

# **READ ARTICLE MANAGEMENT IN DOCUMENT SEARCH PROCESS FOR NTCIR-9 VISEX TASK**

**YASUFUMI TAKAMA**

**SHUNICHI HATTORI**

**RYOSUKE MIYAKE**

**TOKYO METROPOLITAN UNIVERSITY**

# **CONTENTS**

- **Event Collection Task**
- **TM2011 System**
- **Experimental Results**

# EVENT COLLECTION TASK

- **User task: Investigate & report situation about given topic**

- **5 Topics (Topic 0 for training)**

- T0: Incidents all over the world where dinosaur fossils have been excavated.



- T1: Airplane crashes that have happened in Asia.

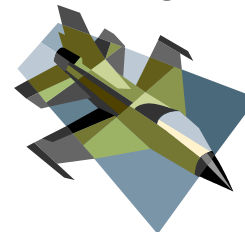


- T2: Incidents that have occurred at Japanese Nuclear Power Plants.



- T3: Nuclear weapons tests that have been carried out by various countries.

- T4: Incidents where NATO has recognized cases of friendly fire.



# TARGET DOCUMENT & INFORMATION OF EACH EVENT

- Document set: Mainichi newspaper articles (1998-2001)
- Common: date & place
- T0: Name of excavation team, category, era of dinosaur found, value of findings
- T1: Airline company name, type of airplane, # of casualties, cause of crash
- T2: Outline of incident, damage
- T3: Outline of test, country carried out the test, reason
- T4: Outline of case (damaged facilities, # of casualties)

# TM2001: DESIGN PRINCIPLE

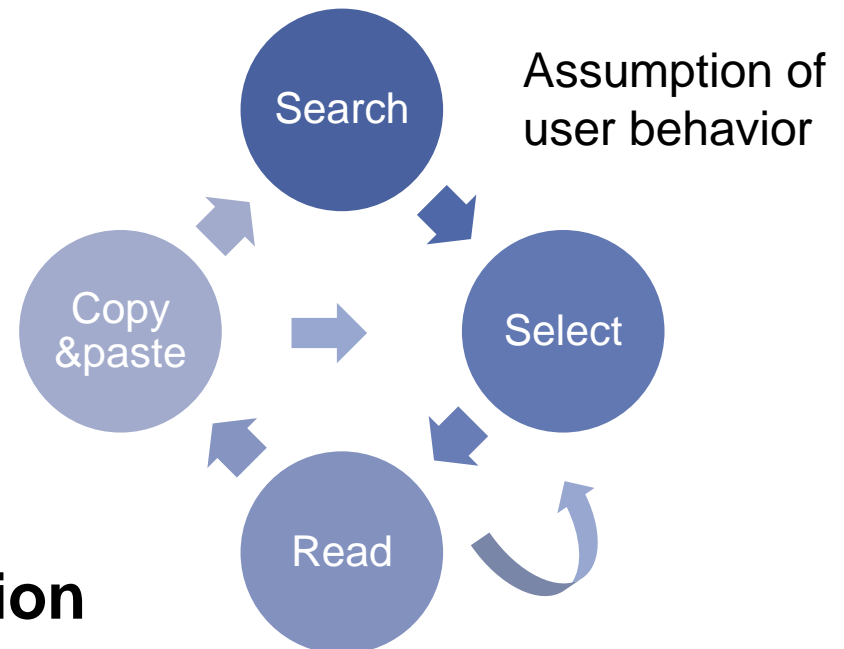


- Aim of VisEx: Evaluate IR systems in context of explorative information access
- ➔ Focus on evaluation of single & simple function for supporting event collection task

## ■ Search result contains

- Unrelated documents
- Already collected events

- ➔ Support for filtering such documents
- ➔ Marking read/unread function



# TM2011: TOP PAGE

[トップページ](#) | [利用マニュアル](#) | [既読記事一覧](#)

