https://doi.org/10.1016/j.ipm.2015.03.006>, <http://www.sciencedirect.com/science/article/pii/S0306457315000448>
- Joho H, Jatowt A, Blanco R, Yu H, Yamamoto S (2016) Overview of NTCIR-12 temporal information access (temporalia-2) task. In: Kando N, Sakai T, Sanderson M (eds) *Proceedings of the 12th NTCIR conference on evaluation of information access technologies*, National Center of Sciences, Tokyo, Japan, 7–10 June 2016, National Institute of Informatics (NII). <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings12/pdf/ntcir/OVERVIEW/01-NTCIR12-OV-TEMPORALIA-JohoH.pdf>
- Kando N, Kishida K, Sugimoto M (eds) (2010) *Proceedings of the 8th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access, NTCIR-8*, National Center of Sciences, Tokyo, Japan, 15–

- 18 June 2010, National Institute of Informatics (NII). <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings8/NTCIR/>
- Kando N, Ishikawa D, Sugimoto M (eds) (2011) Proceedings of the 9th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access, NTCIR-9, National Center of Sciences, Tokyo, Japan, 6–9 December 2011, National Institute of Informatics (NII). http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings9/NTCIR/toc_ntcir.html
- Kanhabua N, Blanco R, NØrvåg K. (2015) Temporal information retrieval. *Found Trends Inf Retr* 9(2):91–208. <https://doi.org/10.1561/15000000043>
- Kazai G, Kamps J, Milic-Frayling N (2013) An analysis of human factors and label accuracy in crowdsourcing relevance judgments. *Inf Retr* 16(2):138–178. <https://doi.org/10.1007/s10791-012-9205-0>
- Kishida K (2010) Vocabulary-based re-ranking for geographic and temporal searching at NTCIR GeoTime task. In: Kando et al 2010, pp 181–184. <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings8/NTCIR/06-NTCIR8-GeoTime-KishidaK.pdf>
- Mani I, Pustejovsky J, Sundheim B (2004) Introduction to the special issue on temporal information processing. *ACM Trans Asian Lang Inf Process (TALIP)* 3(1):1–10. <https://doi.org/10.1145/1017068.1017069>
- Matthews M, Tolchinsky P, Blanco R, Atserias J, Mika P, Zaragoza H (2010) Searching through time in the New York Times. In: HCIR 2010, pp 41–44. http://www.hugo-zaragoza.net/academic/pdf/matthews_HCIR2010.pdf
- Mitamura T, Nyberg E, Shima H, Kato T, Mori T, Lin C, Song R, Lin C, Sakai T, Ji D, Kando N (2008) Overview of the NTCIR-7 ACLIA tasks: advanced cross-lingual information access. In: Kando N (ed) Proceedings of the 7th NTCIR workshop meeting on evaluation of information access technologies: information retrieval, question answering and cross-lingual information access, NTCIR-7, National Center of Sciences, Tokyo, Japan, 16–19 December 2008, National Institute of Informatics (NII). <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings7/pdf/NTCIR7/C1/CCLQA/01-NTCIR7-OV-CCLQA-MitamuraT.pdf>
- Mori T (2010) A method for GeoTime information retrieval based on question decomposition and question answering. In: Kando et al 2010, pp 167–172. <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings8/NTCIR/04-NTCIR8-GeoTime-MoriT.pdf>
- Moulaoui B, Tamine L, Yahia SB (2016) When time meets information retrieval: past proposals, current plans and future trends. *J Inf Sci* 42(6):725–747. <https://doi.org/10.1177/0165551515607277>
- Purves R, Jones C (2004) Workshop on geographic information retrieval, SIGIR 2004. *SIGIR Forum* 38(2):53–56. <https://doi.org/10.1145/1041394.1041406>
- Rizzo SG, Montesi D (2017) Quantification of time in digital libraries: temporal Zipf’s law. In: Proceedings of the 21st international database engineering and applications symposium, IDEAS 2017, pp 143–152. ACM, New York. <https://doi.org/10.1145/3105831.3105866>
- Sakai T, Shima H, Kando N, Song R, Lin C, Mitamura T, Sugimoto M, Lee C (2010) Overview of NTCIR-8 ACLIA IR4QA. In: Kando et al 2010, pp 63–93. <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings8/NTCIR/01-NTCIR8-OV-IR4QA-SakaiT.pdf>
- Sato T (2011) NTCIR-9 GeoTime at Osaka Kyoiku University - toward automatic extraction of place/time terms. In: Kando et al 2011. <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings9/NTCIR/09-NTCIR9-GEOTIME-SatoT.pdf>
- Sekine S, Isahara H (2000) IREX: IR & IE evaluation project in Japanese. In: Proceedings of the 2nd international conference on language resources and evaluation (LREC’00), European Language Resources Association (ELRA), Athens, Greece. <http://www.lrec-conf.org/proceedings/lrec2000/pdf/27.pdf>
- Strötgen J, Gertz M (2012) Event-centric search and exploration in document collections. In: Proceedings of the 12th ACM/IEEE-CS joint conference on digital libraries, JCDL’12, pp 223–232. ACM, New York. <https://doi.org/10.1145/2232817.2232859>

- Strötgen J, Gertz M (2013) Proximity2-aware ranking for textual, temporal, and geographic queries. In: Proceedings of the 22nd ACM international conference on information and knowledge management, CIKM'13, pp 739–744. ACM, New York. <https://doi.org/10.1145/2505515.2505640>
- Sundheim BM (1995) Overview of results of the MUC-6 evaluation. In: Proceedings of the 6th conference on message understanding, association for computational linguistics, MUC6'95, Stroudsburg, PA, USA, pp 13–31. <https://doi.org/10.3115/1072399.1072402>
- Yoshioka M (2010) On a combination of probabilistic and boolean IR models for GeoTime task. In: Kando et al 2010, pp 154–158. <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings8/NTCIR/02-NTCIR8-GeoTime-YoshiokaM.pdf>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.