

Overview of NTCIR-9 1CLICK

Tetsuya Sakai

Makoto P. Kato

Young-In Song

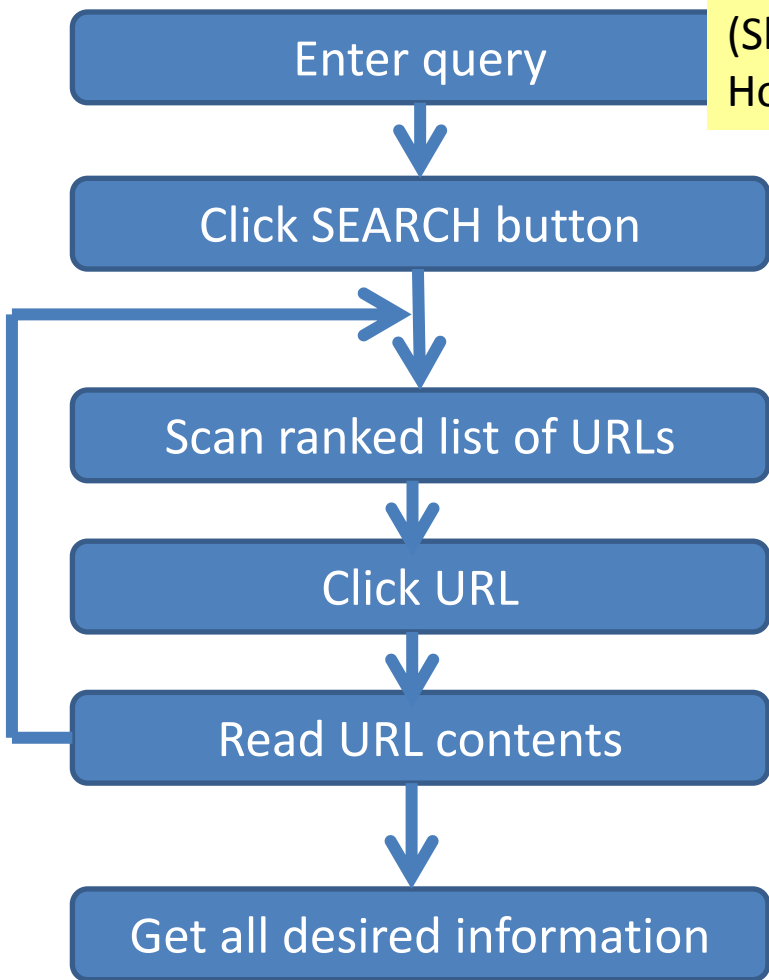
December 9, 2011@NTCIR-9, Tokyo

TALK OUTLINE

- **MOTIVATION**
- **TASK**
- **METRICS**
- **RESULTS**
- **TENTATIVE CONCLUSIONS**

Traditional search vs One Click Access

湘南厚木病院
(Shonan Atsugi
Hospital)



Phone: 046-223-3636. Fax: 046-223-3630. Address: 118-1 Nurumizu, Atsugi, 243-8551. Email: soumu@shonan-atsugi.jp. Visiting hours: general ward Mon-Fri 15-20; Sat&Holidays 13-20 / Intensive Care Unit (ICU) 11-11:30, 15:30, 19-19:30.

Evaluate text based on *nuggets*

“1CLICK” Definition

- Given a query and an output window of size X , return a text (“ X -string”) that
 - Presents important nuggets first; and
 - Minimise the amount of text the user has to read (or the time spent)
- 1CLICK is **direct**: no user operation between clicking SEARCH button and seeing the output
- 1CLICK is **immediate**: very little time spent to locate relevant information within output

TALK OUTLINE

- MOTIVATION
- **TASK**
- METRICS
- RESULTS
- TENTATIVE CONCLUSIONS

1CLICK Queries

- 60 queries; 4 query types based on **SIGIR'09 Good Abandonment paper**

Selected from a mobile query log

- CELEBRITY date of birth, profession, hobbies, etc.
- LOCATION contact info, address, opening hours, etc.
- DEFINITION (of a term)
- QA (Natural language questions)

