Spoken Content Retrieval: Challenges, Techniques and Applications

(Part 5: Accessing Information in Spoken Content)

Gareth J. F. Jones

Centre for Next Generation Localisation
School of Computing, Dublin City University, Dublin, Ireland

Overview

Introduction

- The ultimate usefulness of a Spoken Content Retrieval (SCR) system to users is determined by the user interface.

Introduction

- The ultimate usefulness of a Spoken Content Retrieval (SCR) system to users is determined by the user interface.
- Users must be able to efficiently evaluate these results and identify individual items of interest in order for an SCR system to fulfill its function of satisfying user information needs.

Introduction

- The ultimate usefulness of a Spoken Content Retrieval (SCR) system to users is determined by the user interface.
- Users must be able to efficiently evaluate these results and identify individual items of interest in order for an SCR system to fulfill its function of satisfying user information needs.
- Once identified, the interface must further enable efficient access to information within individual relevant items.

Introduction

- The ultimate usefulness of a Spoken Content Retrieval (SCR) system to users is determined by the user interface.
- Users must be able to efficiently evaluate these results and identify individual items of interest in order for an SCR system to fulfill its function of satisfying user information needs.
- Once identified, the interface must further enable efficient access to information within individual relevant items.
- The interface should make full use of feedback from users in order to refine queries.

Introduction

- In practical terms these overarching general requirements mean that an SCR system should be designed to take into account:

Introduction

- In practical terms these overarching general requirements mean that an SCR system should be designed to take into account:
  - the user

Introduction

- In practical terms these overarching general requirements mean that an SCR system should be designed to take into account:
  - the user
  - the tasks to be undertaken by the user

Introduction

- In practical terms these overarching general requirements mean that an SCR system should be designed to take into account:
  - the user
  - the tasks to be undertaken by the user
  - the environment in which the user will be working

Introduction

- In practical terms these overarching general requirements mean that an SCR system should be designed to take into account:
  - the user
  - the tasks to be undertaken by the user
  - the environment in which the user will be working
  - the hardware which the user will be using

Introduction

Considerations when designing for the user might include:

Introduction

Considerations when designing for the user might include:

- class: amateur $\rightarrow$ professional

Introduction

Considerations when designing for the user might include:

- class: amateur $\rightarrow$ professional
- expertise in the subject area: novice $\rightarrow$ expert

Introduction

Considerations when designing for the user might include:

- class: amateur $\rightarrow$ professional
- expertise in the subject area: novice $\rightarrow$ expert
- experience with SCR systems: none $\rightarrow$ extensive

Introduction

Considerations when designing for the user might include:

- class: amateur $\rightarrow$ professional
- expertise in the subject area: novice $\rightarrow$ expert
- experience with SCR systems: none $\rightarrow$ extensive
- training available: none $\rightarrow$ specialist course

Introduction

Task analysis could include the following:

Introduction

Task analysis could include the following:

- setting
  - e.g. home user, professional

Introduction

Task analysis could include the following:

- setting
  - e.g. home user, professional
- what does the user want as the output?
  - e.g. information, piece of video for reuse

Introduction

Task analysis could include the following:

- setting
  - e.g. home user, professional
- what does the user want as the output?
  - e.g. information, piece of video for reuse
- what will the user want to do or know how to do?
  - e.g. simple text query, complex query using detailed matadata

Introduction

Task analysis could include the following:

- setting
  - e.g. home user, professional
- what does the user want as the output?
  - e.g. information, piece of video for reuse
- what will the user want to do or know how to do?
  - e.g. simple text query, complex query using detailed matadata

Hardware to be used:

- Interaction requirements when using a desktop vs tablet vs smartphone will be very different.

Introduction

- In general terms, the interface components for an SCR system should be designed and constructed taking by making use of principles of human-computer interaction.

Introduction

- In general terms, the interface components for an SCR system should be designed and constructed taking by making use of principles of human-computer interaction.
- In particular, development of SCR interface components should take into consideration good user interface design principles, and be aware of issues in achieving efficient, effective and reliable user interaction with computing systems.

Introduction

- In general terms, the interface components for an SCR system should be designed and constructed taking by making use of principles of human-computer interaction.
- In particular, development of SCR interface components should take into consideration good user interface design principles, and be aware of issues in achieving efficient, effective and reliable user interaction with computing systems.
- A number of good textbooks introduce these topics.

Introduction

- In general terms, the interface components for an SCR system should be designed and constructed taking by making use of principles of human-computer interaction.
- In particular, development of SCR interface components should take into consideration good user interface design principles, and be aware of issues in achieving efficient, effective and reliable user interaction with computing systems.
- A number of good textbooks introduce these topics.
- The examples of existing SCR interface components which follow are just that, *examples*; they provide ideas and pointers, but specific interfaces should be designed for individual applications and their users.

Query Entry

- SCR systems often invite users to enter queries by offering a query box, such as is in widespread use with Web search engines.

Query Entry

- SCR systems often invite users to enter queries by offering a query box, such as is in widespread use with Web search engines.
- The PodCastle system has a simple query entry box. It also offers the user much more information concerning the search system and what it can be used to find.

Query Entry

- SCR systems often invite users to enter queries by offering a query box, such as is in widespread use with Web search engines.
- The PodCastle system has a simple query entry box. It also offers the user much more information concerning the search system and what it can be used to find.
  - For example, the scope of the content indexed (over 100,000 episodes) is explicitly mentioned and links to recommended podcast episodes are provided.

Query Entry

PodCastle is available at the following URL:

`http://en.podcastle.jp/`

Query Entry

- The Google election gadget offered search functionality for political speeches during the 2008 US elections.

Query Entry

- The Google election gadget offered search functionality for political speeches during the 2008 US elections.
- This displayed hints for users.

Query Entry

- The Google election gadget offered search functionality for political speeches during the 2008 US elections.
- This displayed hints for users.
  - It provided information about the scope and the functionality of the system by prompting users with the question: “What did the candidates say?”

Query Entry

- The Google election gadget offered search functionality for political speeches during the 2008 US elections.
- This displayed hints for users.
  - It provided information about the scope and the functionality of the system by prompting users with the question: “What did the candidates say?”
  - There was a choice of constraining the search to “McCain”, “Obama” and “Debates.”
    These choices implicitly supply information about what is present in the content.

Query Entry

- The Google election gadget offered search functionality for political speeches during the 2008 US elections.
- This displayed hints for users.
  - It provided information about the scope and the functionality of the system by prompting users with the question: “What did the candidates say?”
  - There was a choice of constraining the search to “McCain”, “Obama” and “Debates.” These choices implicitly supply information about what is present in the content.
  - If users limit the field of search to a particular category of content, system accuracy and speed can improve.

Query Entry

- The Google election gadget offered search functionality for political speeches during the 2008 US elections.
- This displayed hints for users.
  - It provided information about the scope and the functionality of the system by prompting users with the question: “What did the candidates say?”
  - There was a choice of constraining the search to “McCain”, “Obama” and “Debates.” These choices implicitly supply information about what is present in the content.
  - If users limit the field of search to a particular category of content, system accuracy and speed can improve.
- In general, a trend can be observed towards designing query interfaces that inform and guide the user. This may support users in formulating more effective queries.

Query Entry

- The query input interface of the SpeechFind offers users information about what can be found in the system.

Query Entry

- The query input interface of the SpeechFind offers users information about what can be found in the system.
- SpeechFind displays buttons at the top that show the specific sources of content available in the system and also displays sample content below.

Query Entry

- The query input interface of the SpeechFind offers users information about what can be found in the system.
- SpeechFind displays buttons at the top that show the specific sources of content available in the system and also displays sample content below.
- The observation that longer, richer queries lead to better SCR performance has led to the suggestion that interfaces should encourage users to formulate longer queries.

Query Entry

- The query input interface of the SpeechFind offers users information about what can be found in the system.
- SpeechFind displays buttons at the top that show the specific sources of content available in the system and also displays sample content below.
- The observation that longer, richer queries lead to better SCR performance has led to the suggestion that interfaces should encourage users to formulate longer queries.
- The additional boxes for query entry:

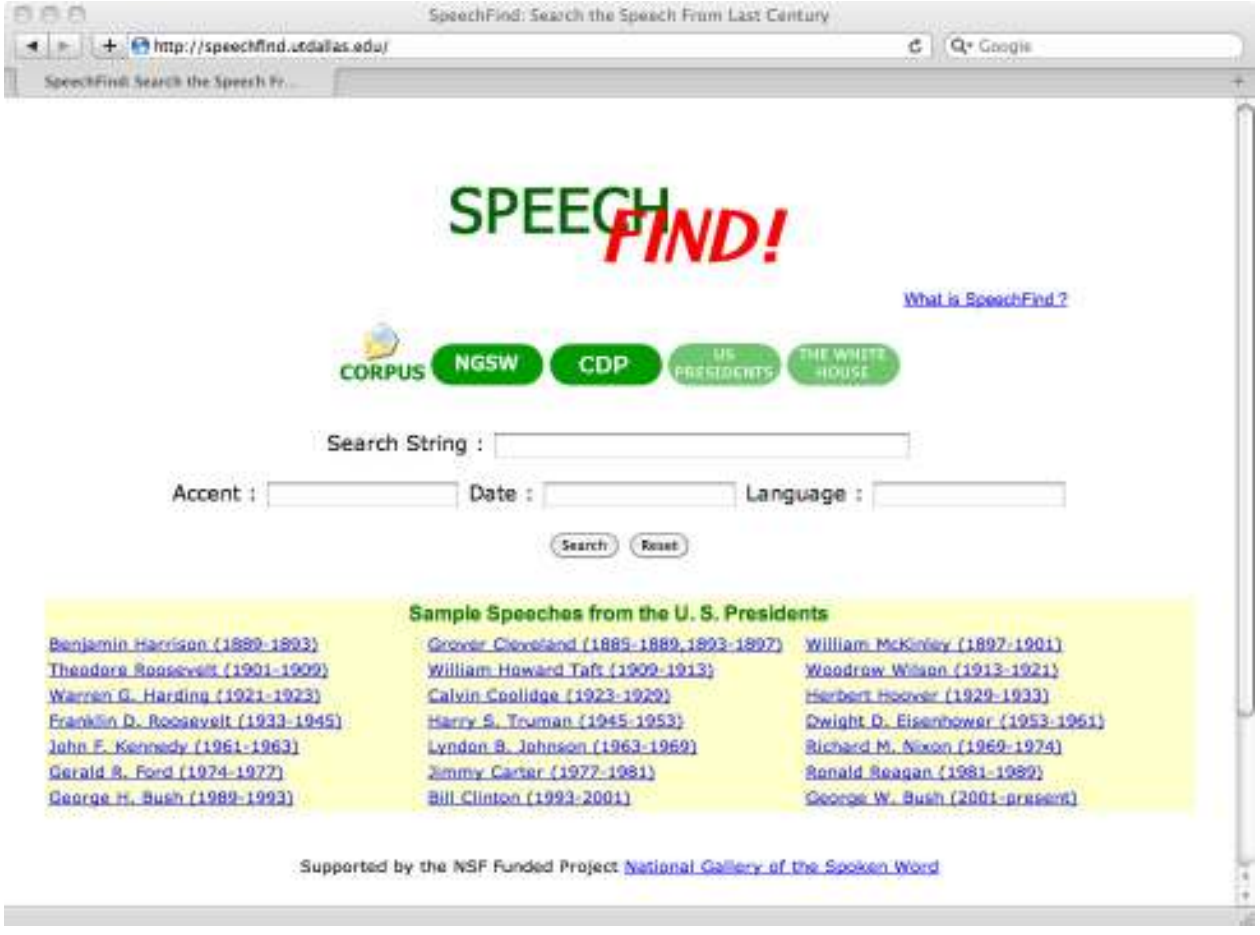
Query Entry

- The query input interface of the SpeechFind offers users information about what can be found in the system.
- SpeechFind displays buttons at the top that show the specific sources of content available in the system and also displays sample content below.
- The observation that longer, richer queries lead to better SCR performance has led to the suggestion that interfaces should encourage users to formulate longer queries.
- The additional boxes for query entry:
  - invite the user to enter information of a specific type;

Query Entry

- The query input interface of the SpeechFind offers users information about what can be found in the system.
- SpeechFind displays buttons at the top that show the specific sources of content available in the system and also displays sample content below.
- The observation that longer, richer queries lead to better SCR performance has led to the suggestion that interfaces should encourage users to formulate longer queries.
- The additional boxes for query entry:
  - invite the user to enter information of a specific type;
  - help the system to disambiguate between query terms that the user expects to hear spoken in the audio content and other characteristics of the speech media, such as the identity of the speaker.

Query Entry


SpeechFind: Search the Speech From Last Century
http://speechfind.utdallas.edu/
SPEECH FIND!
What is SpeechFind ?
CORPUS
NGSW
CDP
US PRESIDENTS
THE WHITE HOUSE
Search String :
Accent :
Date :
Language :
Search
Reset
Sample Speeches from the U. S. Presidents
Benjamin Harrison (1889-1893)
Grover Cleveland (1885-1889,1893-1897)
William McKinley (1897-1901)
Theodore Roosevelt (1901-1909)
William Howard Taft (1909-1913)
Woodrow Wilson (1913-1921)
Warren G. Harding (1921-1923)
Calvin Coolidge (1923-1929)
Herbert Hoover (1929-1933)
Franklin D. Roosevelt (1933-1945)
Harry S. Truman (1945-1953)
Dwight D. Eisenhower (1953-1961)
John F. Kennedy (1961-1963)
Lyndon B. Johnson (1963-1969)
Richard M. Nixon (1969-1974)
Gerald R. Ford (1974-1977)
Jimmy Carter (1977-1981)
Ronald Reagan (1981-1989)
George H. Bush (1989-1993)
Bill Clinton (1993-2001)
George W. Bush (2001-present)
Supported by the NSF Funded Project National Gallery of the Spoken Word

Query Entry

- Ultimately, users are often unwilling or unable to enter a structured query. Thus, systems that automatically differentiate the relationship between query terms and different components of the user information need will prove more effective.

Query Entry

- Ultimately, users are often unwilling or unable to enter a structured query. Thus, systems that automatically differentiate the relationship between query terms and different components of the user information need will prove more effective.
- In addition to text queries, SCR systems can also enable users to specify non-textual queries to initiate the search process.

Query Entry

- Ultimately, users are often unwilling or unable to enter a structured query. Thus, systems that automatically differentiate the relationship between query terms and different components of the user information need will prove more effective.
- In addition to text queries, SCR systems can also enable users to specify non-textual queries to initiate the search process.
  - For example, the Informedia system offers functionality that enables a user to initiate a query by specifying a location on a map. The system returns news stories containing references to places in this region.

Query Entry

- Ultimately, users are often unwilling or unable to enter a structured query. Thus, systems that automatically differentiate the relationship between query terms and different components of the user information need will prove more effective.
- In addition to text queries, SCR systems can also enable users to specify non-textual queries to initiate the search process.
  - For example, the Informedia system offers functionality that enables a user to initiate a query by specifying a location on a map. The system returns news stories containing references to places in this region.
- Other functionality can make it possible to query by visual features such as visual concepts or images.

Display of Results for Selection

- Results returned by the SCR system are generally displayed as a results list.

Display of Results for Selection

- Results returned by the SCR system are generally displayed as a results list.
- It is important that retrieved items are represented in a way that enables users to reliably determine their likely relevance to their information need.

Display of Results for Selection

- The following example shows the ranked retrieval interface from the Video Mail Retrieval Using Voice (VMR) system.

Display of Results for Selection

- The following example shows the ranked retrieval interface from the Video Mail Retrieval Using Voice (VMR) system.
- Individual video messages are shown in the ranked list.

Display of Results for Selection

- The following example shows the ranked retrieval interface from the Video Mail Retrieval Using Voice (VMR) system.
- Individual video messages are shown in the ranked list.
- The horizontal bar lines show the matching score of the message relative to the highest scoring message.

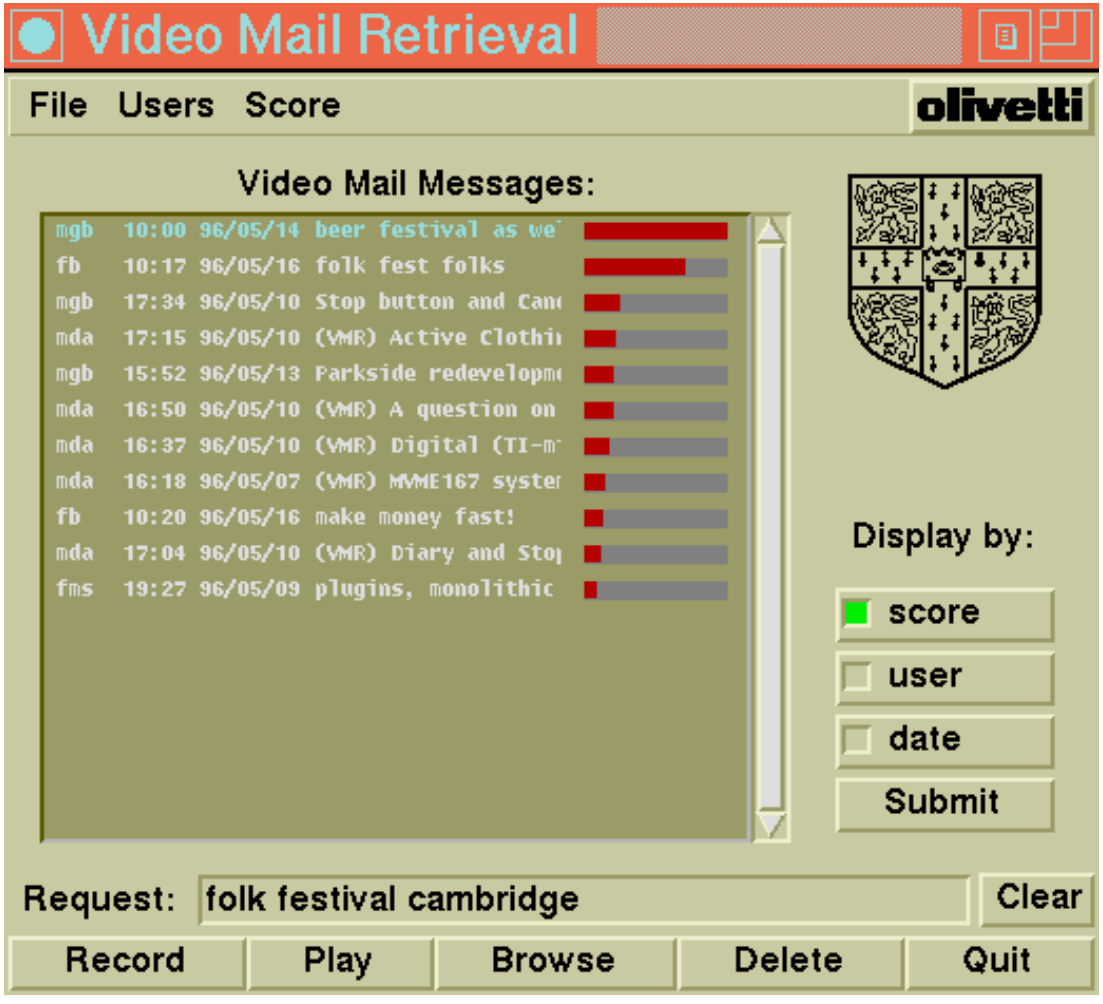
Display of Results for Selection

- The following example shows the ranked retrieval interface from the Video Mail Retrieval Using Voice (VMR) system.
- Individual video messages are shown in the ranked list.
- The horizontal bar lines show the matching score of the message relative to the highest scoring message.
- The search can be narrowed using message dates and identity of the sender.