## VisEx検索システム

### 検索

検索キーワードを入力してください。

キーワード:   Query term

### 検索オプション

表示件数:  ソートキー:  ソート順:  Ascending/descending order

**その他メニュー** Sort by score/ release date

### 利用マニュアル

本検索システムの利用方法についての説明です。

### 既読記事一覧

既読となった記事の一覧を表示します。

Similar to  
Baseline system

[トップページ](#) | [利用マニュアル](#) | [既読記事一覧](#)

# TM2011: RETRIEVED RESULT

[トップページ](#) | [利用マニュアル](#) | [既読記事一覧](#)

## VisEx検索システム

### 検索

キーワード:

表示件数:  ソートキー:  ソート順:

### 誤爆の検索結果

586件の結果が見つかりました。(1番目から20番目を表示)

[次の20件を表示](#)

[米紙報道「誤爆はB2ステルス」--NATO重、中国大使館誤爆](#)

日付:1999-05-11 [\[既読にする\]](#)

【ワシントン10日共同】米紙ワシントン・ポストは10日、ヘオグラードの中国大使館を誤爆したのは米軍が初めて実戦投入したB2ステルス爆撃機であると報じた。米ミズーリ州から2機が出撃を続けていた。

[難民収容所の誤爆を否定--NATO報道官](#)

[クリントン米大統領、改めて謝罪--NATO重、中国大使館誤爆について](#)

日付:1999-05-11 [\[既読にする\]](#)

【ワシントン10日布施広】クリントン米大統領は10日、中国大使館誤爆について、「中国国民と指導部に謝罪する」と語った。コーエン国防長官も同日の会見で、誤爆したのが米軍機であることを公表し「深い遺憾の意」を表明した。大統領は8日に「遺憾の意」を表明、9日には中国の江沢民国家主席に「不運な手違い（誤爆）」を遺憾とする書簡を送った。一方、米國務省は10日、オルブライト國務長官が9日午前0時（日本時間同午後1時）ごろ、ワシントンの中国大使館を訪れ、誤爆を謝罪したうえ、中国外務省あての手紙を手渡したことを

[中国大使館誤爆 誤爆はB2ステルス--米紙報道](#) [【大版】](#)

Mark as read

Title  
Date  
Snippet



[\[既読にする\]](#)

# MARKING READ DISPLAY STYLE

## (a) New article

- Boldface & background color

## (b) Already retrieved

## (c) Mark as read

- Title only & grayed out
- Stored in read-article list
- Change to (b) by unread function

- (a)
- ユーゴ空爆 NATO軍の住宅誤爆犠牲者が20人に  
-----  
日付:1999-04-29 [既読にする]
- 【ベオグラード28日=佐々木良寿】セルビア国営  
少なくとも二十人に達したと伝えた。
- (b)
- 米紙報道「誤爆はB2ステルス」――NATO軍、中国大使館  
-----  
日付:1999-05-11 [既読にする]
- 【ワシントン10日共同】米紙ワシントン・ポストは1  
したB2ステルス爆撃機であると報じた。米ミズーリ州が
- (c)
- 難民収容所の誤爆を否定――NATO報道官  
-----  
包括的、地上部隊、改定、NATO軍、中



