

Multidisciplinary Expertise Retrieval with an Application in R&D Management

Choochart Haruechaiyasak

Alisa Kongthon

Human Language Technology Laboratory (HLT)
National Electronics and Computer Technology Center (NECTEC)
Thailand Science Park, Thailand

Presentation outline

- Background and motivation
- Proposed framework: **MULTI-ER**
- Case study: **Emerging Infectious Diseases (EIDs)**
- Conclusions
- Open discussion issues

Background and motivation

- Previous works in Expertise Retrieval (ER) mainly focused on finding people with a specific knowledge within an organization.
 - e.g., find top-10 programming experts on **Web service** and **SOA**
 - e.g., find a group of research experts on **HIV/AIDS immunology**
- Two types of ER tasks [Balog et.al]:
 - **Expert finding:** identify a list of people who carry some certain knowledge specified by the input query
 - **Expert profiling:** focuses on identifying the area of expertise associated with a given person based on **topical** and **social** information

Background and motivation (cont'd)

- Two main evaluation corpora for ER tasks:
 - The TREC 2005 Enterprise Track [Craswell et. al.]
 - Various contents from the W3C Web site: Web pages, emails, and source codes
 - Expert search task is to identify a list of W3C people who are experts for each given topical query.
 - The CSIRO Enterprise Research Collection (CERC) [Bailey et. al.]
 - Represents real-world search activities within enterprise
 - Use of internal staffs called *science communicators* to create some topics and perform the judgment

Multidisciplinary ER (MULTI-ER)

- **MULTI-ER** is a process which identifies and forms a group of expert candidates whose combined expertise is required to solve a multidisciplinary R&D problem.
 - For example, organizing a research forum to discuss on the global warming issue
 - Require experts with various knowledge such as scientists in various fields, sociologists and policy makers.

