

# **Interactive Healthcare Information Retrieval for Aging Population**

**Mingkun Gao**

**University of Illinois at Urbana-Champaign**

# Overview

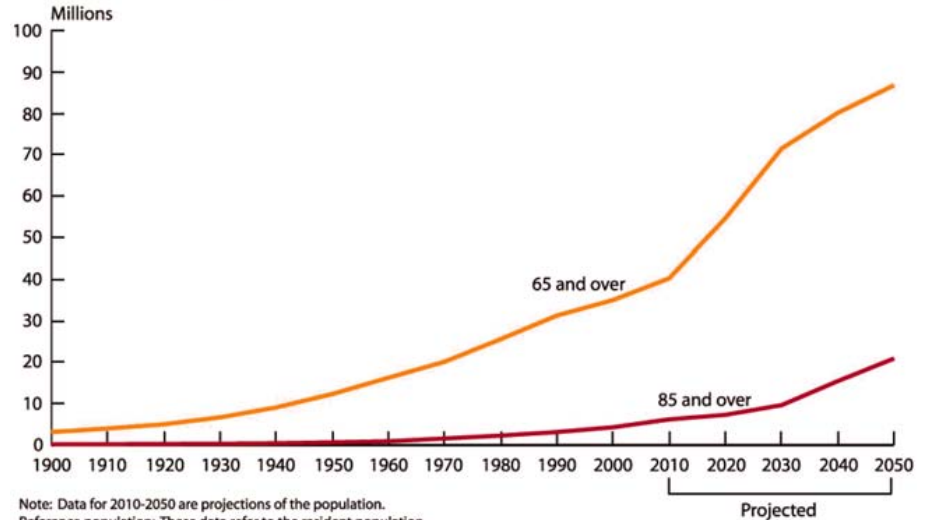
- Background
- Problems
- Research Question
- Proposed Framework

# Background

# Background

## The world is getting older

Number of people age 65 and over, by age group, selected years 1900-2000 and projected 2010-2050



Note: Data for 2010-2050 are projections of the population.  
Reference population: These data refer to the resident population.  
Source: U.S. Census Bureau, Decennial Census and Projections.

# Background

## Old People & Techs [1,2,3]

- **Common Belief: They don't like techs**
- **Study shows: They are receptive to using techs with:**
  - **more anxiety**
  - **less confidence**



[1] Czaja, Sara J., and Chin Chin Lee. **“The impact of aging on access to technology.”** *Universal Access in the Information Society* 5.4 (2007): 341-349.

[2] Czaja, Sara J., et al. **“Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE).”** *Psychology and aging* 21.2 (2006): 333.

[3] Marquié, Jean Claude, L. Jourdan-Boddaert, and Nathalie Huet. **“Do older adults underestimate their actual computer knowledge?”** *Behaviour & Information Technology* 21.4 (2002): 273-280.

# Background

Recent study [1] shows that:

- Internet would be a potentially valuable source of health information for older adults
- They indicated that they would use the Internet to seek health-related information and advice

[1] Czaja, Sara J., et al. "**Older adults and internet health information seeking.**" *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 53. No. 2. SAGE Publications, 2009.

# Problems

# Problems

(1) Degeneration of physical and cognitive ability.



# Problems

(2) Hard to reformulate accurate queries efficiently for the aging population. [1]

[1] Kubeck, Jean E. "FINDING INFORMATION ON THE WORLD WIDE WEB: EXPLORING OLDER ADULTS'EXPLORATION." *Educational Gerontology*25.2 (1999): 167-183.

## Research Questions

- How new interactive searching components, such as keyword hints, could improve the efficiency and accuracy when the aging population search for health information online.

# Proposed Framework

- **Data**
- **Retrieval Model**
- **Interface Design**
- **Evaluation**

# Data

- We crawled the data from **WebMD.com**
  - **WebMD.com** is website publishing healthcare information, such as symptom checklist, pharmacy information, drugs information...
  - On WebMD.com, there are categories for different diseases
- We used webpages in four diseases categories:
  - ADHD (187)
  - Allergies (368)
  - Alzheimers (75)
  - Anxiety-Panic (52)

# Retrieval Model

## Vector Space Model:

- Represent each page using feature vector

# Features

- Professional Key Words: extracted from professional annotations of pages
- Non-Professional Key Words: extracted from the text in the page
- Category-level Key Words: Key words used to distinguish between different category of diseases
- Page-level Key Words: Key Words used to distinguish between different pages within each category

	<b>Non-professional</b>	<b>Professional</b>
<b>Category-level</b>	✓	✓
<b>Page-level</b>	✓	

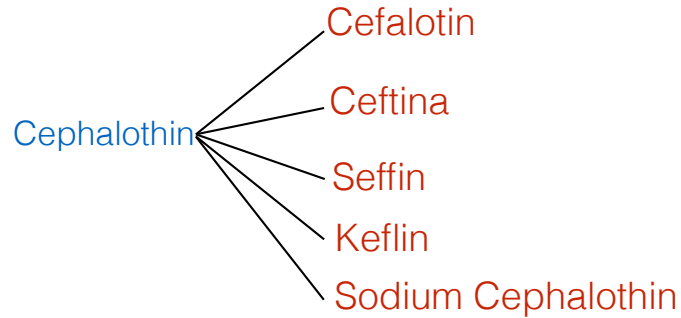
# Feature Extraction

## Professional Keywords Features:

- Using MeSH RDF to annotate each page
- MeSH RDF is a professional resource description framework released by national library of medicine
- It has (professional) descriptor/term pairs
- For each page, if the term appears, we use the descriptor to annotate that page