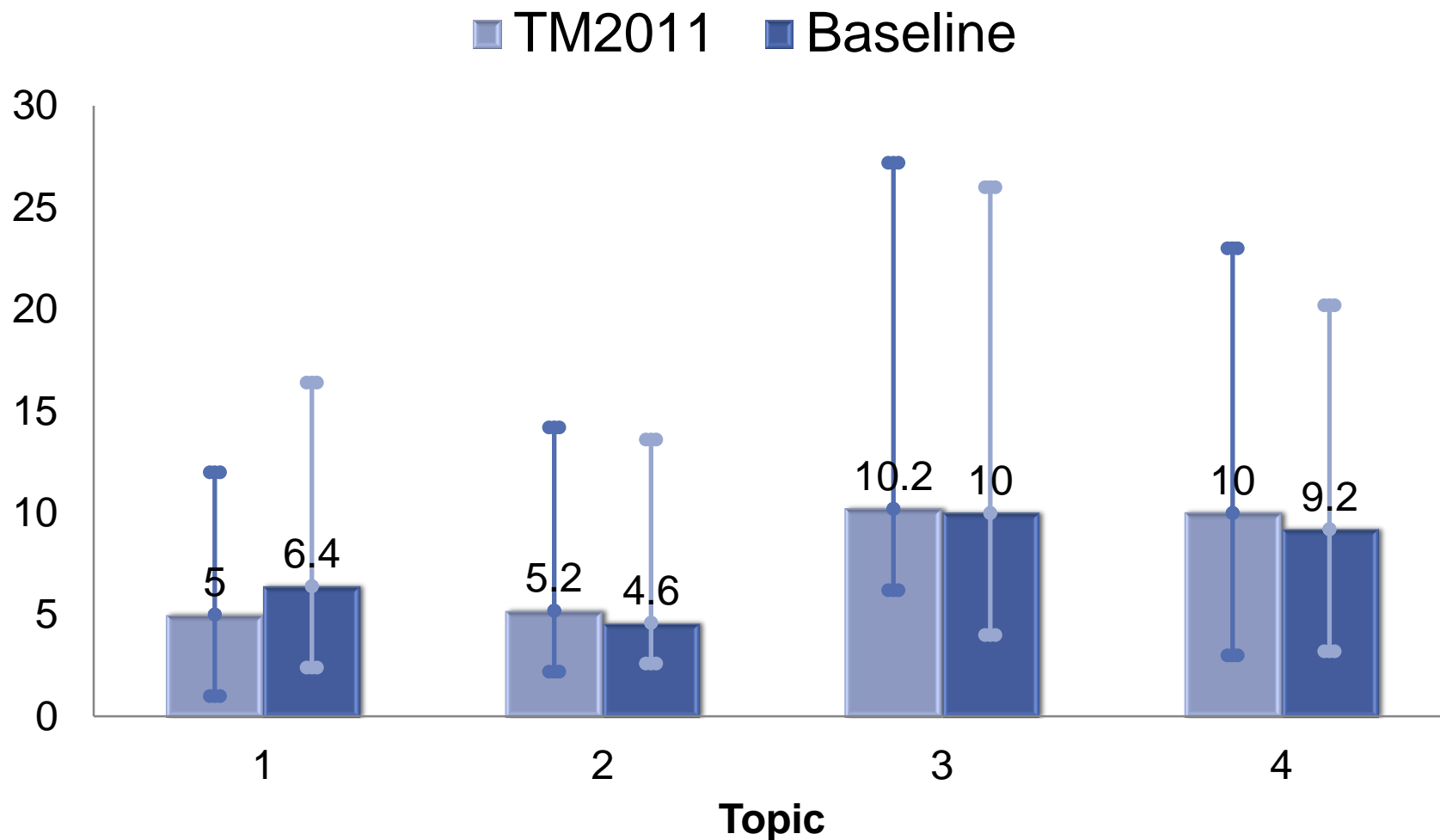
# **USAGE OF MARKING READ FUNCTION**

- 1. Articles to be included in a report are marked as read.**
- 2. Articles related with already-included event in a report are marked as read without reading those.**
- 3. Unrelated articles with target topic are marked read for improving readability of subsequent retrieval results.**