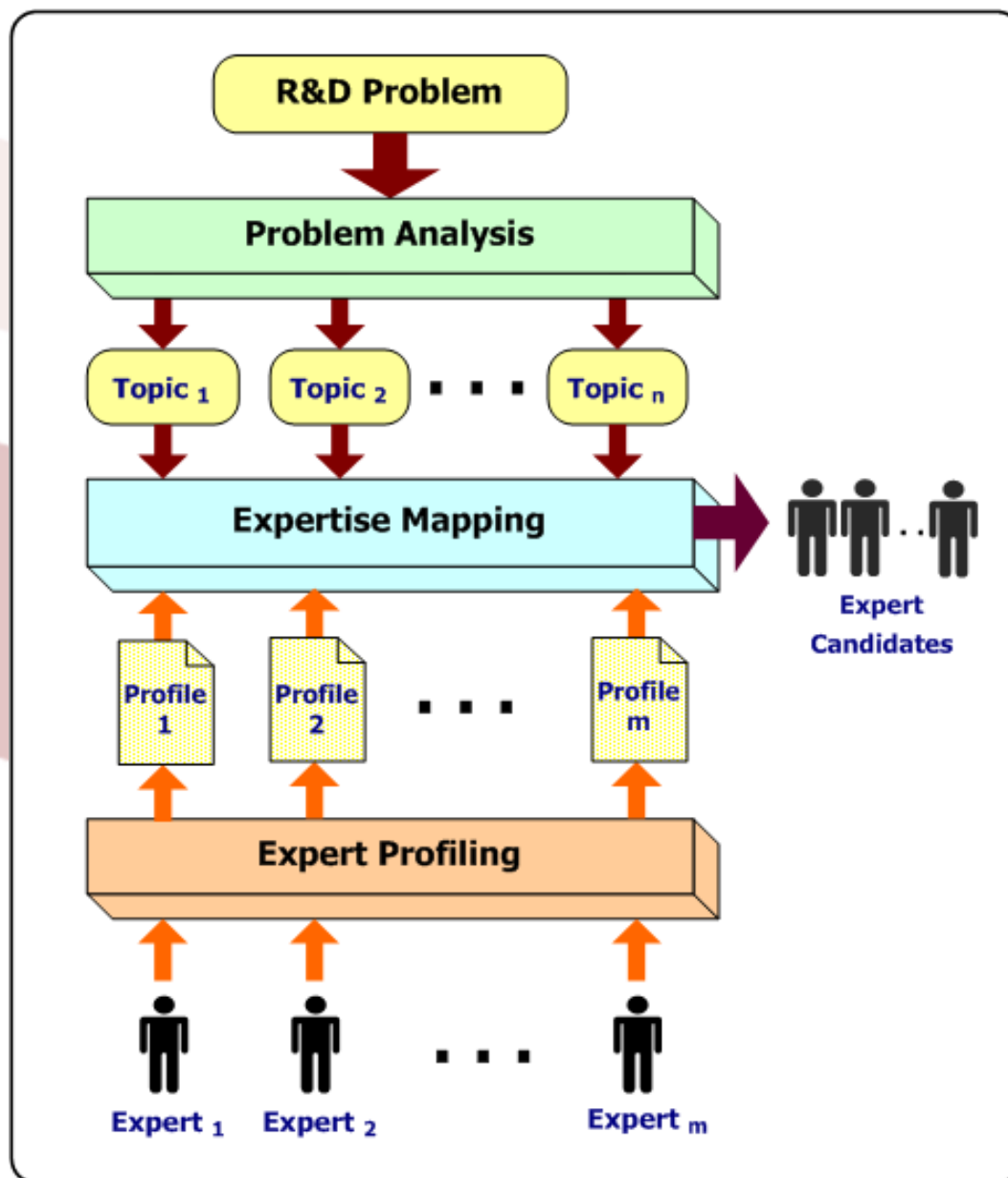
MULTI-ER: Support R&D management

- Example of R&D management tasks:
 - Organizing a forum to discuss on a certain problem issue
 - Forming a research workgroup to collaborate on a project.
- Typical steps:
 - **Analyze** the problem to identify all related topics.
 - **Search** for a list of potential expert candidates in each topic.
 - **Map** between each topic and the candidates based on profiles.
- MULTI-ER could provide a **decision support** function to assist users in performing the R&D management tasks by using techniques in **IR** and **text mining**.

