

# Using metadata to combine fragments of audiovisual material

**Lynda Hardman**

<http://www.cwi.nl/~lynda>

CWI, Interactive Information Access  
UvA, Institute for Informatics



# Finding & Combining

- Users need support for
  - the processing of information-bearing content
  - in one or more media types
  - for their specific task
- We need to be aware that there is more than the information “expressed” by the media asset itself, e.g.
  - the intended purpose of the creator
  - the context in which the media asset was created



# Understanding Multimedia Applications Workflow

- Identify and define a number of *canonical processes of media production*
- Community effort
  - 2005: Dagstuhl seminar
  - 2005: ACM MM Workshop on Multimedia for Human Communication
  - 2008: Multimedia Systems Journal Special Issue (core model and companion system papers)  
*editors: Frank Nack, Zeljko Obrenovic and Lynda Hardman*

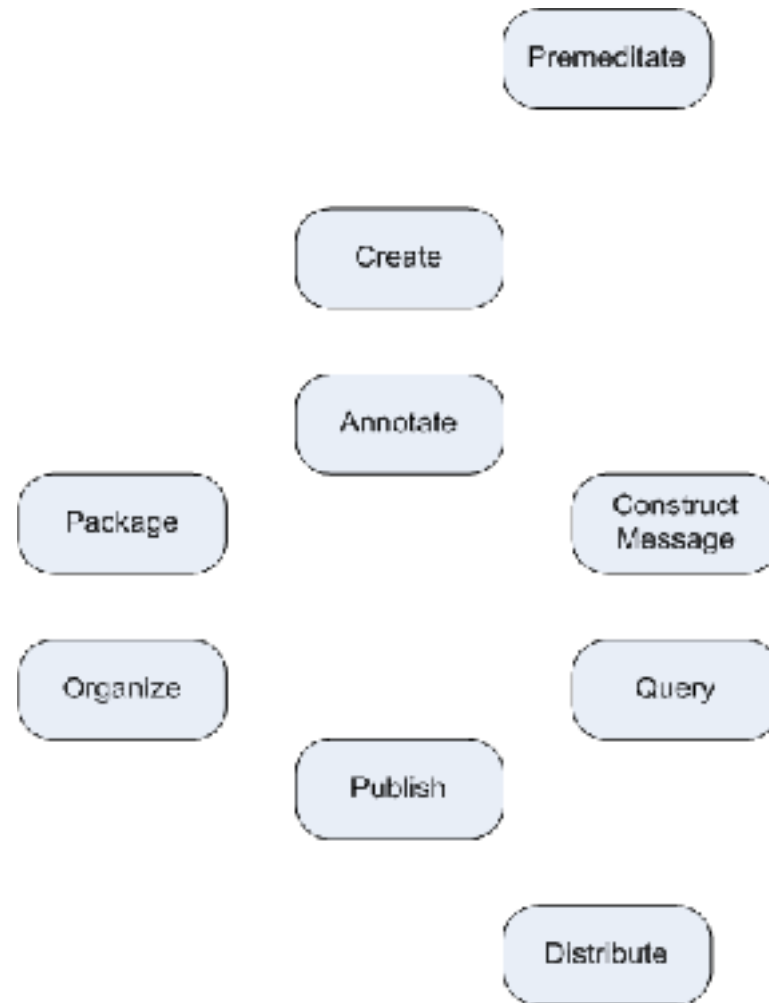


---

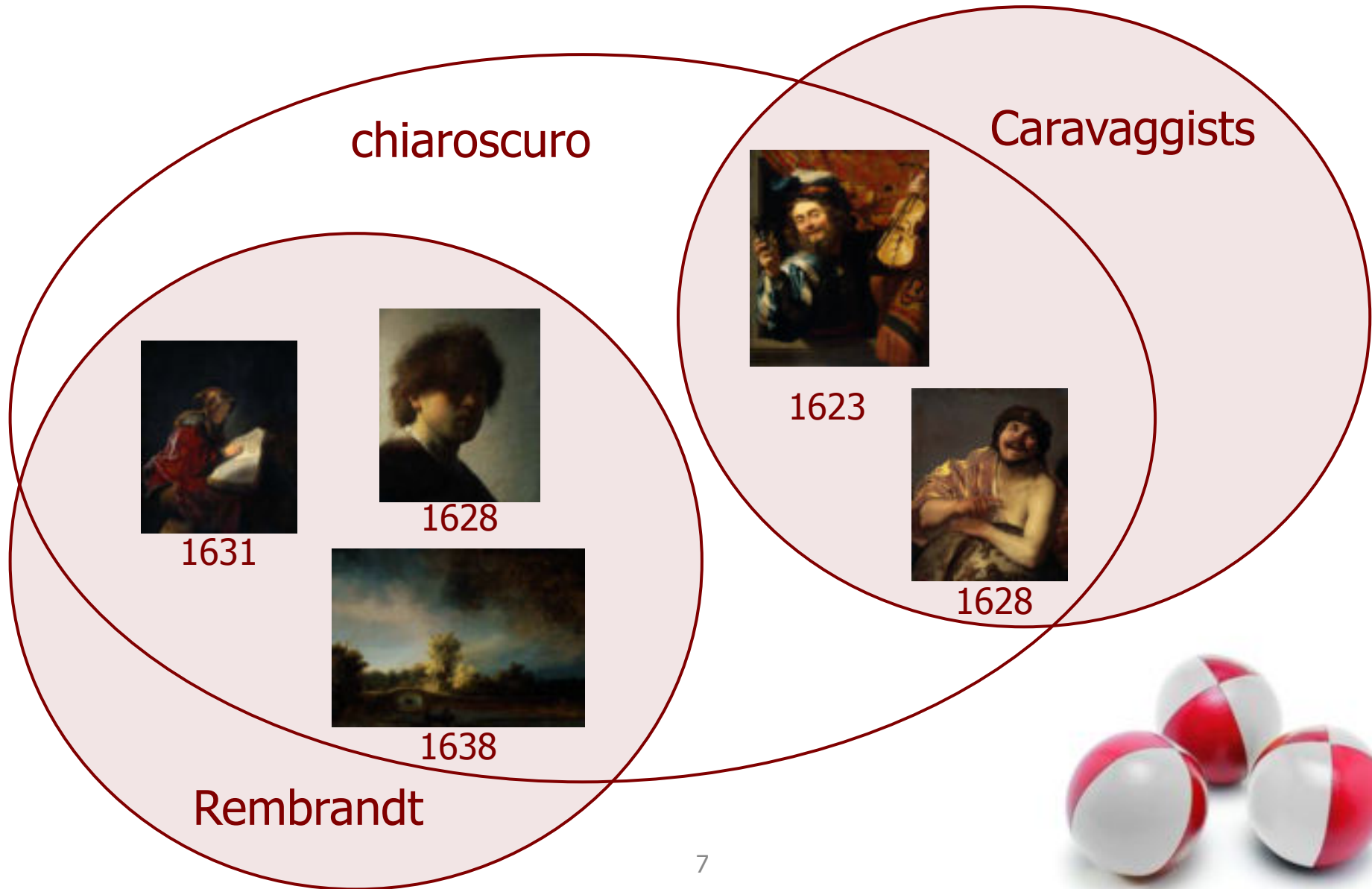
**Canonical** Reduced to the simplest  
and most significant form possible,  
without loss of generality