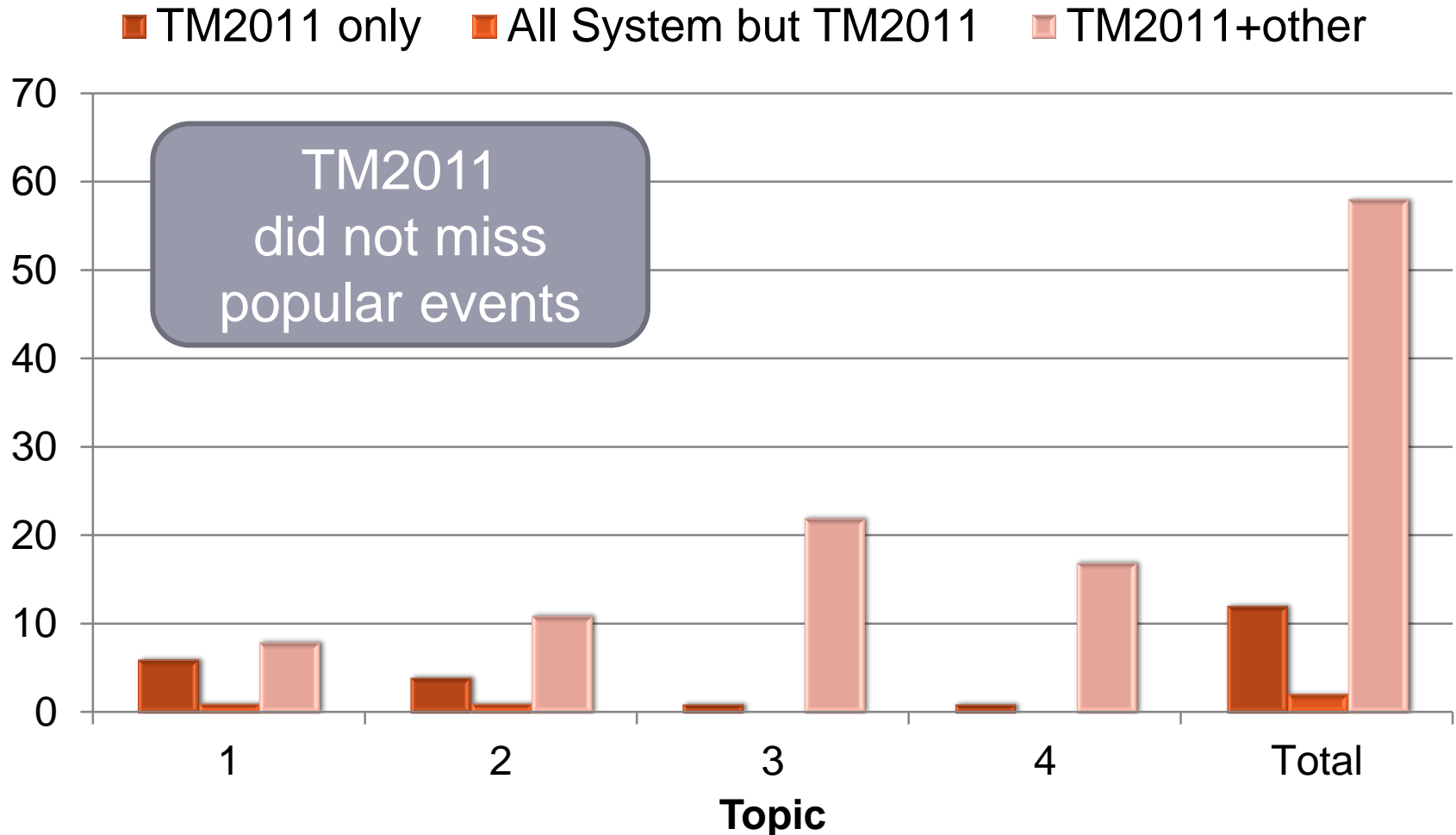
# OUTLINE OF EXPERIMENT

- **Date: TM2011(2011.8.15), Baseline(2011.8.18)**
- **Between-subjects experiment design**
- **1 hour for each topic (total 4 topics)**
- **Questionnaire**
  - Before experiment ... experience & skill of Web search & writing reports, interest in current issues
  - After each topic ... impression about task (topic)
  - After experiment ... impression about system & tasks