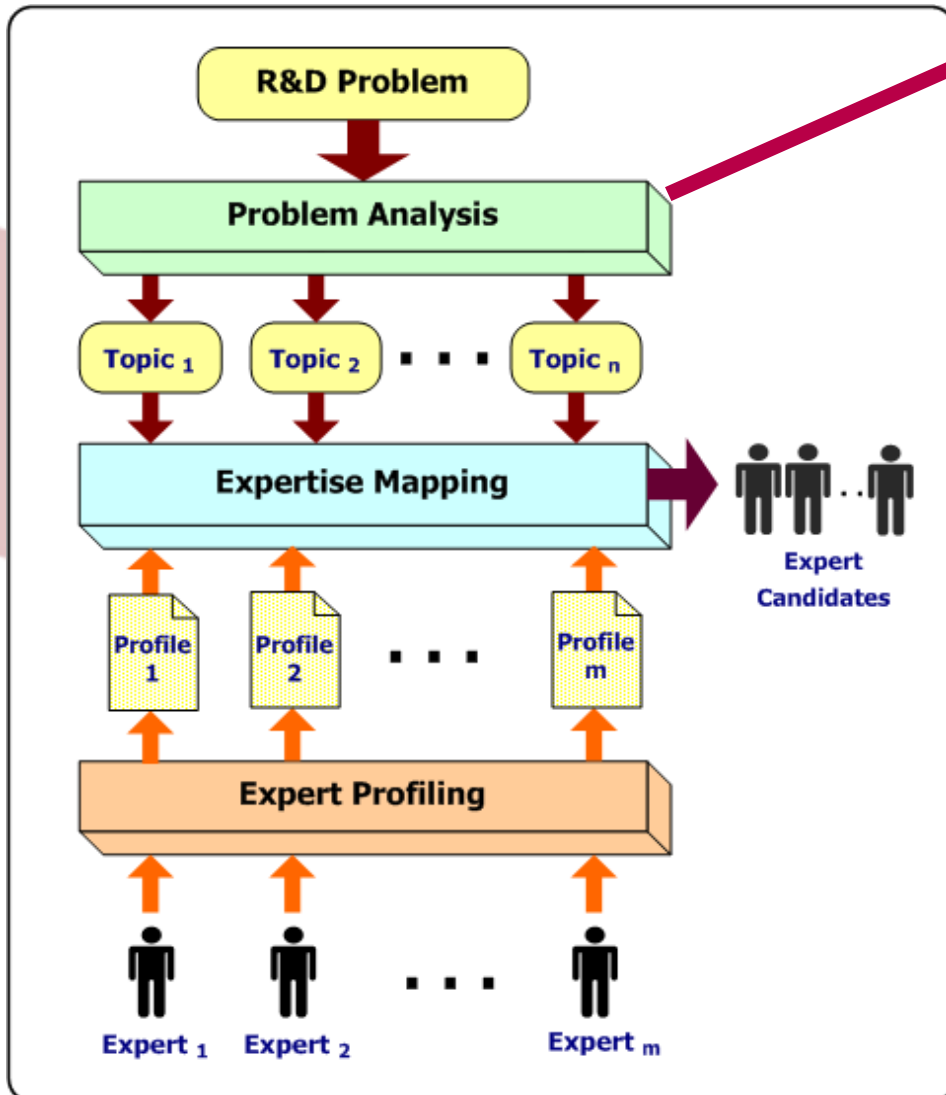
MULTI-ER: Characteristics

- Compared to the previous ER task, the MULTI-ER has two following differences.
 - **Larger problem scale:**
 - Global warming issue
 - Emerging Infectious Diseases (EIDs)
 - Alternative energy resources
 - International terrorism
 - **Broader expert scope:**
 - Interorganizational
 - National level
 - Global level

