

Automatically Extracting Sentences from Medline Citations to Support Clinicians' Information Needs

Siddhartha Jonnalagadda, PhD

Research Associate, NLP program

Division of Biomedical Statistics and Informatics,
Department of Health Sciences Research

Poll Question #1

- What is your primary role in VA? **“select all that apply”**
 - student, trainee, or fellow
 - Clinician or clinical researcher
 - informatics researcher
 - Programmer or developer
 - Other

Up to five answer options. Can be “select one” or “select all that apply”

Poll Question #2

- Which best describes your research experience? “select one”
 - have not done research
 - have collaborated on research
 - have conducted research myself
 - have applied for research funding
 - have led a funded research grant

Poll Question #3

- Do you know about information needs at the point of care? “select one”
 - Yes
 - No

Poll Question #3

- Would you be willing to participate or help in recruiting for an international survey on information needs? “select one”
 - Yes
 - No

Team

Siddhartha Jonnalagadda, PhD¹; Guilherme Del Fiol, MD, PhD²; Richard Medlin, MD³; Charlene Weir, RN, PhD²; Marcelo Fiszman, MD, PhD⁴; Javed Mostafa, PhD³; Hongfang Liu, PhD¹

¹Department of Health Sciences Research, Mayo Clinic, Rochester, MN;

²Department of Biomedical Informatics, University of Utah, Salt Lake City, UT;

³Information and Library Science, University of North Carolina, Chapel Hill, NC;

⁴Lister Hill Center, National Library of Medicine, Bethesda, MD



The need to tackle information needs

- Two questions for every three patients and 70% unanswered
- Online health knowledge resources provide answers
- Finding relevant information incompatible with the busy clinical workflow
- Promising approach – automatic extraction and summarization

Major question types

Based on Ely's taxonomy, questions are typically classified as:

- **Treatment: 43.7%**
 - Drug (dosage, efficacy, prevention, etc)
 - Other treatment
- **Diagnosis: 37.7%**
 - finding (sign/symptom/test) → condition
 - condition → finding (list, efficacy, accuracy, etc.)
 - name finding, orientation, etc.
- **Management: 9.2%**
- **Epidemiology: 5.9%**
- **Non-clinical: 3.8%**