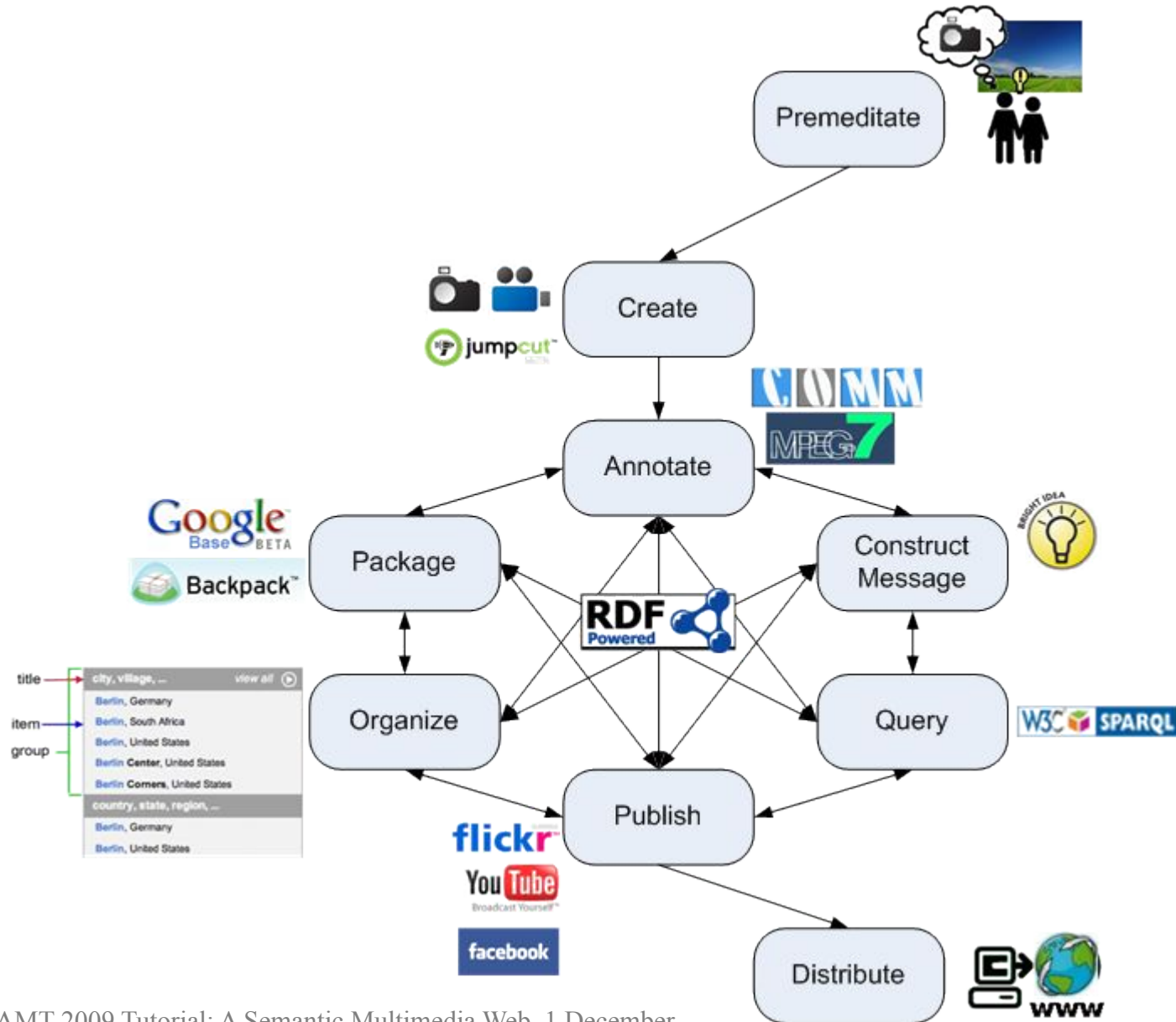
# Overview of Canonical Processes



# Organize using domain annotations



# Canonical Processes





# How can metadata help?

- What can be expressed explicitly?
  - the message to be conveyed
  - objects that are depicted in a media asset
  - domain information (e.g., art, painter)
  - human communication roles (discourse)
  - ...
- What can they be used for?
  - disambiguating queries when selecting media items
  - grouping similar items for conveying topic breadth
  - ordering presentation of items
  - linking to related items
  - visualizing items for presentation, e.g. timeline, map
  - finding similar items (topic or media attributes)
  - finding items that are not quite related
  - ...





**Generating video documentaries  
from annotated media repositories**  
[www.interviewwithamerica.com](http://www.interviewwithamerica.com)

**Stefano Bocconi, Frank Nack (CWI, Amsterdam)**

# Example: What do you think of the war in Afghanistan?



*“I am never a fan of military action, in the big picture I don’t think it is ever a good thing, but I think there are circumstances in which I certainly can’t think of a more effective way to counter this sort of thing...”*

# What do you think of the war in Afghanistan?

**War has never solved anything**

**Two billions dollar bombs on tents**



**I am not a fan of military actions**

**I cannot think of a more effective solution**



# The annotations

- Rhetorical
  - Rhetorical Statement  
(mostly verbal, but visual also possible)
  - Argumentation model: Toulmin model
- Descriptive
  - Question asked
  - Interviewee (social)
  - Filmic *next slide*



# Filmic annotations

Continuity, e.g.