Selected from Japanese Yahoo! Answers

- 2,839 nuggets
 - facts as of Dec 31, 2010

Query type	Mean	Range
CE	126.5	[38, 368]
LO	30.9	[10, 125]
DE	26.3	[2, 174]
QA	5.6	[2, 26]
Total	47.3	[2, 368]

0001 LO 1 東山動物園
0002 QA 1 スタジオジブリの長編映画で、風の谷のナウシカからハウルの動く城までの全タイトルが知りたい
0003 LO 1 墨田都税事務所
0004 QA 1 上野動物園以外で、パンダが見られる日本の動物園は
0005 LO 1 メナード青山リゾート
0006 LO 1 神戸市立中央図書館
0007 QA 2 湯村温泉はどこにあるか
0008 LO 1 東京地方裁判所
0009 CE 1 山崎邦正
0010 CE 1 浅田真央
0011 DE 1 オンドゥル語
0012 CE 1 細木教子
0013 CE 1 中野美奈子
0014 DE 2 スカルプチャー
0015 QA 1 血液検査でSCC、NSEは何を意味するか
0016 QA 1 通天閣と宇都宮タワーはどちらが高いか
0017 LO 1 港北区役所
0018 CE 1 釈由美子
0019 DE 1 ファーストピアス
0020 CE 1 大沢たかお
0021 QA 1 足し算を「加法」、掛け算を「乗法」というが、引き算と割り算は何法というか
0022 CE 1 真鍋かをり
0023 DE 5 カラシニコフ
0024 LO 1 熱田神宮宮庁
0025 CE 1 小泉進次郎
0026 DE 1 ドーハの悲劇
0027 DE 1 日本国民の三大義務
0028 QA 1 常盤貴子と深津絵里と稲森いずみの中で、一番年上と年下は誰か
0029 LO 1 東京ドーム
0030 CE 1 中川昭一
0031 CE 1 手塚治虫
0032 QA 1 SFメトロカードのSFは何の略か
0033 LO 2 福岡工業高等学校
0034 QA 1 ANA国内線で機内に持ち込める手荷物のサイズ
0035 QA 1 人間の血液が流れる速度
0036 LO 1 ホテルアーバイン鎌田
0037 LO 2 三本松駅
0038 CE 1 天海祐希
0039 LO 1 金沢大学
0040 QA 1 六角形と円錐をそれぞれ英語でなんとというか
0041 LO 1 東京ディズニーランド
0042 DE 1 敵に塩を送る
0043 QA 1 日本で一番長い苗字
0044 DE 2 デウス・エクス・マキナ
0045 QA 1 記憶容量の単位でMB、GB、TBより大きいのは
0046 LO 1 湘南厚木病院
0047 DE 1 排他的経済水域
0048 DE 1 一国一城令
0049 DE 3 イカ天
0050 CE 1 酒井法子
0051 CE 1 田口ランディ
0052 LO 1 大丸神戸店
0053 CE 1 島田紳助
0054 DE 1 むしずが走る
0055 DE 1 エレクトーン
0056 DE 1 ラムサール条約
0057 CE 1 GACKT
0058 DE 1 ぐる
0059 QA 1 obrigadolは何語でどういう意味か
0060 QA 1 ホスピタリティ

Mao Asada

Apart from Ueno Zoo, where in Japan can I see a panda?

Which is taller, Tsutenkaku or Utsunomiya Tower?

Kaori Manabe

Kalashnikov

The three duties of a Japanese citizen

Osamu Tezuka

Sanbonmatsu station

Kanazawa University

What are the storage capacity units larger than MB, GB and TB?

1st field: queryID
2nd field: query type
CE: CELEBRITY
LO: LOCAL
DE: DEFINITION
QA: Question Answering (15 queries each)
3rd field: number of intents
4th field: query string

1CLICK Nugget Record

Nugget = 5-tuple

– ID: **N001**

– Weight: **15**

– Nugget semantics:

- **Osamu Tezuka was born on Nov 3, 1928.**

– Vital string:

- **1928.11.03**

– Source URL:

- **<http://tezukaosamu.net/jp/about/>**

An atomic factual statement in natural language. Used by nugget match assessors for identifying the nugget.