Display of Results for Selection

- The following example shows the ranked retrieval interface from the Video Mail Retrieval Using Voice (VMR) system.
- Individual video messages are shown in the ranked list.
- The horizontal bar lines show the matching score of the message relative to the highest scoring message.
- The search can be narrowed using message dates and identity of the sender.
  - Such use of metadata to restrict search can be effective to improve efficiency of information access in SCR.

Display of Results for Selection



Display of Results for Selection

- The Broadcast News Retrieval system of the University of Twente illustrates some key issues of results display.

Display of Results for Selection

- The Broadcast News Retrieval system of the University of Twente illustrates some key issues of results display.
  - The example shows results returned in response to the query “Amsterdam.” These are shown as a ranked list of jump-in points each associated with a speech media fragment that the system has matched with the query.

Display of Results for Selection

- The Broadcast News Retrieval system of the University of Twente illustrates some key issues of results display.
  - The example shows results returned in response to the query “Amsterdam.” These are shown as a ranked list of jump-in points each associated with a speech media fragment that the system has matched with the query.
- Each item is represented by a surrogate comprising a short excerpt of the ASR transcript, a keyframe, the name and date of the program and a timecode locating the result within the program.

Display of Results for Selection

- The Broadcast News Retrieval system of the University of Twente illustrates some key issues of results display.
  - The example shows results returned in response to the query “Amsterdam.” These are shown as a ranked list of jump-in points each associated with a speech media fragment that the system has matched with the query.
- Each item is represented by a surrogate comprising a short excerpt of the ASR transcript, a keyframe, the name and date of the program and a timecode locating the result within the program.
  - The surrogate is designed to give the user the information necessary to evaluate the relevance of the result and to decide whether to review it in more depth.

Display of Results for Selection

- The Broadcast News Retrieval system of the University of Twente illustrates some key issues of results display.
  - The example shows results returned in response to the query “Amsterdam.” These are shown as a ranked list of jump-in points each associated with a speech media fragment that the system has matched with the query.
- Each item is represented by a surrogate comprising a short excerpt of the ASR transcript, a keyframe, the name and date of the program and a timecode locating the result within the program.
  - The surrogate is designed to give the user the information necessary to evaluate the relevance of the result and to decide whether to review it in more depth.
  - Surrogates are particularly important for SCR systems, since reviewing a spoken media result requires listening to an audio file or watching a video file.

Display of Results for Selection

- This process of reviewing speech media is considerably more time consuming that skimming a page of text.

Display of Results for Selection

- This process of reviewing speech media is considerably more time consuming that skimming a page of text.
- Text snippets can form effective surrogates, but this is related to the quality of the underlying speech transcripts — it has been observed that the accuracy of surrogates determines their usefulness.

Display of Results for Selection

- This process of reviewing speech media is considerably more time consuming that skimming a page of text.
- Text snippets can form effective surrogates, but this is related to the quality of the underlying speech transcripts — it has been observed that the accuracy of surrogates determines their usefulness.
- Note that in the example, the query word has been highlighted in each snippet. The presence of the query word is strong evidence for the user that the result maybe relevant to their information need.

Display of Results for Selection

- This process of reviewing speech media is considerably more time consuming that skimming a page of text.
- Text snippets can form effective surrogates, but this is related to the quality of the underlying speech transcripts — it has been observed that the accuracy of surrogates determines their usefulness.
- Note that in the example, the query word has been highlighted in each snippet. The presence of the query word is strong evidence for the user that the result maybe relevant to their information need.
- Interface design should take into consideration the user's expectation level of seeing the query word in the snippet and hearing the query word very quickly after the playback of the result is initiated.

Display of Results for Selection

- For video applications, presenting a keyframe may provide the user with an additional hint as to the content of the result.

Display of Results for Selection

- For video applications, presenting a keyframe may provide the user with an additional hint as to the content of the result.
- A further dimension to the selection of the appropriate form of surrogates is the background knowledge of the user, which has been observed to have an impact on the types of surrogates that are preferred.

Display of Results for Selection

- It is important to note how display of results for selection may limit the application of advanced IR techniques.

Display of Results for Selection

- It is important to note how display of results for selection may limit the application of advanced IR techniques.
- Application of techniques such as query expansion may return result items that are relevant to the user's original information need, but where none of the *original* query words are actually uttered in the spoken content.

Display of Results for Selection

- It is important to note how display of results for selection may limit the application of advanced IR techniques.
- Application of techniques such as query expansion may return result items that are relevant to the user's original information need, but where none of the *original* query words are actually uttered in the spoken content.
- Unless there is a mechanism to convince the user of the relevance of such results without showing evidence of the query word being directly associated with the content, users may pass over such results.

Display of Results for Selection

- It is important to note how display of results for selection may limit the application of advanced IR techniques.
- Application of techniques such as query expansion may return result items that are relevant to the user's original information need, but where none of the *original* query words are actually uttered in the spoken content.
- Unless there is a mechanism to convince the user of the relevance of such results without showing evidence of the query word being directly associated with the content, users may pass over such results.
- One of the challenges of displaying a ranked list of results is to effectively communicate to the user the relationship between the results and the structure of the speech media in the underlying collection.

Display of Results for Selection

- Recall multiple levels of units may be used by the SCR system, and that the retrieval unit is not necessarily the only important or useful unit of structure within the collection.

Display of Results for Selection

- Recall multiple levels of units may be used by the SCR system, and that the retrieval unit is not necessarily the only important or useful unit of structure within the collection.
  - For example, as shown in the next figure, there is a tension between the retrieval unit (a fragment) and a larger, natural unit in the collection, e.g. in a news collection - a news item.

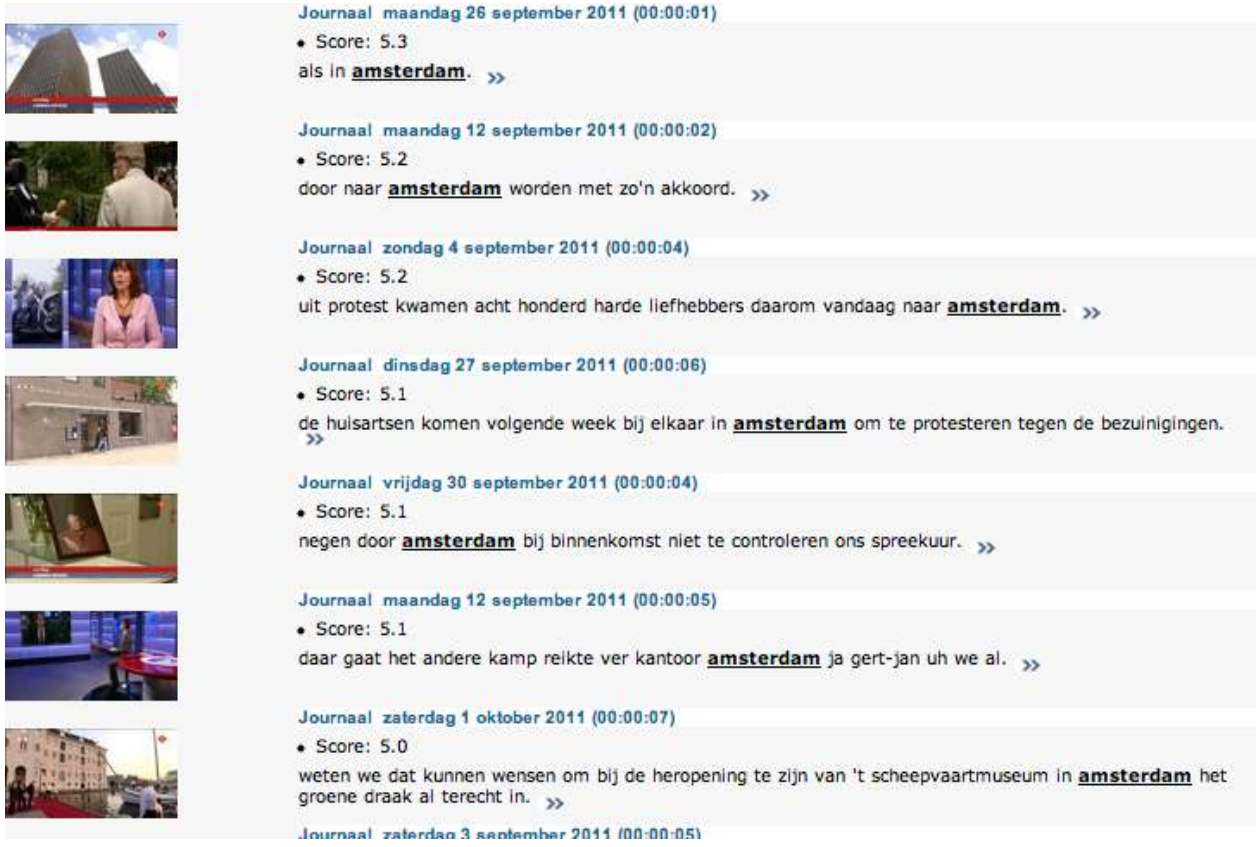
Display of Results for Selection

- Recall multiple levels of units may be used by the SCR system, and that the retrieval unit is not necessarily the only important or useful unit of structure within the collection.
  - For example, as shown in the next figure, there is a tension between the retrieval unit (a fragment) and a larger, natural unit in the collection, e.g. in a news collection - a news item.
  - Two results are returned from the same news program on Monday, 12 September 2011. Depending on the application, two results from the same program might confuse a user, who may consider them actually to constitute a single, duplicated result.