- lighting conditions
- background sound
- gaze direction of speaker
  - left, centre, right
- framing continuity
  - close-up, medium shot, long shot
- camera movement
  - none, pan left/right, shaking,  
tilt up/down, zoom in/out



# Statement encoding

- Statement formally annotated:
  - <subject> <modifier> <predicate>
  - E.g. “**war best solution**”
- A thesaurus containing:
  - Terms (155)
  - Relations between terms: *similar* (72), *opposite* (108), *generalization* (10), *specialization* (10)
  - E.g. **war opposite diplomacy**



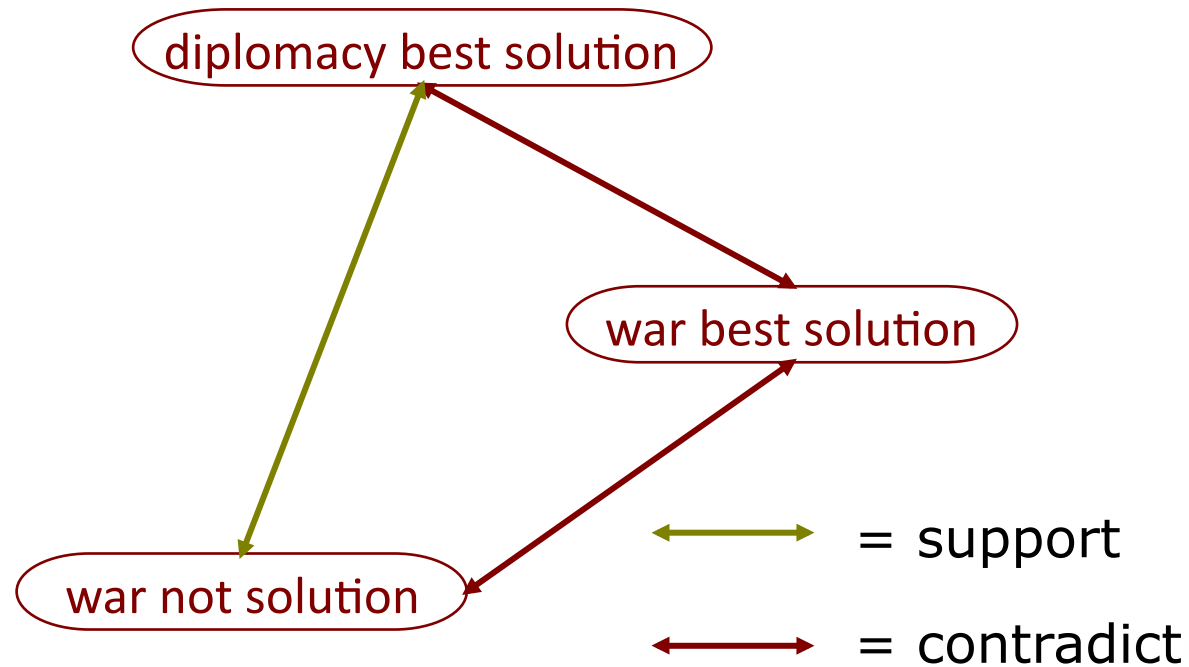
# Connect statements

- Using the thesaurus, generate related statements and query the repository  
*“war best solution”,*  
*“diplomacy best solution”,*  
*“war not solution”*
- Create a **graph** of related statements
  - nodes are the statements  
(corresponding to video segments)
  - edges are either *support* or *contradict*