MULTI-ER: The Proposed Framework

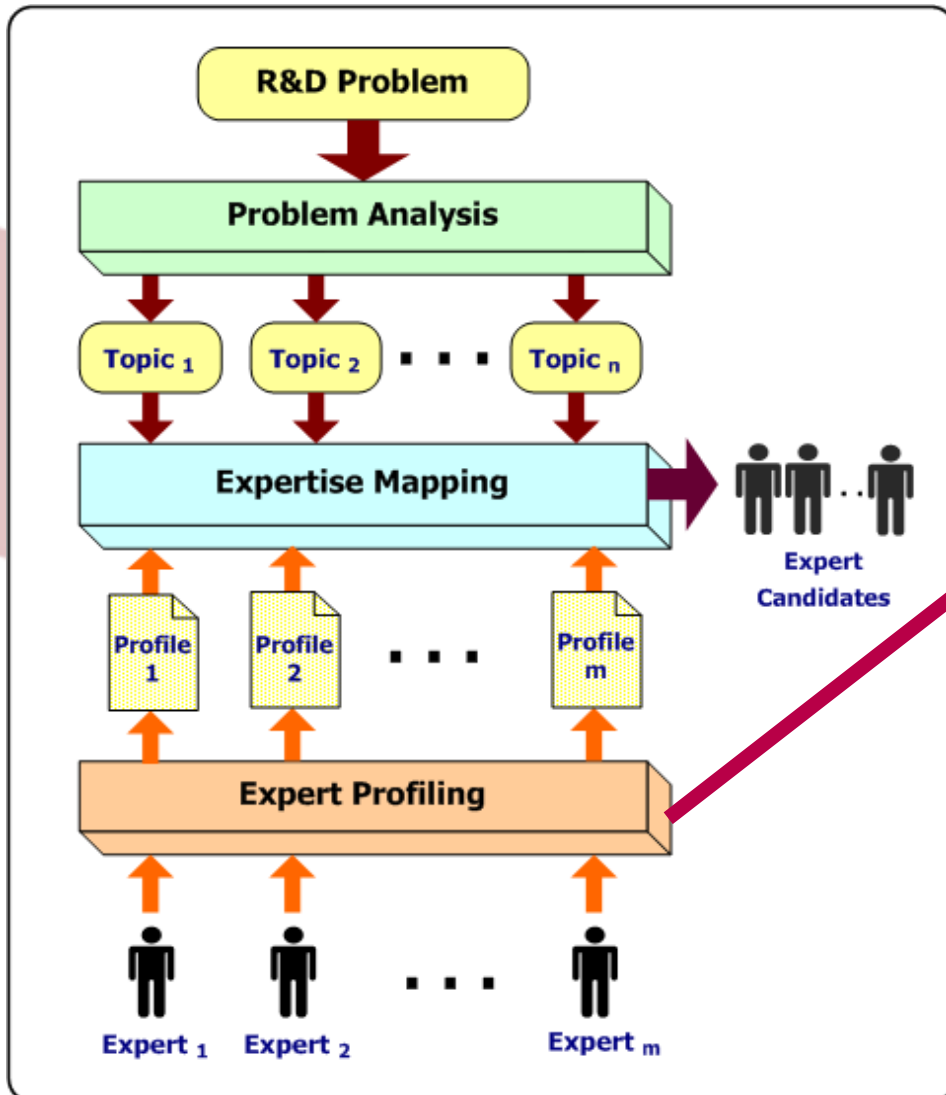


MULTI-ER: (1) Problem Analysis



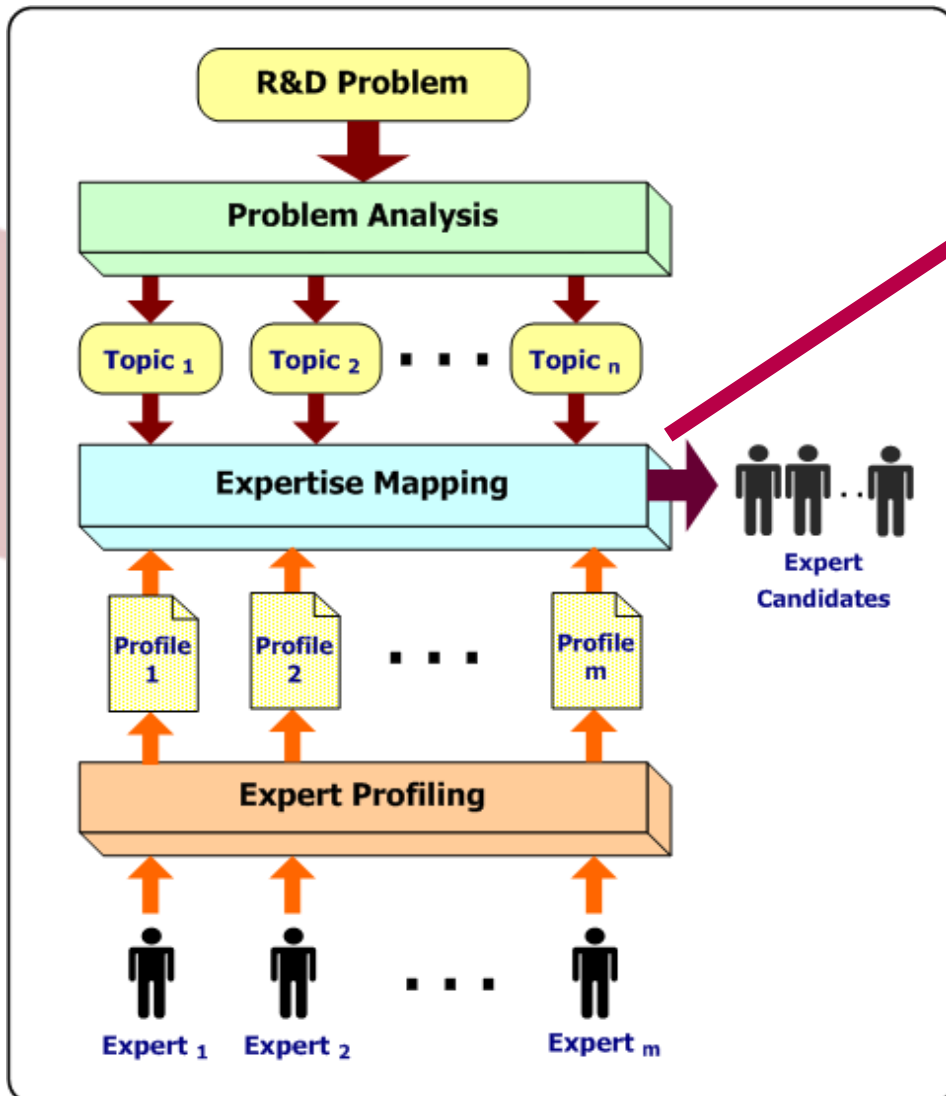
- Analyze a given R&D problem and identify a set of n related topics.
- Use resources such as **publication** and **patent** databases.
- **Difficult to evaluate** since the success on solving a given problem depends on many factors besides the topical coverage.
- **Topic verification** done by a group of initial experts