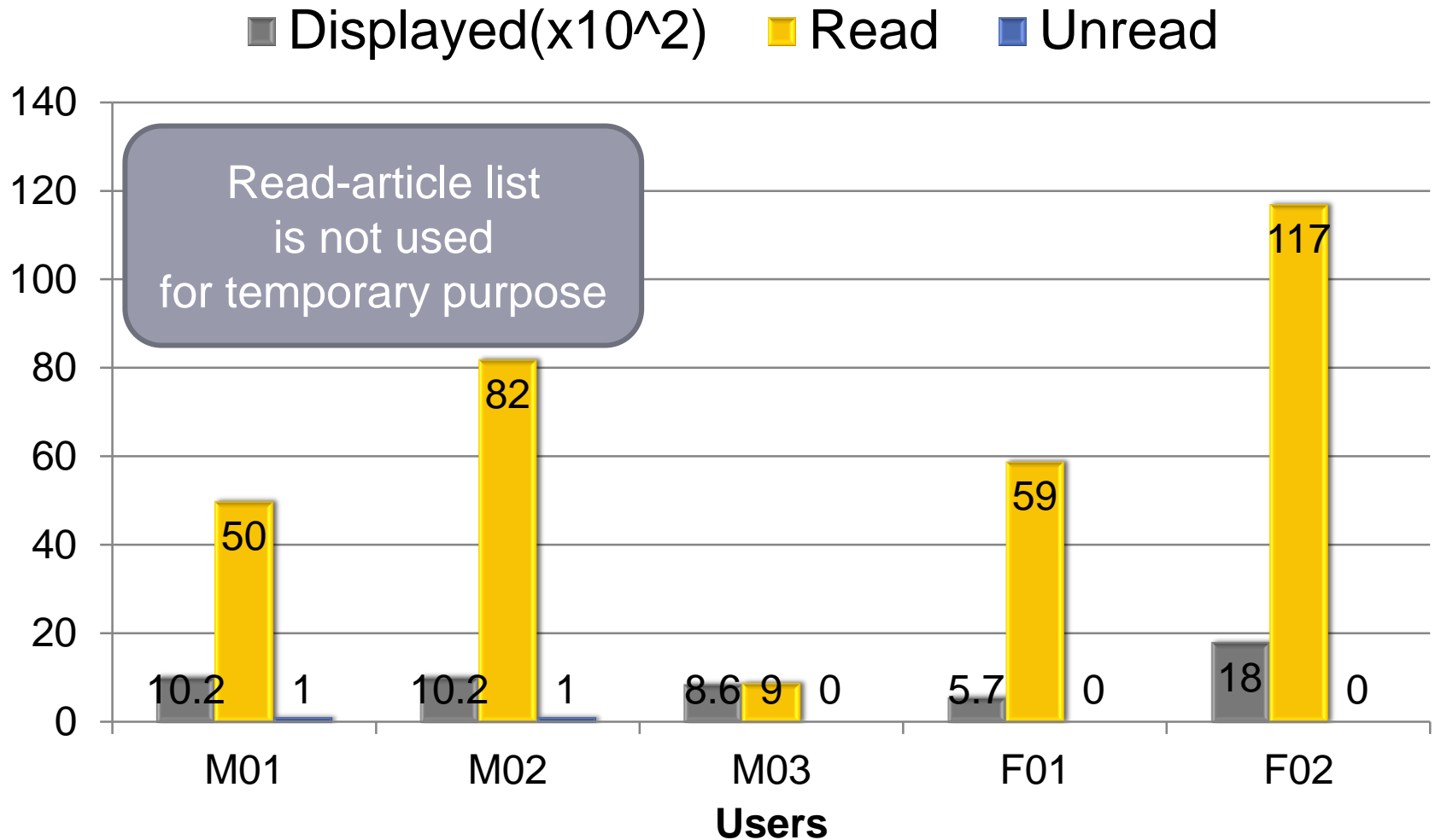
# RESULTS: # OF EVENTS ANSWERED



# CLASSIFICATION OF ANSWERED EVENTS



# USAGE OF MARK READ/UNREAD FUNCTIONS



# USAGE OF MARK READ FUNCTIONS / TOPIC

Used in report

Marked & used

Marked

	M01	M02	M03	F01	F02
	5	15	10	6	13
1	4	0	1	6	12
	12	19	1	15	49
	9	4	9	9	12
2	5	0	2	4	2
	8	12	5	21	17
	14	13	8	8	18
3	11	10	0	1	11
	11	28	0	6	19
	19	14	11	11	16
4	17	14	2	5	16
	19	23	3	17	32
	47	46	38	34	59
Total	37	24	5	16	41
	50	82	9	59	117