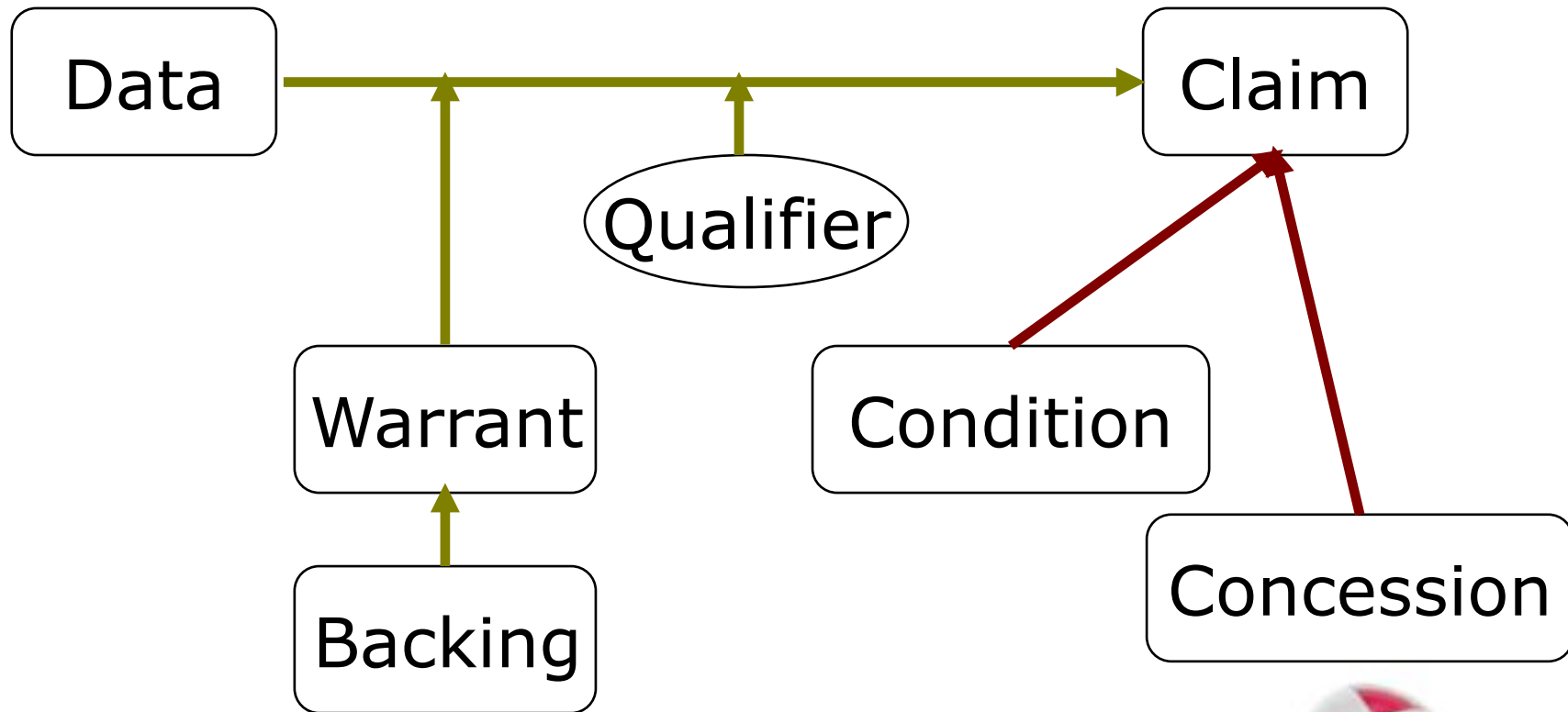




# Semantic Graph



# Toulmin model



*57 Claims, 16 Data, 4 Concessions,  
3 Warrants, 1 Condition*



# Analysis of the Example

**Two billions dollar bombs on tents**



Claim

*contradict*



Claim

**I cannot think of a more effective solution**

*weaken*

Concession

**I am not a fan of military actions**



Claim

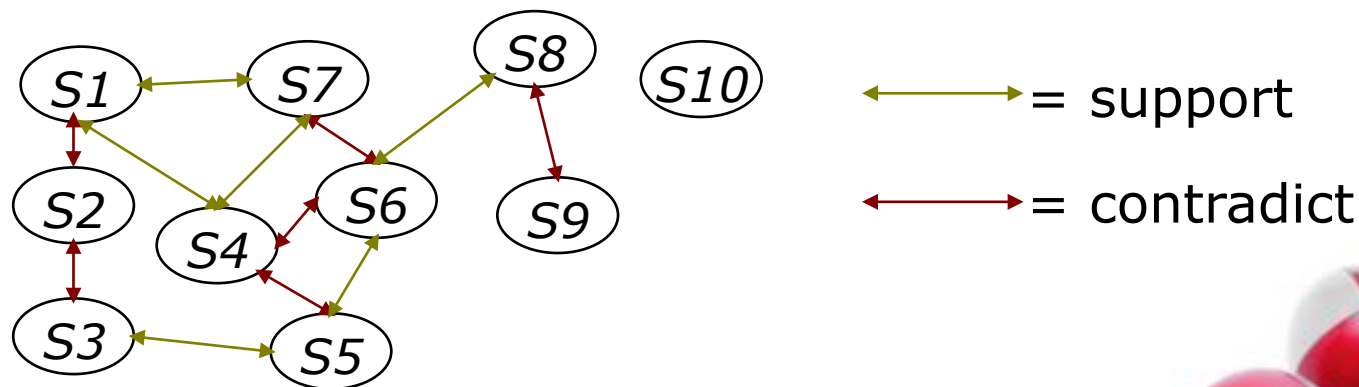
*support*

**War has never solved anything**



# Facts and features

- Annotations: 1 hour annotated, 15 interviews, 60 interview segments, 120 statements
- Partially **tunable**: examining the Segment graph gives feedback on the quality of the annotations and the thesaurus



# Controlling the Bias

- Video documentaries are not neutral account of reality: the selection and editing of the footage expresses a point of view
- Editing strategy:
  - Balanced
  - Pro opinion X
  - Against opinion X
- We use:
  - Logos (the statements)
  - Ethos (based on user profile)
  - Film editing (e.g. framing, gaze)



# Vox Populi interface

Question	Interviewee	Opinion					
Why did they do what they did? What do you think of the casualties among civilians? What do you think of the Afghanistan war? What are the consequences of the war? What are the roots of the problem? What do you think about the Anthrax?	Cameron Parking Guard at Standard Lawyer in Harvard	War in Afghanistan - Pro	Position				
Age	Education	Employment	GeoLocation	Race	Religion	Sex	First Character
