

Summarization in Question Answering



Text Summarization

- **Goal:** produce an abridged version of a text that contains information that is important or relevant to a user.
- **Summarization Applications**
 - **outlines or abstracts** of any document, article, etc
 - **summaries** of email threads
 - **action items** from a meeting
 - **simplifying** text by compressing sentences



What to summarize?

Single vs. multiple documents

- **Single-document summarization**
 - Given a single document, produce
 - abstract
 - outline
 - headline
- **Multiple-document summarization**
 - Given a group of documents, produce a gist of the content:
 - a series of news stories on the same event
 - a set of web pages about some topic or question



Query-focused Summarization & Generic Summarization

- Generic summarization:
 - Summarize the content of a document
- Query-focused summarization:
 - summarize a document with respect to an information need expressed in a user query.
 - a kind of complex question answering:
 - Answer a question by summarizing a document that has the information to construct the answer



Summarization for Question Answering: Snippets

- Create **snippets** summarizing a web page for a query
 - Google: 156 characters (about 26 words) plus title and link

This is a screenshot of a Google search interface. At the top, the Google logo is on the left, and a search bar contains the text 'what is die brücke?'. Below the search bar, it says 'Search' and 'About 5,910,000 results (0.28 seconds)'. On the left side, there is a vertical list of filters: 'Everything', 'Images', 'Maps', 'Videos', 'News', 'Shopping', 'Applications', and 'More'. The main content area shows three search results. The first result is titled 'Die Brücke - Wikipedia, the free encyclopedia' with a green link to 'en.wikipedia.org/wiki/Die_Brücke'. The snippet below it says: 'Die Brücke (The Bridge) was a group of German expressionist artists formed in Dresden in 1905, after which the Brücke Museum in Berlin was named. Founding ... You've visited this page 5 times. Last visit: 4/16/12'. The second result is titled 'Die Brücke (film) - Wikipedia, the free encyclopedia' with a green link to 'en.wikipedia.org/wiki/Die_Brücke_(film)'. The snippet says: 'Die Brücke (English: The Bridge) is a 1959 West German film directed by Austrian filmmaker Bernhard Wicki. It is based on the eponymous 1958 novel by ...'. The third result is titled 'Die Brücke - Die Brücke Art' with a green link to 'www.huntfor.com/arhistory/c20th/diebrucke.htm'. The snippet says: 'Die Brücke was the association of artist expressionists from Dresden, Germany. ... Die Brücke made use of a technique that was controlled, intentionally ...'.



Summarization for Question Answering: Multiple documents

Create **answers** to complex questions summarizing multiple documents.

- Instead of giving a snippet for each document
- Create a cohesive answer that combines information from each document

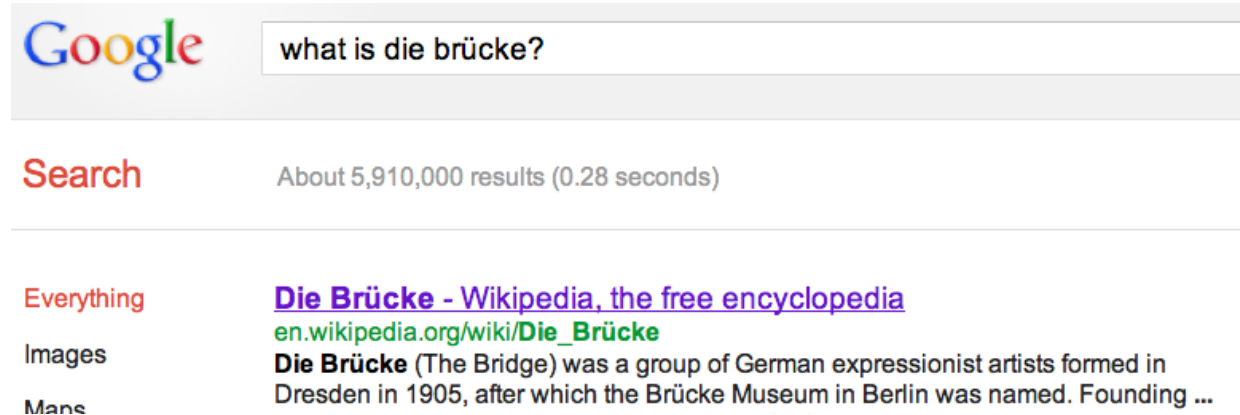


Extractive summarization & Abstractive summarization

- **Extractive summarization:**
 - create the summary from phrases or sentences in the source document(s)
- **Abstractive summarization:**
 - express the ideas in the source documents using (at least in part) different words



Simple baseline: take the first sentence



Die Brücke

From Wikipedia, the free encyclopedia

For other uses, see [Die Brücke \(disambiguation\)](#).

Die Brücke (The Bridge) was a group of [German expressionist](#) artists formed in [Dresden](#) in 1905, after which the [Brücke Museum in Berlin](#) was named. Founding members were [Fritz Bleyl](#), [Erich Heckel](#), [Ernst Ludwig Kirchner](#) and [Karl Schmidt-Rottluff](#). Later members were [Emil Nolde](#), [Max Pechstein](#) and [Otto Mueller](#). The seminal group had a major impact on the evolution of [modern art](#) in the 20th century and the creation of expressionism.^[1]

8

Die Brücke is sometimes compared to the [Fauves](#). Both movements shared interests in [primitivist](#) art. Both

[illegible]

Summarization in Question Answering