Display of Results for Selection

- Recall multiple levels of units may be used by the SCR system, and that the retrieval unit is not necessarily the only important or useful unit of structure within the collection.
  - For example, as shown in the next figure, there is a tension between the retrieval unit (a fragment) and a larger, natural unit in the collection, e.g. in a news collection - a news item.
  - Two results are returned from the same news program on Monday, 12 September 2011. Depending on the application, two results from the same program might confuse a user, who may consider them actually to constitute a single, duplicated result.
  - Even if they contain different spoken content, it is difficult to indicate the difference clearly, since, as illustrated by this example, displayed results often depend on program level metadata, in this case the date, which is the same.

Display Of Results for Selection



Display of Results for Selection

- A typical approach is to choose the largest unit with which metadata is associated as the retrieval unit.

Display of Results for Selection

- A typical approach is to choose the largest unit with which metadata is associated as the retrieval unit.
- This is ranked and displayed in the results list, e.g. a complete news broadcast or lecture.

Display of Results for Selection

- A typical approach is to choose the largest unit with which metadata is associated as the retrieval unit.
- This is ranked and displayed in the results list, e.g. a complete news broadcast or lecture.
- This approach is taken by PodScope (`http://www.podscope.com`), a spoken-content-based podcast search engine.

Display of Results for Selection

- A typical approach is to choose the largest unit with which metadata is associated as the retrieval unit.
- This is ranked and displayed in the results list, e.g. a complete news broadcast or lecture.
- This approach is taken by PodScope (`http://www.podscope.com`), a spoken-content-based podcast search engine.
- Each result shown is an individual podcast episode displayed together with its metadata.

Display of Results for Selection

- A typical approach is to choose the largest unit with which metadata is associated as the retrieval unit.
- This is ranked and displayed in the results list, e.g. a complete news broadcast or lecture.
- This approach is taken by PodScope (`http://www.podscope.com`), a spoken-content-based podcast search engine.
- Each result shown is an individual podcast episode displayed together with its metadata.
- Information about the relevance of individual fragments within the podcast episode is contained in the surrogate.

Display of Results for Selection

- On the left is a scrolling list of jump-in points, displayed as time codes.

Display of Results for Selection

- On the left is a scrolling list of jump-in points, displayed as time codes.
  - This enables the user to initiate playback of a particular fragment level result directly from the episode-level results list.

Display of Results for Selection

- On the left is a scrolling list of jump-in points, displayed as time codes.
  - This enables the user to initiate playback of a particular fragment level result directly from the episode-level results list.
- Displaying both episodes and fragments together in the results list gives greater flexibility for result review.

Display of Results for Selection

- On the left is a scrolling list of jump-in points, displayed as time codes.
  - This enables the user to initiate playback of a particular fragment level result directly from the episode-level results list.
- Displaying both episodes and fragments together in the results list gives greater flexibility for result review.
- This benefit can be offset by:

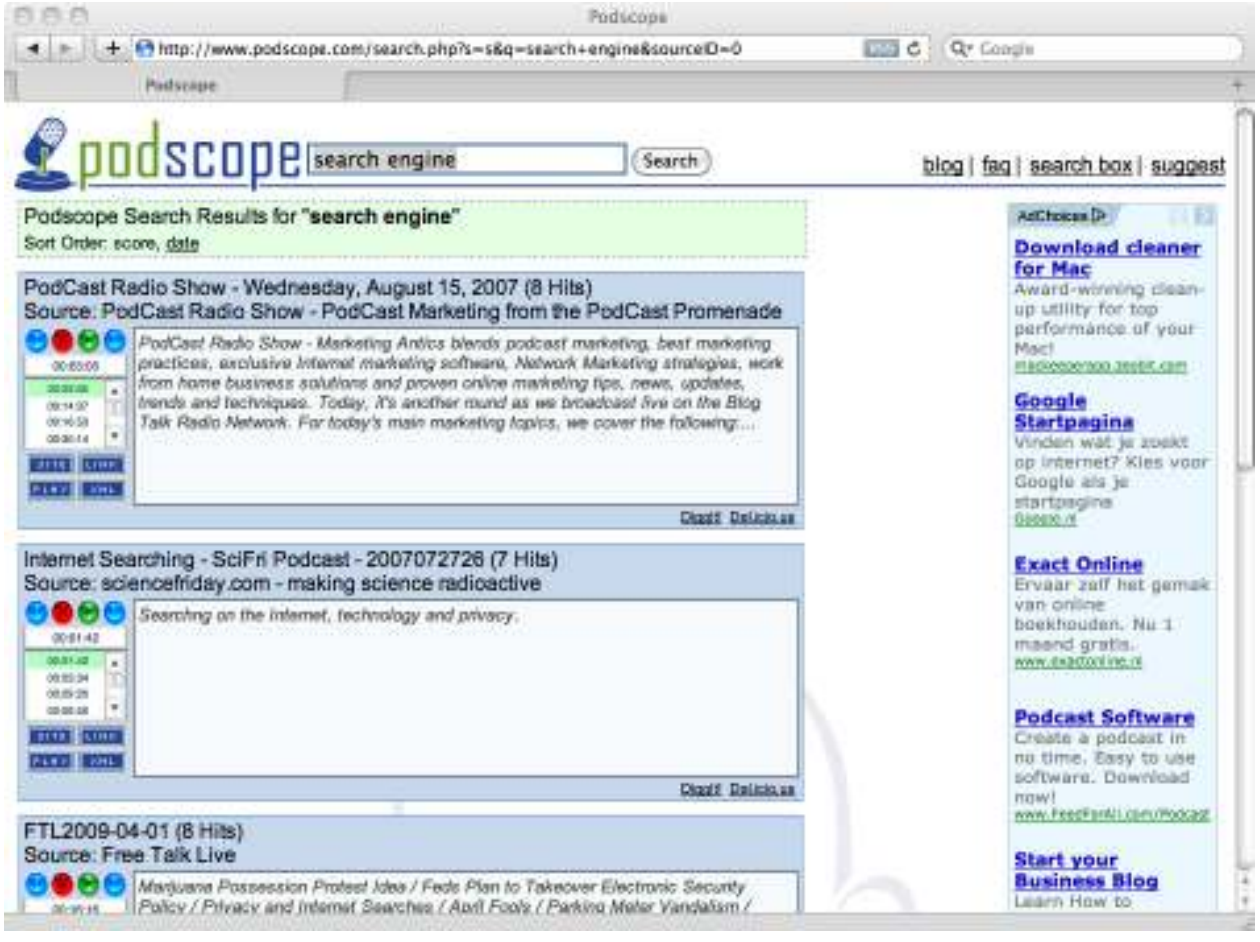
Display of Results for Selection

- On the left is a scrolling list of jump-in points, displayed as time codes.
  - This enables the user to initiate playback of a particular fragment level result directly from the episode-level results list.
- Displaying both episodes and fragments together in the results list gives greater flexibility for result review.
- This benefit can be offset by:
  - the relatively large amount of space required by the surrogate,

Display of Results for Selection

- On the left is a scrolling list of jump-in points, displayed as time codes.
  - This enables the user to initiate playback of a particular fragment level result directly from the episode-level results list.
- Displaying both episodes and fragments together in the results list gives greater flexibility for result review.
- This benefit can be offset by:
  - the relatively large amount of space required by the surrogate,
  - the fact that time codes provide little information to the user about which fragment would be most interesting to select.

Display of Results for Selection



Display Of Results for Selection

- Another approach to results display is used by application developed by the University of Twente for the collection at `http://www.buchenwald.nl` containing interviews with Dutch survivors of the Buchenwald concentration camp.

Display Of Results for Selection

- Another approach to results display is used by application developed by the University of Twente for the collection at `http://www.buchenwald.nl` containing interviews with Dutch survivors of the Buchenwald concentration camp.
- A results list returned in response to the query “bezetting” (Eng. “occupation”) is shown in the example.

Display Of Results for Selection

- Another approach to results display is used by application developed by the University of Twente for the collection at `http://www.buchenwald.nl` containing interviews with Dutch survivors of the Buchenwald concentration camp.
- A results list returned in response to the query “bezetting” (Eng. “occupation”) is shown in the example.
- Each result is an individual interview.

Display Of Results for Selection

- Another approach to results display is used by application developed by the University of Twente for the collection at `http://www.buchenwald.nl` containing interviews with Dutch survivors of the Buchenwald concentration camp.
- A results list returned in response to the query “bezetting” (Eng. “occupation”) is shown in the example.
- Each result is an individual interview.
- The user is not presented with the relevant fragments directly, but rather merely supplied with information on how many there are.