# RELATION BETWEEN MARKING & SEARCHING ACTIONS

Search

Search + mark

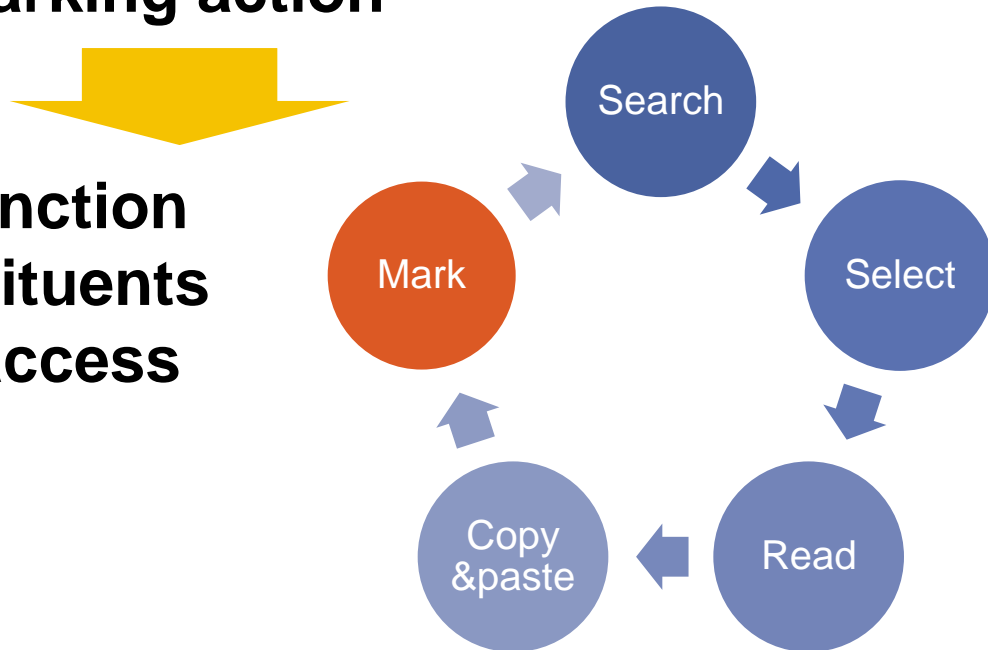
Search + multiple mark

	M01	M02	M03	F01	F02
	13	35	16	13	53
1	9	20	1	13	31
	3	1	0	2	7
	24	40	30	52	58
2	8	11	5	20	13
	0	1	0	1	2
	21	24	46	54	42
3	11	18	0	6	12
	0	9	0	0	3
	29	29	16	54	37
4	20	20	3	17	17
	0	3	0	0	6
	87	128	108	173	190
Total	48	69	9	56	73
	3	14	0	3	18

# DISCUSSION

- **Mark read function is used for supposition 1):**  
Articles to be included in a report are marked as read
- **Frequently observed behavior: Do next search**  
**JUST AFTER** marking action

- **Marking read function**  
**as typical constituents**  
**of information access**  
**cycle**