Middleage Old Teenager Young	HighEducated LowEducated MediumEducated	HighIncomeJob LowIncomeJob MiddleIncomeJob Retired Student	NotUSA USA	AmericanIndian Asian Black Hispanic White	Atheist Christian Muslim	Female Male	
Age	Education	Employment	GeoLocation	Race	Religion	Sex	Second Character
Middleage Old Teenager Young	HighEducated LowEducated MediumEducated	HighIncomeJob LowIncomeJob MiddleIncomeJob Retired Student	NotUSA USA	AmericanIndian Asian Black Hispanic White	Atheist Christian Muslim	Female Male	
Strategy	Bandwidth	Intercut	Caption				
<input type="radio"/> None <input checked="" type="radio"/> Create Clash <input type="radio"/> Create Support <input type="radio"/> Vox Populi	<input type="radio"/> Low Bandwidth <input checked="" type="radio"/> Medium Bandwidth <input type="radio"/> High Bandwidth	<input checked="" type="radio"/> True <input type="radio"/> False	<input type="radio"/> On (can cause problems) <input checked="" type="radio"/> Off				
Done	Reset						

# Conclusions

- Automatic generation of video interviews augmented with supporting and/or contradicting material
- The **user** can determine the subject and the bias of the presentation
- The **documentarist** can add material and let the system generate new documentaries



# Pointers & Acknowledgments

- Demo available at:

<http://www.cwi.nl/~media/demo/VoxPopuli/>

- VoxPopuli research was funded by the Dutch national ToKeN I<sup>2</sup>RP and CHIME projects





# Acknowledgements



Jacco van  
Ossenbruggen



Frank Nack



Raphaël Troncy



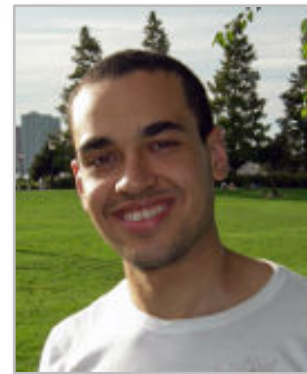
Stefano  
Bocconi



Alia Amin



Michiel  
Hildebrand



Andre Fialho