MULTI-ER: (2) Expert Profiling



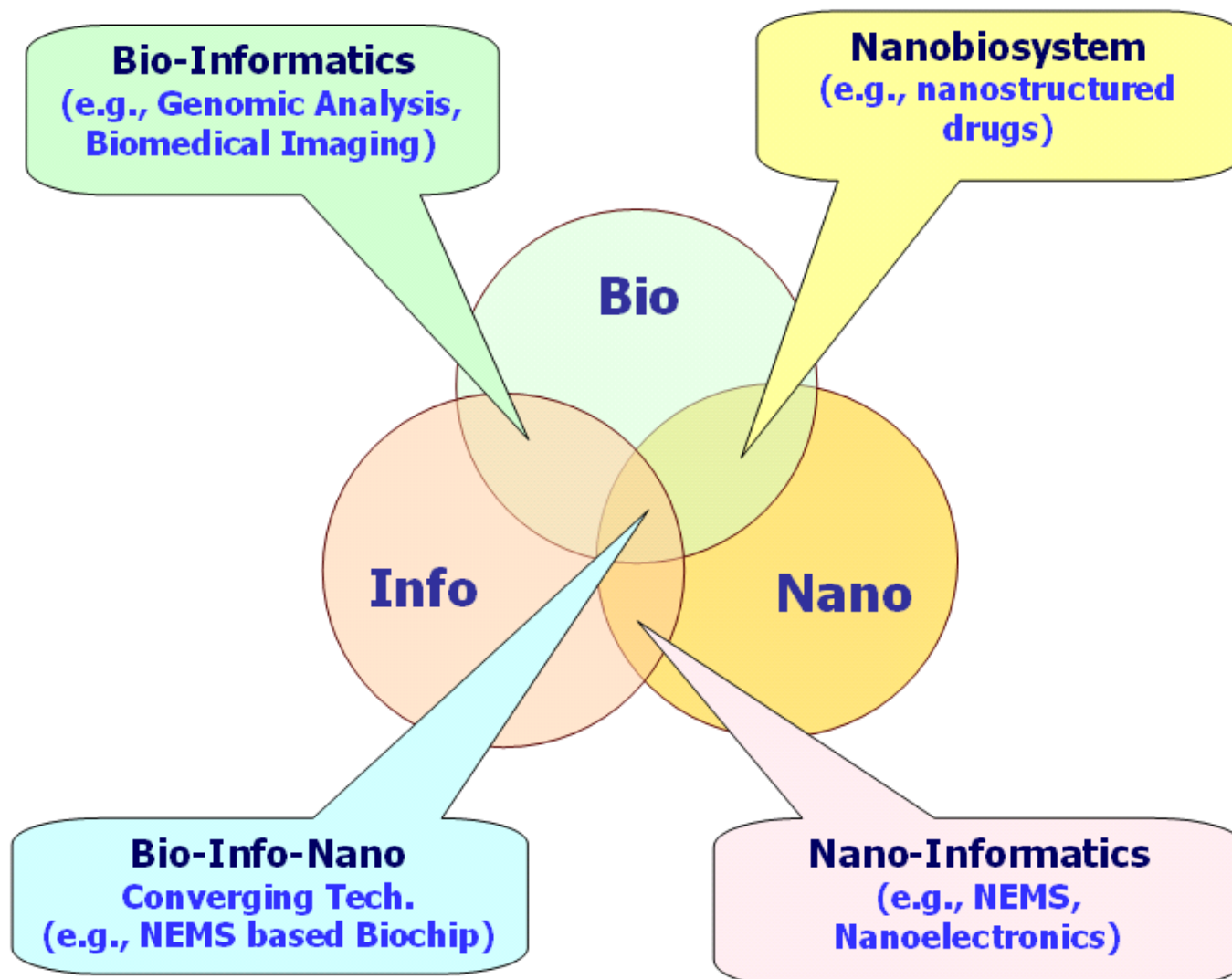
- Identify the expertise associated with an expert based on **topical** and **social** information.
- The output from the expert profiling is a set of m profiles describing each expert.
- Using individual information such as **CV** and **publications**.
- May need to acquire the personal information from outside organization.