A Brief History of Modern Question Answering: From Google to Watson

- Google – provides a list of documents that contain an answer

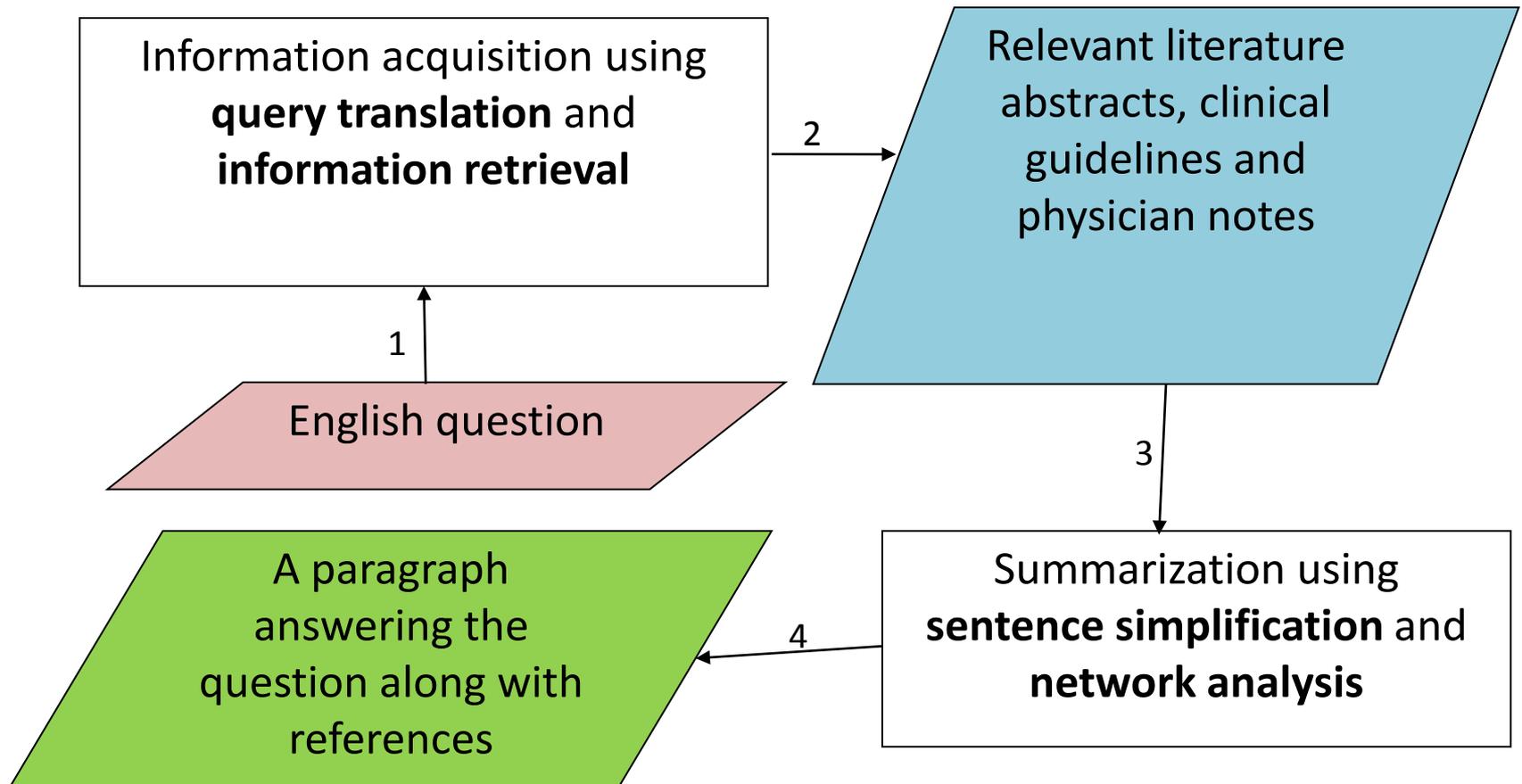
I am a deeply religious person. I believe in the One that is Omniscient. Others call Her Google. My Existential Dilemma is if something cannot be found with Google, does it really exist?

- Wolfram | Alpha – answers factual queries directly by computing the answer from structured data
- Watson – answers open domain questions based on information present in documents as well as databases

Medical Question Answering

- IBM Watson's Jeopardy! being adapted for Doctor's Dilemma.
- AskHermes – a proto-type online medical question answering system
- Medline Plus – Reliable, up-to-date health information, anytime, anywhere for public
- AskMayoExpert – Physician aided shared decision for evidence based care
- MedKS, MiPACQ, Infobot

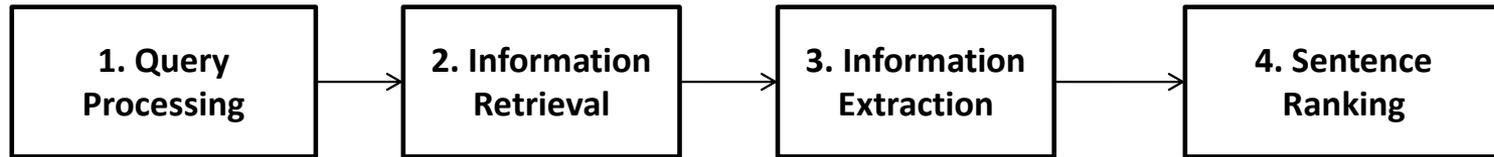
“Is Drug D safe for a patient with problem P, if he is positive on test T and has history of undergoing treatment Tr?”