Display Of Results for Selection

- Another approach to results display is used by application developed by the University of Twente for the collection at `http://www.buchenwald.nl` containing interviews with Dutch survivors of the Buchenwald concentration camp.
- A results list returned in response to the query “bezetting” (Eng. “occupation”) is shown in the example.
- Each result is an individual interview.
- The user is not presented with the relevant fragments directly, but rather merely supplied with information on how many there are.
  - (e.g., three fragments is indicated with “3 fragmenten”).

Display Of Results for Selection

- Another approach to results display is used by application developed by the University of Twente for the collection at `http://www.buchenwald.nl` containing interviews with Dutch survivors of the Buchenwald concentration camp.
- A results list returned in response to the query “bezetting” (Eng. “occupation”) is shown in the example.
- Each result is an individual interview.
- The user is not presented with the relevant fragments directly, but rather merely supplied with information on how many there are.
  - (e.g., three fragments is indicated with “3 fragmenten”).
- The user enters an interview in order to explore the fragments that it contains.

Display Of Results for Selection

- In this system. as in earlier examples, segmentation interacts with the IR models used.

Display Of Results for Selection

- In this system. as in earlier examples, segmentation interacts with the IR models used.
- In order to rank interviews, fragment-level and overall interview-level relevance scores are combined.

Display Of Results for Selection

- In this system. as in earlier examples, segmentation interacts with the IR models used.
- In order to rank interviews, fragment-level and overall interview-level relevance scores are combined.
- The exact balance to be used when making this computation is important for a system is important.

Display Of Results for Selection

- In this system. as in earlier examples, segmentation interacts with the IR models used.
- In order to rank interviews, fragment-level and overall interview-level relevance scores are combined.
- The exact balance to be used when making this computation is important for a system is important.
  - For example, in some cases a unit containing a single highly relevant fragment should out-rank a unit containing multiple less relevant fragments.

Display Of Results for Selection

- In this system. as in earlier examples, segmentation interacts with the IR models used.
- In order to rank interviews, fragment-level and overall interview-level relevance scores are combined.
- The exact balance to be used when making this computation is important for a system is important.
  - For example, in some cases a unit containing a single highly relevant fragment should out-rank a unit containing multiple less relevant fragments.
  - In other cases, the opposite could be expected to hold.

Display Of Results for Selection

- In this system. as in earlier examples, segmentation interacts with the IR models used.
- In order to rank interviews, fragment-level and overall interview-level relevance scores are combined.
- The exact balance to be used when making this computation is important for a system is important.
  - For example, in some cases a unit containing a single highly relevant fragment should out-rank a unit containing multiple less relevant fragments.
  - In other cases, the opposite could be expected to hold.
  - This combination itself should depend on how users use the system.

Display Of Results for Selection

Review and Playback of Results

- SCR systems include a player that allows the user to play back individual results.

Review and Playback of Results

- SCR systems include a player that allows the user to play back individual results.
- Playback is an important functionality for both content review and content consumption.

Review and Playback of Results

- SCR systems include a player that allows the user to play back individual results.
- Playback is an important functionality for both content review and content consumption.
- Playback should have the goal of saving the user time and reducing memory load.

Review and Playback of Results

- SCR systems include a player that allows the user to play back individual results.
- Playback is an important functionality for both content review and content consumption.
- Playback should have the goal of saving the user time and reducing memory load.
  - The latter is important since humans generally have poor facilities for extracting and remembering detailed information from audio content.

Review and Playback of Results

- SCR systems include a player that allows the user to play back individual results.
- Playback is an important functionality for both content review and content consumption.
- Playback should have the goal of saving the user time and reducing memory load.
  - The latter is important since humans generally have poor facilities for extracting and remembering detailed information from audio content.
- As a design guideline, players should provide as much information as possible about the content, without clutter, and give users flexible control over navigation.

Review and Playback of Results

- An important part of accessing spoken content is actually listening to audio material, and even more so in applications where little or no vision display of feature relating to the content is possible.

Review and Playback of Results

- An important part of accessing spoken content is actually listening to audio material, and even more so in applications where little or no vision display of feature relating to the content is possible.
- This process can be made more efficient by enabling fast playback.

Review and Playback of Results

- An important part of accessing spoken content is actually listening to audio material, and even more so in applications where little or no vision display of feature relating to the content is possible.
- This process can be made more efficient by enabling fast playback.
  - It has been shown that speech can still be intelligible if the speed of delivery is doubled. However, the cognitive load of listening to speech at this speed is considerably higher than natural speech, leading to the listener rapidly losing focus or becoming overloaded.

Review and Playback of Results

- An important part of accessing spoken content is actually listening to audio material, and even more so in applications where little or no vision display of feature relating to the content is possible.
- This process can be made more efficient by enabling fast playback.
  - It has been shown that speech can still be intelligible if the speed of delivery is doubled. However, the cognitive load of listening to speech at this speed is considerably higher than natural speech, leading to the listener rapidly losing focus or becoming overloaded.
- Other approaches for compressing speech involve not only altering the speech rate, but also removing unimportant words and segments (called excision).

Review and Playback of Results

- Another option available to SCR systems is presenting users with ASR transcripts to read.

Review and Playback of Results

- Another option available to SCR systems is presenting users with ASR transcripts to read.
- Scanmail is an example of a system that presents the user with the ASR transcript directly, in this case a voicemail message.

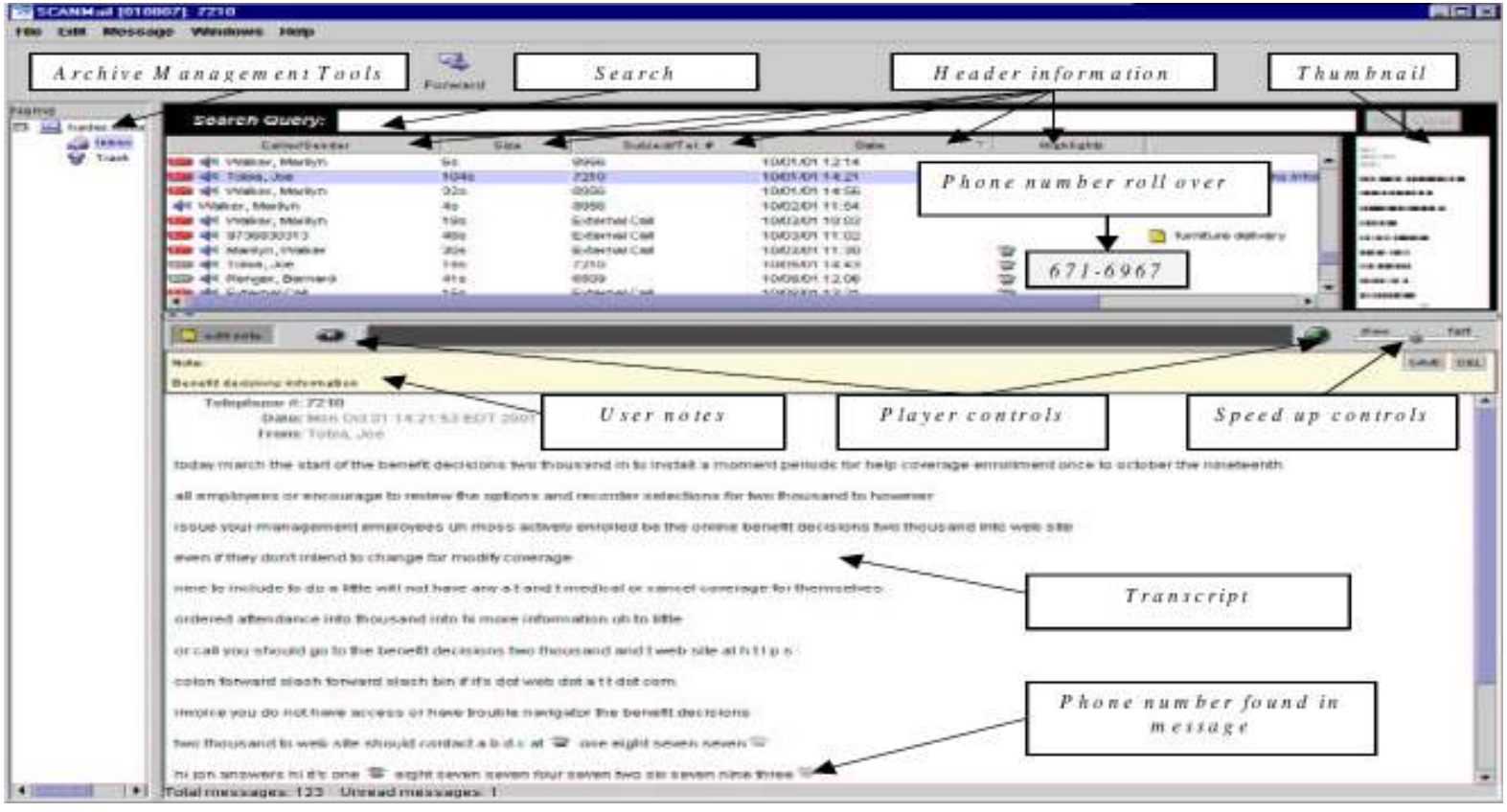
Review and Playback of Results

- Another option available to SCR systems is presenting users with ASR transcripts to read.
- Scanmail is an example of a system that presents the user with the ASR transcript directly, in this case a voicemail message.
- A user study suggested that this feature was appreciated. However, these is a relationship between the error levels of ASR transcripts and their usefulness in the system, with higher error rates being less useful.