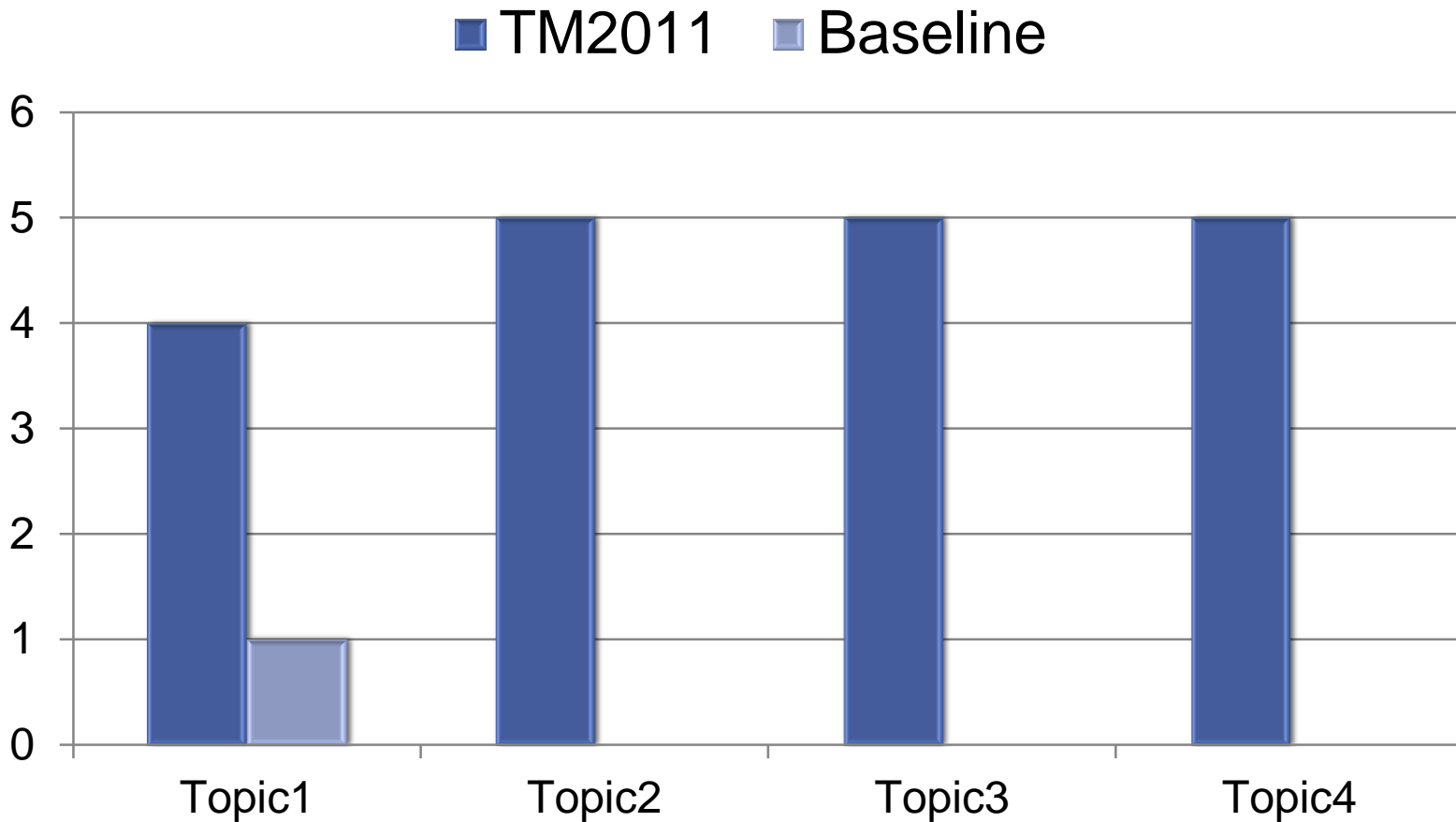




# RESULTS OF QUESTIONNAIRES

- **Experience & skill of Web search: Frequently used**
- **Interests in current issue: depends on users**
- **Familiarity of topics:**
  - Familiar: 2, Unfamiliar: 4
  - No relation with performance of tasks?
- **Overall impression of systems**
  - TM2011&Baseline: simple, easy to use, quick response
  - Marking functions: useful (M02, F01), although some problems
- **Big difference between TM2011 & Baseline:  
# of new findings / surprises during tasks**

# # OF USERS WHO HAD NEW FINDINGS/SURPRISES



# CONCLUSION

- **Result of Taking part in NTCIR-9 VisEx Task**
- **TM2011: marking read / unread functions**
  - Usage of functions investigated based on log data
- **Future works**
  - Investigate usages of marking read functions other than supposition 1)