# Feature Extraction

Examples of annotating page with Mesh RDF:



Page 1 **Content:** ... How to use **Seffin** injection.... **Label:** Cephalothin



# Feature Extraction

## Professional Keywords Features (Category-level):

- For each category of disease, there will be a group of professional descriptors of the pages within this category
- Extract key descriptors for each disease **category** using TF-IDF
- For each category, we extracted the top 50 most descriptive words

# Feature Extraction

## Non-Professional Keywords Features (Category Level):

- Tokenize the title and content in for all pages in each category
- Extract descriptive keys words for each **category** using TF-IDF
- For each category, we extracted the top **50** most descriptive words

# Feature Extraction

## Non-Professional Keywords Features (Page Level):

- Tokenize the title and content in each web page
- Extract descriptive keys words for each **page** using TF-IDF
- For each page, we extracted the top **5** most descriptive words

# Feature Extraction

Features = professional Features + non-professional features

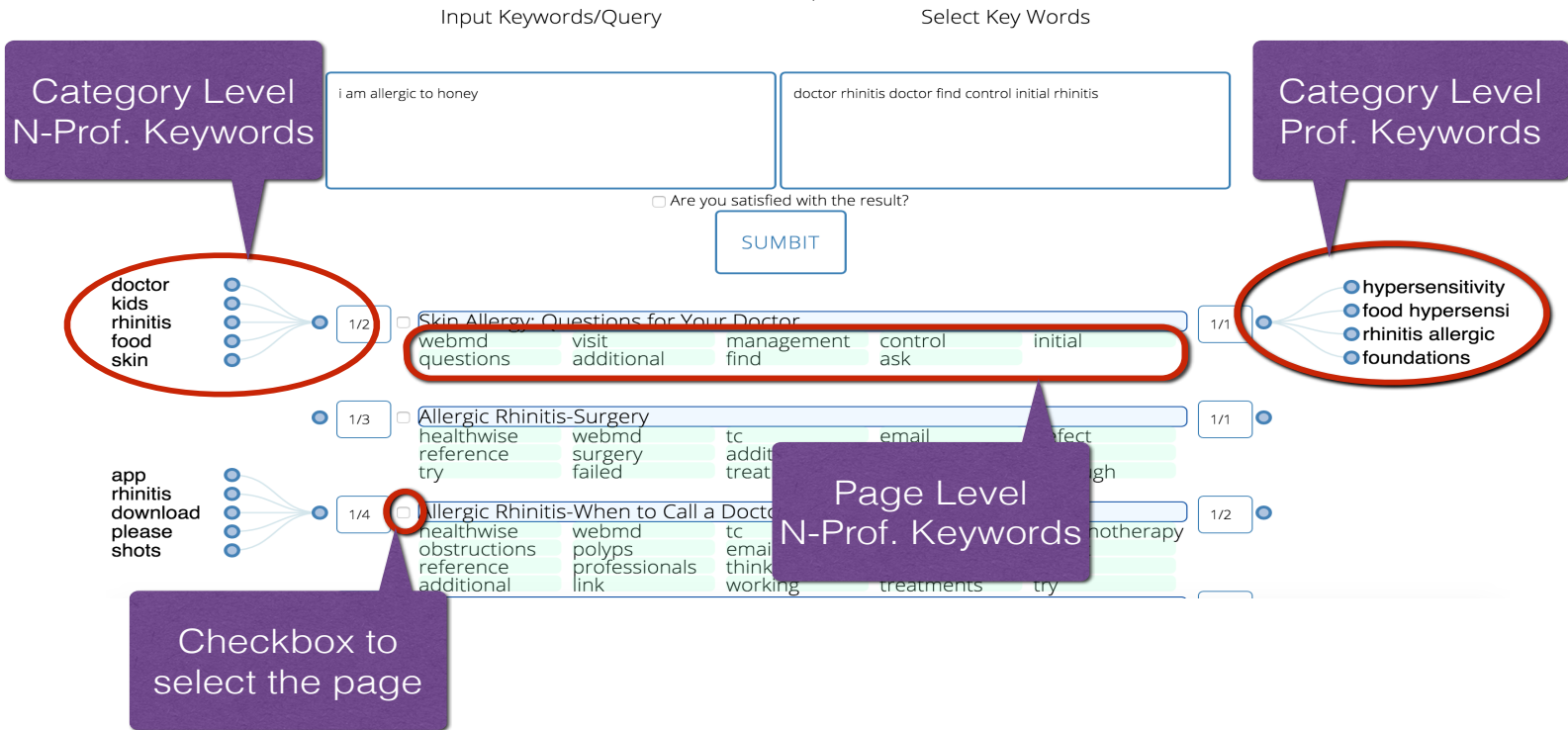
# Rank Pages

Cosine Similarity between the query and pages

# Interface Design

- Show non-professional and professional key words for each page on different sides
- Users could modify their query by selecting key words or editing
- Users could also select the page they prefer to look for more similar pages

# Interface Design



# Evaluation

- Recruit older adults to participate (Age > 65)



# Evaluation

- Ask each user to perform searching tasks
- Our interface vs. Normal searching interface
- Task Example:

Tom spent his holiday with his friends on a vegetable farm last month, they really had a nice time until their arms bit by some unknown insects. They all developed a raised welt around the sting site, and a tiny white mark was visible in the middle of the welt where the stinger punctured skin.

# Evaluation

- Record and compare the pages users went through and the number of iteration
- Do questionnaire to compare usability
- Give feedback

**Thank you**

Any Feedback?