Review and Playback of Results

- Another option available to SCR systems is presenting users with ASR transcripts to read.
- Scanmail is an example of a system that presents the user with the ASR transcript directly, in this case a voicemail message.
- A user study suggested that this feature was appreciated. However, these is a relationship between the error levels of ASR transcripts and their usefulness in the system, with higher error rates being less useful.
- A user study on the usefulness and usability of ASR transcripts for a web archive found that:
  - transcripts with WER $>$ 45% were unsatisfactory,
  - while transcripts with WER $<$ 25% were useful and usable.

Review and Playback of Results



ScanMail user interface

Review and Playback of Results

- It is important to keep in mind, that an SCR system must not allow users to develop an unfounded trust in the ASR transcripts.

Review and Playback of Results

- It is important to keep in mind, that an SCR system must not allow users to develop an unfounded trust in the ASR transcripts.
- One study showed that professional users were found to have significant confidence in the SCR system, the transcripts and their own ability to work with them.
- This resulted in the users failing to seek relevant content not explicitly reflected by the transcripts, reducing the recall of their results.

Review and Playback of Results

- It is important to keep in mind, that an SCR system must not allow users to develop an unfounded trust in the ASR transcripts.
- One study showed that professional users were found to have significant confidence in the SCR system, the transcripts and their own ability to work with them.
- This resulted in the users failing to seek relevant content not explicitly reflected by the transcripts, reducing the recall of their results.
- The same effect was reported in user of Scanmail.
- Recall is more critical for voicemail search and misplaced trust in the ASR transcripts caused users in the study to miss crucial information that was not recognized by the ASR system.

Review and Playback of Results

- If ASR transcripts have word- or sentence-level time codes, these can be offered to users in the interface linked to the speech stream so that users can click in the transcript to jump directly to listening to the stream, when they find a portion that interests them.

Review and Playback of Results

- If ASR transcripts have word- or sentence-level time codes, these can be offered to users in the interface linked to the speech stream so that users can click in the transcript to jump directly to listening to the stream, when they find a portion that interests them.
- SCR players typically combine functionality that allows users random access to the speech stream with a tape-recorder metaphor.

Review and Playback of Results

- If ASR transcripts have word- or sentence-level time codes, these can be offered to users in the interface linked to the speech stream so that users can click in the transcript to jump directly to listening to the stream, when they find a portion that interests them.
- SCR players typically combine functionality that allows users random access to the speech stream with a tape-recorder metaphor.
- Interfaces often use a timeline metaphor with time running from left to right, with events positioned along the timeline proportionately to where they occur in the audio file.

Review and Playback of Results

- Many interfaces incorporate a slider bar that enables the user to manipulate the point of playback.

Review and Playback of Results

- Many interfaces incorporate a slider bar that enables the user to manipulate the point of playback.
- The timeline can contain markers indicating the positions of particularly relevant material, e.g. words, phrases.

Review and Playback of Results

- Many interfaces incorporate a slider bar that enables the user to manipulate the point of playback.
- The timeline can contain markers indicating the positions of particularly relevant material, e.g. words, phrases.
- An early example of this functionality is illustrated in the Video Mail Retrieval system browsing interface.

Review and Playback of Results

- Many interfaces incorporate a slider bar that enables the user to manipulate the point of playback.
- The timeline can contain markers indicating the positions of particularly relevant material, e.g. words, phrases.
- An early example of this functionality is illustrated in the Video Mail Retrieval system browsing interface.
- The interface shows a graphical timebar with individual hits on query words in the audio file highlighted. Search term confidence is indicated by the brightness of the results when displayed.

Review and Playback of Results

- Many interfaces incorporate a slider bar that enables the user to manipulate the point of playback.
- The timeline can contain markers indicating the positions of particularly relevant material, e.g. words, phrases.
- An early example of this functionality is illustrated in the Video Mail Retrieval system browsing interface.
- The interface shows a graphical timebar with individual hits on query words in the audio file highlighted. Search term confidence is indicated by the brightness of the results when displayed.
- The user can click to start playback at any point on the timeline. This example is only 20 seconds in length — for longer files, clusters of search term hits can direct the user to regions which are likely to be relevant.

Review and Playback of Results

Review and Playback of Results

- If the speech media is a video, a storyboard of clickable keyframes can be used to depict the temporal progression graphically for the user.

Review and Playback of Results

- If the speech media is a video, a storyboard of clickable keyframes can be used to depict the temporal progression graphically for the user.
- An example of this principle in use is the CMU News-On-Demand system.

Review and Playback of Results

- If the speech media is a video, a storyboard of clickable keyframes can be used to depict the temporal progression graphically for the user.
- An example of this principle in use is the CMU News-On-Demand system.
- This approach is also adopted by many of the video retrieval systems developed for the TRECVid benchmark.

Review and Playback of Results

- The Dutch podcast search engine *Kunststofzuiger* developed at the University of Amsterdam illustrates typical strategies for SCR players.

Review and Playback of Results

- The Dutch podcast search engine *Kunststofzuiger* developed at the University of Amsterdam illustrates typical strategies for SCR players.
- The player page is displayed when the user clicks on a podcast episode presented in a results list in response to a query. In this case “nederland” (Eng. “Netherlands”).

Review and Playback of Results

- The Dutch podcast search engine *Kunststofzuiger* developed at the University of Amsterdam illustrates typical strategies for SCR players.
- The player page is displayed when the user clicks on a podcast episode presented in a results list in response to a query. In this case “nederland” (Eng. “Netherlands”).
- The player page has a query-independent representation of the episode, in the form of the podcast title, broadcast date and description, and also a term cloud that has been extracted from the transcript of the podcast.

Review and Playback of Results

- The Dutch podcast search engine *Kunststofzuiger* developed at the University of Amsterdam illustrates typical strategies for SCR players.
- The player page is displayed when the user clicks on a podcast episode presented in a results list in response to a query. In this case “nederland” (Eng. “Netherlands”).
- The player page has a query-independent representation of the episode, in the form of the podcast title, broadcast date and description, and also a term cloud that has been extracted from the transcript of the podcast.
- It also has a query-biased representation of the episode in the form of the player, which contains markers pointing to the moments within the podcast at which the query word occurs.

Review and Playback of Results

- These markers depict the overall frequency of the query word.

Review and Playback of Results

- These markers depict the overall frequency of the query word.
- Clicking one of the markers moves the user to the point in the speech stream at which the query word is spoken.

Review and Playback of Results

- These markers depict the overall frequency of the query word.
- Clicking one of the markers moves the user to the point in the speech stream at which the query word is spoken.
- Note that playback should begin a few words before the spoken word in order to allow the user to process the speech.

Review and Playback of Results

- These markers depict the overall frequency of the query word.
- Clicking one of the markers moves the user to the point in the speech stream at which the query word is spoken.
- Note that playback should begin a few words before the spoken word in order to allow the user to process the speech.
- The time lag before the spoken word should be approximately constant so that a few interactions with the system will inform the user how long to listen at a particular jump-in point before concluding that point was a false alarm and contained no mention of the keywords.

Review and Playback of Results

- Note the interface of the Kunststofzuiger search engine takes a “finding content” approach to retrieving podcasts.

Review and Playback of Results

- Note the interface of the Kunststofzuiger search engine takes a “finding content” approach to retrieving podcasts.
- Once the content has been located, the occurrence of certain words within the podcast is made visible to the user.

Review and Playback of Results

- Note the interface of the Kunststofzuiger search engine takes a “finding content” approach to retrieving podcasts.
- Once the content has been located, the occurrence of certain words within the podcast is made visible to the user.
- In other words, a “finding mentions” point of view is used to support the user in navigating within the episode.

Review and Playback of Results

Review and Playback of Results

- Two variants on these strategies are:

Review and Playback of Results

- Two variants on these strategies are:
  - A term cloud spread out along the player to give the user a general idea of the topical development over the course of the speech media.

Review and Playback of Results

- Two variants on these strategies are:
  - A term cloud spread out along the player to give the user a general idea of the topical development over the course of the speech media.
  - A heat map display that uses shading or colour to reflect the relative likelihood of a position along the timeline being relevant to a query, rather than showing position of specific words.

Review and Playback of Results

- Two variants on these strategies are:
  - A term cloud spread out along the player to give the user a general idea of the topical development over the course of the speech media.
  - A heat map display that uses shading or colour to reflect the relative likelihood of a position along the timeline being relevant to a query, rather than showing position of specific words.
    - This approach is adopted in the VMR Broadcast News browser.
    - In order to create this representation, the ASR transcript is divided into equal-length segments each of which is scored against the query.

Review and Playback of Results

- The brightness of the shade indicates the strength of the match of each segment, allowing identification of likely relevant regions of the broadcast.

Review and Playback of Results

- The brightness of the shade indicates the strength of the match of each segment, allowing identification of likely relevant regions of the broadcast.
- This example illustrates the typical structure of a news broadcast where the story of interest is mentioned in the headlines and covered in detail in the main broadcast and then appears later when the broadcast headlines are repeated.

Review and Playback of Results

- The player bar of the Radio Oranje application illustrates the use of a magnifying glass metaphor.

Review and Playback of Results

- The player bar of the Radio Oranje application illustrates the use of a magnifying glass metaphor.
- The player displays the entire speech in a timeline, as well as a magnified view showing a window of 45 seconds around the current position of the cursor.