Examples of Different sources

- **Medline abstracts**
- Full- text of these abstracts
 - Pubmed Central (1 year embargo)
 - Elsevier, Oxford, etc.
- <http://text.soe.ucsc.edu/progress.html>
- Medical Text books
- Wikipedia
- Medpedia

System Architecture



Query –
Alzheimer's

MeSH –
Alzheimer
Disease

UMLS -
C0002395_T04
7

Eutils –
(systematic[sb]
OR
Therapy/Narrow[
filter])
AND Alzheimer
Disease
[MeSH]

PMIDS –
10026388,
10066203, and
1334 more

SemRep:

Treatment name	#
Donepezil	45
Cholinesterase Inhibitors	38
Galantamine	34
rivastigmine	29
Memantine	29
Intervention regimes	9
Vitamin E	8
... and 90 more	

Top 4 sentences:

- Evidence that free radicals may contribute to the pathological processes of cognitive impairment including alzheimer's disease (ad) has led to interest in the use of vitamin e in the treatment of alzheimer's disease and mild cognitive impairment (mci). (PMID: 18646084)
- Finally, a randomized controlled trial questions the clinical benefit of atypical neuroleptics in alzheimer's disease and a comprehensive review of pharmacological trials in mild cognitive impairment reports no benefit of any of the tested drugs on conversion rate to alzheimer's disease. (PMID: 17354654)
- Efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of alzheimer's disease: a systematic review and meta-analysis. (PMID: 18686744)
- A long-term comparison of galantamine and donepezil in the treatment of alzheimer's disease. (PMID: 12875613)

Query Processing

- Tokenization
- Lexical Normalization
- Dictionary Lookup
- Semantic groups
 - treatment
 - disorder
- Screened concepts mapped to MeSH

Algorithm 1: Information retrieval strategy to retrieve abstracts relevant to a particular treatment topic

```
pmidSET ← []
if topic not treatment type then
  goto END
end if

MIN PMIDS = 100
join = AND
BEGIN:
query = ''
for each concept in the topic search do
  if concept is disorder or treatment then
    query = query + join + mesh-form(concept) [MeSH]
  end if
end for

BEGIN1:
pmidSET ← eutils(systematic[sb] AND query)
pmidSET ← eutils(Therapy/Narrow[filter] AND query)
if pmidSET.size() < MIN PMIDS then
  pmidSET ← eutils(Therapy/Broad[filter] AND query)
end if

if pmidSET.size() < MIN PMIDS && join=AND then
  join = OR
  goto BEGIN
end if

if pmidSET.size() < MIN PMIDS then
  query = topic
  goto BEGIN1
end if

END:
```

Algorithm 2: Information extraction method to retrieve semantic predications relevant to a particular treatment topic

```
if topic not treatment type then
  goto END
end if

subjects ← []
objects ← []
for each concept in the topic search do
  if concept is disorder then
    objects ← objects + concept
  else if concept is treatment then
    subjects ← subjects + concept
  end if
end for

if subjects.size()==0 && objects.size()> 0 then
  return predications whose object's CUI is one of the objects' CUIs

else if subjects.size()> 0 && objects.size()> 0 then
  answers ← predications whose object's CUI is one of the objects' CUIs AND
subject's CUI is one of the subjects' CUIs
  if answers.size() > MIN_ANSWERS then
    return answers
  end if
  return answers + predications whose object's CUI is one of the objects' CUIs OR
subject's CUI is one of the subjects' CUIs

else if subjects.size()> 0 then
  return predications whose subject's CUI is one of the subjects' CUIs

else then
  return predications whose object's name is one of the objects' UMLS preferred
```

Sentence Ranking

- Exclude sections (i.e., *objectives*, *selection criteria*, and *methods*) that typically do not contain background statements or study conclusions
- Adapted the TextRank algorithm to rank the sentences retrieved
 - each unique sentence is a vertex in a graph
 - each pair of sentences is connected with an edge whose weight is determined by the cosine word similarity between the sentences
- This approach allows us to not only take into account the similarity between the query and the sentence, but also that among the individual sentences