Used for approximating the minimal string length to convey the meaning of the nugget in X-string.

1CLICK Participants and Runs

9

- **DESKTOP** runs: $X=500$
 - (about five search engine snippets “above the fold”)
- **MOBILE** runs: $X=140$
 - (satisfy user with one mobile screen or a “tweet”)
- Participants: 25 teams registered, **3 submitted**
(10 runs)

Kyoto U (Information Extraction)

Tokyo Institute of Tech (Summarisation)

MSRA (Passage Retrieval)

TALK OUTLINE

10

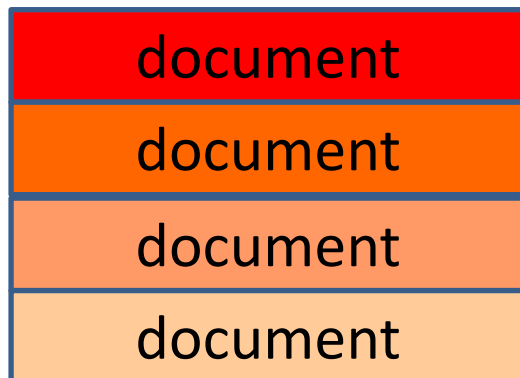
- MOTIVATION
- TASK
- METRICS
- RESULTS
- TENTATIVE CONCLUSIONS

Analogy with nDCG

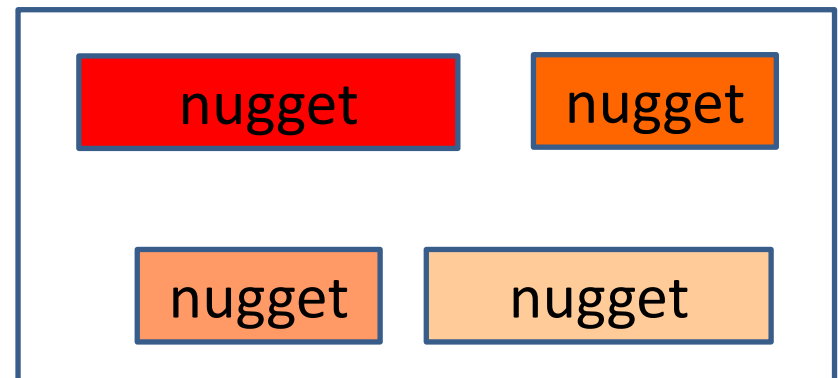
nDCG **discounts**
documents
based on ranks

S-measure **discounts**
nuggets
based on offsets
(positions in *X*-string)

Ranked list of documents



X-string



Offset-based linear discounting

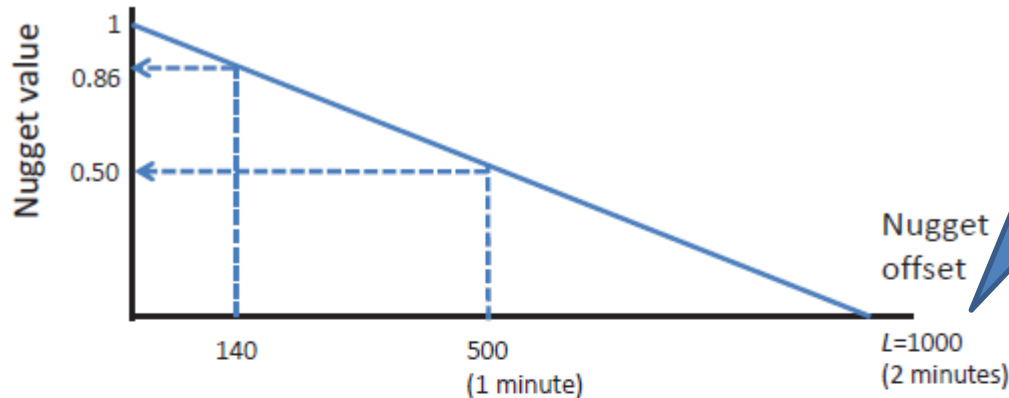


Figure 2: Discounting nugget values.