Review and Playback of Results

- The player bar of the Radio Oranje application illustrates the use of a magnifying glass metaphor.
- The player displays the entire speech in a timeline, as well as a magnified view showing a window of 45 seconds around the current position of the cursor.
- Above the play bar, the transcript of the currently-playing segment is displayed, with the query word in bold and a moving underline tracking the progression of the playback.

Review and Playback of Results

- The player bar of the Radio Oranje application illustrates the use of a magnifying glass metaphor.
- The player displays the entire speech in a timeline, as well as a magnified view showing a window of 45 seconds around the current position of the cursor.
- Above the play bar, the transcript of the currently-playing segment is displayed, with the query word in bold and a moving underline tracking the progression of the playback.
- The magnified view makes it possible to also depict segmentation information for the entire program in a relatively compact space without losing detail.

Review and Playback of Results

- Segmentation structure is a key characteristic of speech media used for depiction in players.

Review and Playback of Results

- Segmentation structure is a key characteristic of speech media used for depiction in players.
- Segmentation patterns act as an identifying fingerprint for speech media.

Review and Playback of Results

- Segmentation structure is a key characteristic of speech media used for depiction in players.
- Segmentation patterns act as an identifying fingerprint for speech media.
- A global pattern may serve to implicitly convey the nature of the media to the user, e.g., if it is a conversational interview or a political speech.

Review and Playback of Results

- Displaying segments within the playback interface provides the user with the context of the results being examined; this can be an important aid to interpretation.

Review and Playback of Results

- Displaying segments within the playback interface provides the user with the context of the results being examined; this can be an important aid to interpretation.
- Segmentation provides an alternative to rewinding and fast-forwarding: a user can jump back to the beginning of a segment boundary.

Review and Playback of Results

- Displaying segments within the playback interface provides the user with the context of the results being examined; this can be an important aid to interpretation.
- Segmentation provides an alternative to rewinding and fast-forwarding: a user can jump back to the beginning of a segment boundary.
- Such jumps can be considered to be “intelligent” in so far as the underlying segmentation provides a good representation of useful semantic structure of the speech media.

Review and Playback of Results

- Displaying segments within the playback interface provides the user with the context of the results being examined; this can be an important aid to interpretation.
- Segmentation provides an alternative to rewinding and fast-forwarding: a user can jump back to the beginning of a segment boundary.
- Such jumps can be considered to be “intelligent” in so far as the underlying segmentation provides a good representation of useful semantic structure of the speech media.
- The most appropriate use of segmentation structure will depend on the segmentation information available, its quality and also the types of user needs and tasks the SCR system is designed to support.

Review and Playback of Results

- A large number of segments can be simultaneously displayed on the player interface by adopting a playlist metaphor and listing segments vertically.

Review and Playback of Results

- A large number of segments can be simultaneously displayed on the player interface by adopting a playlist metaphor and listing segments vertically.
- Depending on the use case, space can be conserved by dropping length information and representing each segment as an equal-height line.

Review and Playback of Results

- A large number of segments can be simultaneously displayed on the player interface by adopting a playlist metaphor and listing segments vertically.
- Depending on the use case, space can be conserved by dropping length information and representing each segment as an equal-height line.
- Horizontal layout with tracks is a choice preferred for situations in which multiple overlapping segmentations exist.

Review and Playback of Results

- A large number of segments can be simultaneously displayed on the player interface by adopting a playlist metaphor and listing segments vertically.
- Depending on the use case, space can be conserved by dropping length information and representing each segment as an equal-height line.
- Horizontal layout with tracks is a choice preferred for situations in which multiple overlapping segmentations exist.
  - An early example is the PARC audio browser, which displays separate tracks for announcer, speaker, audience, silence, and applause.

Review and Playback of Results

- In the ideal case, manual segmentation information should be included in the metadata of the spoken content.

Review and Playback of Results

- In the ideal case, manual segmentation information should be included in the metadata of the spoken content.
- Since this is often not the case, automatic methods such as TextTiling must usually be applied to generate segmentation boundary points.

Review and Playback of Results

- In the ideal case, manual segmentation information should be included in the metadata of the spoken content.
- Since this is often not the case, automatic methods such as TextTiling must usually be applied to generate segmentation boundary points.
- Such methods inevitably make errors, dropping real boundary markers in some places and inserting false boundary markers in others.

Review and Playback of Results

- In the ideal case, manual segmentation information should be included in the metadata of the spoken content.
- Since this is often not the case, automatic methods such as TextTiling must usually be applied to generate segmentation boundary points.
- Such methods inevitably make errors, dropping real boundary markers in some places and inserting false boundary markers in others.
- The utility of the browsing interface may potentially be impacted by these errors, particularly if they are numerous or occur at significant points in the semantic flow of the content.

Review and Playback of Results

- Colours in the playback interface can be used to represent speaker characteristics such as gender and speaker age (i.e., child or adult) or even speaker identity or speech/non-speech differences.

Review and Playback of Results

- Colours in the playback interface can be used to represent speaker characteristics such as gender and speaker age (i.e., child or adult) or even speaker identity or speech/non-speech differences.
- The patterns of segment alternation are a potential source of valuable information that can aid the user in the selection process.

Review and Playback of Results

- Colours in the playback interface can be used to represent speaker characteristics such as gender and speaker age (i.e., child or adult) or even speaker identity or speech/non-speech differences.
- The patterns of segment alternation are a potential source of valuable information that can aid the user in the selection process.
- Care must be taken in using this approach, since alternating patterns may be difficult for users to interpret.

Review and Playback of Results

- Colours in the playback interface can be used to represent speaker characteristics such as gender and speaker age (i.e., child or adult) or even speaker identity or speech/non-speech differences.
- The patterns of segment alternation are a potential source of valuable information that can aid the user in the selection process.
- Care must be taken in using this approach, since alternating patterns may be difficult for users to interpret.
- In general, the more immediately obvious it is to the user why the system “chose” the particular object as relevant, the more comfortable the user will feel using the interface.

Review and Playback of Results

- System designers need to pay careful attention to how the player links to the media file.

Review and Playback of Results

- System designers need to pay careful attention to how the player links to the media file.
- The Radio Oranje application was implemented by using ASR to create a forced alignment of human-generated transcripts with spoken content.

Review and Playback of Results

- System designers need to pay careful attention to how the player links to the media file.
- The Radio Oranje application was implemented by using ASR to create a forced alignment of human-generated transcripts with spoken content.
- Several other types of alignment that can be used to improve the ability of an interface to visually represent the spoken content of a speech media item or otherwise support the user's process of reviewing or examining results.

Review and Playback of Results

- Alignment is not just limited to transcripts, but rather for any sort of speech media, if there is a parallel resource containing text available, the two can potentially be aligned with the help of ASR.

Review and Playback of Results

- Alignment is not just limited to transcripts, but rather for any sort of speech media, if there is a parallel resource containing text available, the two can potentially be aligned with the help of ASR.
- Slides of presentations can be aligned with the speech media.

Review and Playback of Results

- Alignment is not just limited to transcripts, but rather for any sort of speech media, if there is a parallel resource containing text available, the two can potentially be aligned with the help of ASR.
- Slides of presentations can be aligned with the speech media.
  - The slides provide structure for the speech stream and can act as surrogates for displaying spoken content in the interface.

Review and Playback of Results

- The IBM WASABI system transcribes broadcast news in real time, analyzes it for named entities and topic, formulates a set of queries, and uses those queries to extract information from other information sources, e.g., newspapers, WWW.
- Excerpts of information are inserted into the news stream in order to provide viewers with background information on the topics treated.

Review and Playback of Results

- The IBM WASABI system transcribes broadcast news in real time, analyzes it for named entities and topic, formulates a set of queries, and uses those queries to extract information from other information sources, e.g., newspapers, WWW.
- Excerpts of information are inserted into the news stream in order to provide viewers with background information on the topics treated.
- A similar system for Dutch-language broadcast news is the InfoLink system.

Review and Playback of Results

- The IBM WASABI system transcribes broadcast news in real time, analyzes it for named entities and topic, formulates a set of queries, and uses those queries to extract information from other information sources, e.g., newspapers, WWW.
- Excerpts of information are inserted into the news stream in order to provide viewers with background information on the topics treated.
- A similar system for Dutch-language broadcast news is the InfoLink system.
- These types of interfaces invite exploration and support the user in browsing activity, to which we now turn in more detail.

Review and Playback of Results

- The browsing interface of the Google election gadget is designed to enable the user to quickly navigate to specific section of a video.

Review and Playback of Results

- The browsing interface of the Google election gadget is designed to enable the user to quickly navigate to specific section of a video.
- Time-aligned textual information is integrated with video playback for targeted in-video navigation.

Review and Playback of Results

- The browsing interface of the Google election gadget is designed to enable the user to quickly navigate to specific section of a video.
- Time-aligned textual information is integrated with video playback for targeted in-video navigation.
- The interface uses the concept of the text snippet and applies to spoken content using ASR transcripts.

Review and Playback of Results

- The browsing interface of the Google election gadget is designed to enable the user to quickly navigate to specific section of a video.
- Time-aligned textual information is integrated with video playback for targeted in-video navigation.
- The interface uses the concept of the text snippet and applies to spoken content using ASR transcripts.
  - These are intended to skimmed and no read carefully.

Review and Playback of Results

- The browsing interface of the Google election gadget is designed to enable the user to quickly navigate to specific section of a video.
- Time-aligned textual information is integrated with video playback for targeted in-video navigation.
- The interface uses the concept of the text snippet and applies to spoken content using ASR transcripts.
  - These are intended to skimmed and no read carefully.
- For rapid navigation the snippets index into the video and provide points to start playback.