Case Study Evaluation

- Treatment of two conditions:
 - *depression* and
 - *Alzheimer's disease*
- Topics selected after the system was developed

Attributes of the sentences assessed

- *Topic-relevant* sentences describe one or more treatment alternatives for the condition of interest
- *Conclusive* sentences comprise a statement about one or more treatment alternatives for the condition of interest, either as background information (e.g., current state of knowledge) or study conclusion
- *Comparative* sentences contrast two or more treatment approaches for the condition of interest
- *Contextually-constrained* sentences include specific clinical situations in which a treatment alternative is applicable

Examples

(1) **Not topic-relevant:** “There is insufficient randomized evidence to support the routine use of antidepressants for the prevention of depression or to improve recovery from stroke.” (Pubmed ID: 15802637)

(2) **Conclusive and contextually-constrained:** “There is marginal evidence to support the use of tricyclic antidepressants in the treatment of depression in adolescents, although the magnitude of effect is likely to be moderate at best.” (Pubmed ID: 10908557)

(3) **Comparative:** “Escitalopram versus other antidepressive agents for depression.” (Pubmed ID: 19370639)

(4) **Conclusive and comparative:** “We found no strong evidence that fluvoxamine was either superior or inferior to any other antidepressants in terms of efficacy and tolerability in the acute phase treatment of depression.” (Pubmed ID: 20238342)

(5) **Not topic-relevant:** “Observational studies suggest that some preventive approaches, such as healthy lifestyle, ongoing education, regular physical activity, and cholesterol control, play a role in prevention of AD.” (Pubmed ID: 16529393)

Case study ratings

(kappa = 0.78)	Depression / N	Alzheimer's / N
Relevant abstract	92.6%	84.5%
Relevant sentence	95.6%	88.4%
Conclusive	31.2%	35.0%
Comparative	17.9%	4.4%
Contextually-constrained	44.0%	6.4%

Results contd...

- In both case studies, most of the non-relevant sentences were related to the condition of interest, but the focus was on the diagnosis or prevention [sentences (1) and (5)]
- In both case studies, the TextRank probability was not associated significantly with relevant or conclusive
- Conclusive sentences were located closer to the end of the abstract than non-conclusive sentences (0.51 vs. 0.26; $p < 0.00001$)

Discussion

- High rate of relevant sentences
- Only about one third of the sentences retrieved included a conclusive statement
- Conclusive sentences were located much closer to the end of the abstract than non-conclusive sentences. The percentage of structured abstracts (with conclusion section) in Medline increased from 2.4% in 1992 to 20.3% in 2005
- Sentences with treatment and comparative predications (e.g., *treatment A HIGHER_THAN treatment B*) may be more likely to be conclusive sentences
- A small percentage of the sentences retrieved by the system compared treatment alternatives
- Although almost half of the sentences in the depression case study contained contextual constraints, a much smaller fraction of these sentences were retrieved for Alzheimer's disease

Limitations and Future work

- Generalizability of our findings
- Further studies are needed to identify a threshold that achieves optimal recall for IR
- Broadening from treatment to other predications
- Full-fledged measurement study for overall recall
- Facilitate a tighter definition of relevancy

Conclusion

For the two case studies, the **system retrieved a high percentage of topic-relevant sentences**. Yet, a smaller percentage of sentences were **conclusive, comparative, or contextually constrained**.

Overall, this **seems to be a feasible approach** to constructing context-specific semantic knowledge summaries to support clinicians' patient care decision-making.