MULTI-ER: (3) Expertise Mapping



- Perform the **matching** between the profiles and the topics.
- The output from this process is **a group of candidates** whose combined expertise are needed to solve the given problem.
- Making the decision to form a group of candidates is not straightforward as it depends on many factors such as
 - Experience levels of the experts
 - The success level of the previously performed works.

Case study: EIDs



EIDs in MULTI-ER Context

Problem Analysis

- ❑ Study reports and research articles
- ❑ Conduct focused interview with an initial group of experts
- ❑ Perform text mining using R&D bibliographic databases

Expert Profiling

- ❑ Identify scientists and technologists
 - expert referral
 - CVs and research papers
 - research social activities

Expertise Mapping

- ❑ Conduct three roundtable meetings among the program committee and target experts

Conclusions

- We proposed a new task called **MULTI-ER** as a process to identify a group of expert candidates to solve a multidisciplinary R&D problem.

Comparison between the ordinary ER and the proposed MULTI-ER

Factors	ER	MULTI-ER
Problem Scale	Small and specific	Large and multidisciplinary
Expert Scope	Within a single organization	Across multiple organizations
Data Resource	Intranet and internal DBs	The Internet and Outside DBs
Organizational level	Intra-organizational, e.g., W3C and CSIRO	Interorganizational, e.g., UN, WHO and FAO

Q & A

- **Problem analysis:**

- What types of information should be considered?
- What techniques should be applied?
- How many subtopics should we select?
- How to make sure that all relevant topics are included to solve the problem successfully?
- How can we prioritize those selected topics?
- ...

Open discussion issues (cont'd)

- **Expert profiling:**

- What types of information are needed to describe the expertise of a person?
- How to obtain related information on an expert?
- How to model an expert's experience for profile construction?
- ...

Open discussion issues (cont'd)

- **Expertise mapping:**

- How to rank the experts on a given topic?
- Should we consider the social and the topical information equally important?
- Should we also consider including other non-measurable and subjective factors such as interpersonal skill, experience-level into this process? If yes, how?
- ...

- **Corpus construction and evaluation metrics:**
 - Should we consider each process within the framework separately and evaluate each individually?
 - Where and how to obtain the information resources to build a real-world corpus to evaluate the framework?
 - ...

Open discussion issues (cont'd)

- **User involvement:**

- The experience and feedback from the users are very useful to develop a successful framework.
- Which process is considered the most important for the users?
- ...

Thank you for your attention