Review and Playback of Results

- Markers are placed on the playbar at the start time of each snippet.

Review and Playback of Results

- Markers are placed on the playbar at the start time of each snippet.
- The snippet appears on mouseover of the marker in the playbar.

Review and Playback of Results

- Markers are placed on the playbar at the start time of each snippet.
- The snippet appears on mouseover of the marker in the playbar.
- Clicking on the marker begins playback two seconds before the start time of the snippet.

Review and Playback of Results

- Markers are placed on the playbar at the start time of each snippet.
- The snippet appears on mouseover of the marker in the playbar.
- Clicking on the marker begins playback two seconds before the start time of the snippet.
  - The offset is designed to prevent missing the start of the audio due to potential problems in text/audio alignment.

Review and Playback of Results

Content Browsing

- The search process is only one way in which users seek information, alternatives include methods such as spatial visualization of document collections.

Content Browsing

- The search process is only one way in which users seek information, alternatives include methods such as spatial visualization of document collections.
- An example of an approach that captures the entire document collection in one structure is a Chinese broadcast news prototype, which offers, in addition to a bottom up retrieval functionality, the option of top-down browsing.

Content Browsing

- The search process is only one way in which users seek information, alternatives include methods such as spatial visualization of document collections.
- An example of an approach that captures the entire document collection in one structure is a Chinese broadcast news prototype, which offers, in addition to a bottom up retrieval functionality, the option of top-down browsing.
  - A list of 20 news categories is provided on the portal page. These categories provide an entry point for a 2D topic tree.

Content Browsing

- The search process is only one way in which users seek information, alternatives include methods such as spatial visualization of document collections.
- An example of an approach that captures the entire document collection in one structure is a Chinese broadcast news prototype, which offers, in addition to a bottom up retrieval functionality, the option of top-down browsing.
  - A list of 20 news categories is provided on the portal page. These categories provide an entry point for a 2D topic tree.
  - Clicking on a category reveals a grid representing the latent topical structure of that category. Clicking on a grid cell reveals a finer breakdown of that category.

Content Browsing

- If the user has retrieved an item via a query submitted to the system, a click on this item will reveal its position within the 2D tree structure.

Content Browsing

- If the user has retrieved an item via a query submitted to the system, a click on this item will reveal its position within the 2D tree structure.
- Exploration is supported by a browsing interface that visualizes the structure of items and connections within the collection.

Content Browsing

- If the user has retrieved an item via a query submitted to the system, a click on this item will reveal its position within the 2D tree structure.
- Exploration is supported by a browsing interface that visualizes the structure of items and connections within the collection.
- Browsers that go above and beyond audio and video content itself, for example, by integrating slides and notes, have been designated *artifact* browsers.

Content Browsing

- If the user has retrieved an item via a query submitted to the system, a click on this item will reveal its position within the 2D tree structure.
- Exploration is supported by a browsing interface that visualizes the structure of items and connections within the collection.
- Browsers that go above and beyond audio and video content itself, for example, by integrating slides and notes, have been designated *artifact* browsers.
- The design of suitable interfaces to support interaction is again crucial to the success of such systems.

Content Browsing

Content Browsing

- Note that there is not a hard boundary separating playback interfaces, discussed in the previous subsection, and browser interfaces.

Content Browsing

- Note that there is not a hard boundary separating playback interfaces, discussed in the previous subsection, and browser interfaces.
- The difference lies in the emphasis that browser interfaces put on presenting a complete picture and on supporting discovering.

Content Browsing

- Note that there is not a hard boundary separating playback interfaces, discussed in the previous subsection, and browser interfaces.
- The difference lies in the emphasis that browser interfaces put on presenting a complete picture and on supporting discovering.
- Much research effort on browsing has been devoted to the domain of meetings.

Content Browsing

- Note that there is not a hard boundary separating playback interfaces, discussed in the previous subsection, and browser interfaces.
- The difference lies in the emphasis that browser interfaces put on presenting a complete picture and on supporting discovering.
- Much research effort on browsing has been devoted to the domain of meetings.
- Interfaces can provide access to spoken audio via time specific links to a meeting's agenda, to images made during the meeting, for example of the whiteboard, or to automatically identify elements such as topic, functional category (presentation, discussion, break) or "hot spots".

Content Browsing

- The JFerret meeting browser was developed by the AMI/AMIDA project.

Content Browsing

- The JFerret meeting browser was developed by the AMI/AMIDA project.
- The browser supports the user in finding elements of interest within a recorded meeting.

Content Browsing

- The JFerret meeting browser was developed by the AMI/AMIDA project.
- The browser supports the user in finding elements of interest within a recorded meeting.
- A range of data types is displayed for this purpose, including speaker diarization, linked slides and transcripts. A curve that plots speaker dominance within the meeting is also displayed.

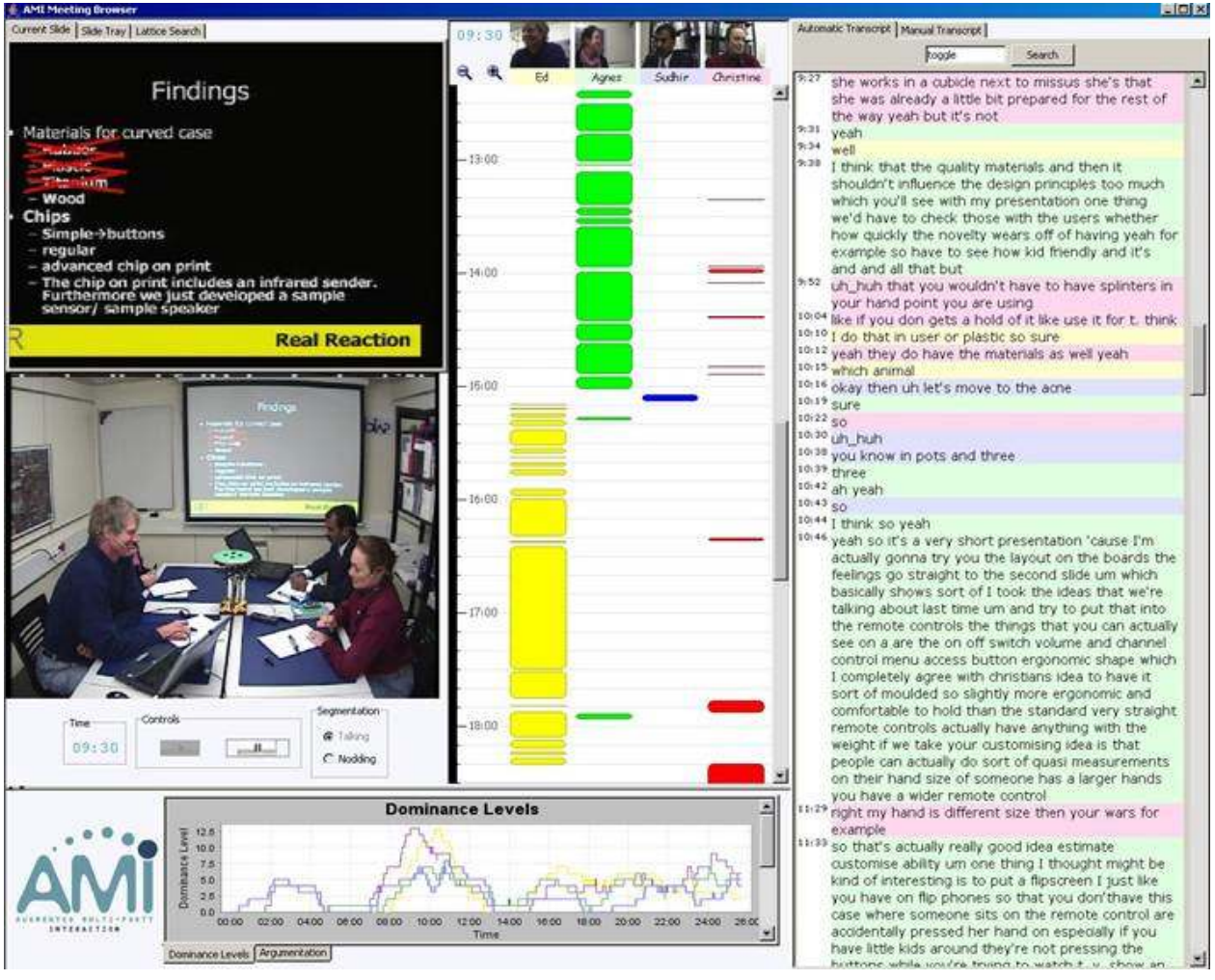
Content Browsing

- The JFerret meeting browser was developed by the AMI/AMIDA project.
- The browser supports the user in finding elements of interest within a recorded meeting.
- A range of data types is displayed for this purpose, including speaker diarization, linked slides and transcripts. A curve that plots speaker dominance within the meeting is also displayed.
- This browser is intended as one part of a large system that to enable users to explore a corpus of meetings.

Content Browsing

- The JFerret meeting browser was developed by the AMI/AMIDA project.
- The browser supports the user in finding elements of interest within a recorded meeting.
- A range of data types is displayed for this purpose, including speaker diarization, linked slides and transcripts. A curve that plots speaker dominance within the meeting is also displayed.
- This browser is intended as one part of a large system that to enable users to explore a corpus of meetings.
  - e.g. to search for the most relevant pieces across meetings to allow the user to answer a specific question.

Content Browsing



JFerret meeting browser