K99/R00 Research Plan

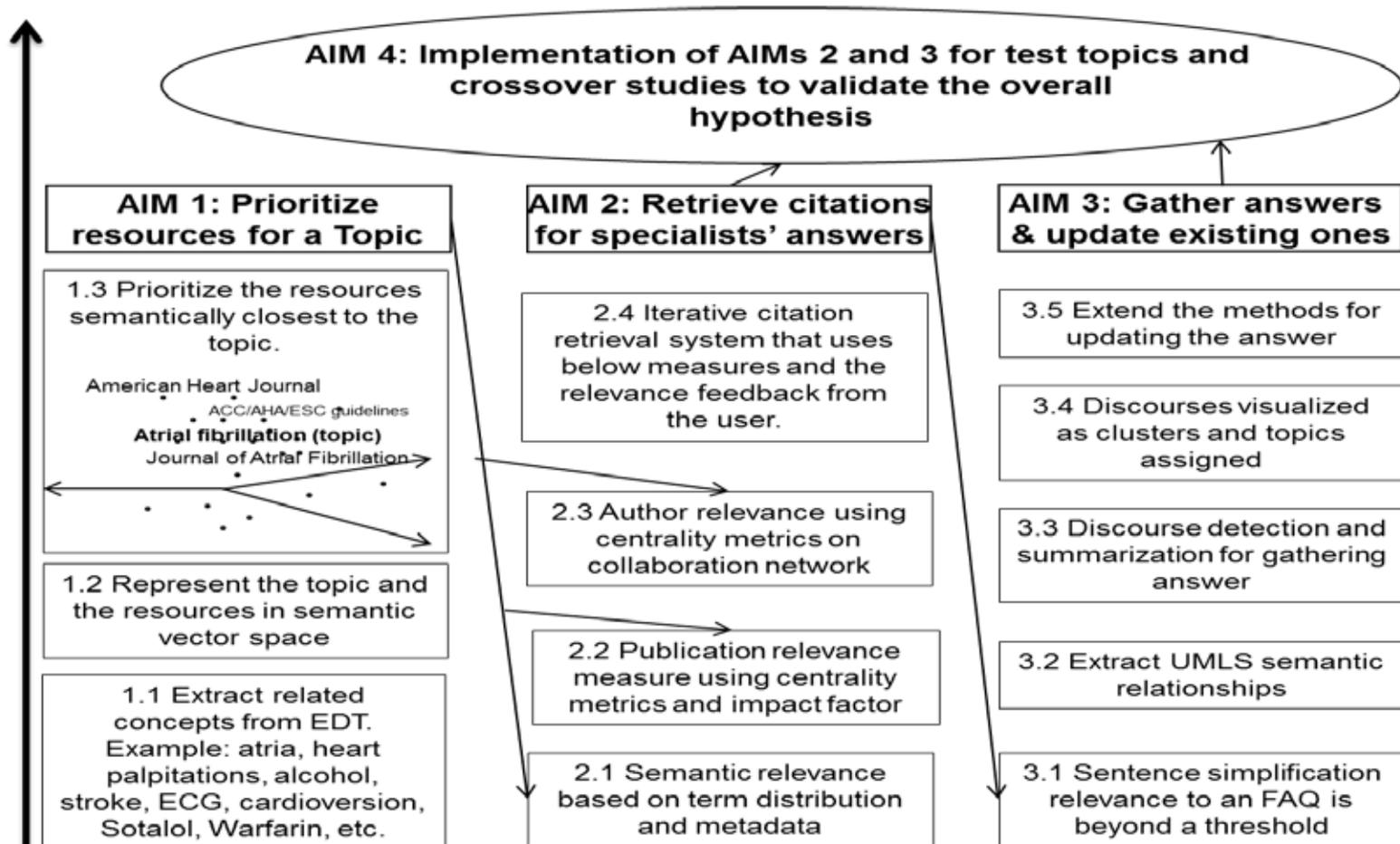


Figure 1: Overview of the methods proposed

Acknowledgements

- The two anonymous reviewers for the highly encouraging comments
- JAMIA for its rapid publication process
- Agency for Healthcare Research and Quality for supporting Del Fiol through grant number K01HS018352.
- NLM for providing access to SemMedDB

Questions?

Corresponding author:
siddhartha@mayo.edu