Assumption: user's reading speed is constant \rightarrow nugget value wears out linearly

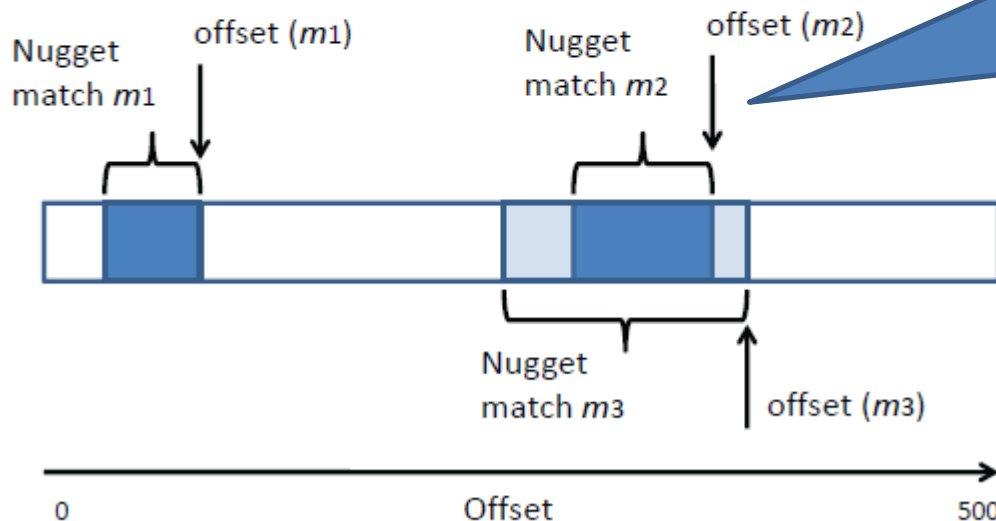


Figure 3: Nugget matches and offsets.

Position of a nugget match is taken as the END offset (user has to read THROUGH the nugget)

Manual nugget matching

13

Submission Evaluation

Job ID: 1501
Query: 北川景子
Description: 1C0-0001 CE 1 北川景子

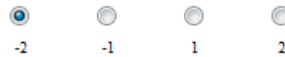
[Back to job list](#)

Content: 407 chars

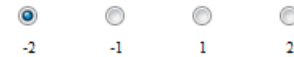
2011年3月1日...公式サイト。、出演...ブログ。/
の公式サイト。ニュース、フォト...の売れ筋商
品、出演DVD、写真集、CD、関連書籍
ど。http://www.amazon.co.jp...eRosebud。ファン
ンサイト。笑顔の写真集、事典、な.../2008年7月
14日...名前：(きたがわけいこ) 出身地：兵庫県神戸
市愛称：けいこちゃん、きーたん 生年月日：1986年8月
22日 血液型：o型身長：160cm体重：非公開所属事務所：
スターダストプロモーション.../2011年3月1日...
Yahoo!人物名鑑。2011年3月3日...の、画像、動画、
ニュース、作品情報、ブログ、Feペディア (出典：
Wikipedia) 等、Yahoo!人物名鑑でチェック！
http://talent.yahoo.co.jp/pf/detail/.../20
10年1月8日...きたがわ・けいこ。ヒートアイランド・
(c) 2007「ヒートアイランド」製作委員会。木村了/城
田優/。■生年月日：1986/08/22。■出身地：兵庫
県。IMDBで検索 [IMDBSearch]。をYahoo

Assessor
selects a
nugget
match area
and saves

Readability



Trustworthiness



Done

Memo

- 1986年8月22日生
- 生年月日: 1986年8月22日 [184,199]
- 兵庫県出身
- 血液型O型
- twitterアカウント kitagawa_keiko
- 公式ブログ http://star-studio.jp/kitagawa-keiko/
- 女優
- 公式モバイルサイト http://stardust-web.net/
- 身長160cm
- 2003年 ミスSEVENTEENに選ばれデビュー
- 北川景子は本名
- スターダストプロモーション所属
- 出演映画 Dear Friends ディア フレンズ
- 出演TV モップガール
- 2006年 SEVENTEENモデル卒業
- 2009年 明治大学卒業
- 趣味・特技 映画鑑賞

Offset recorded

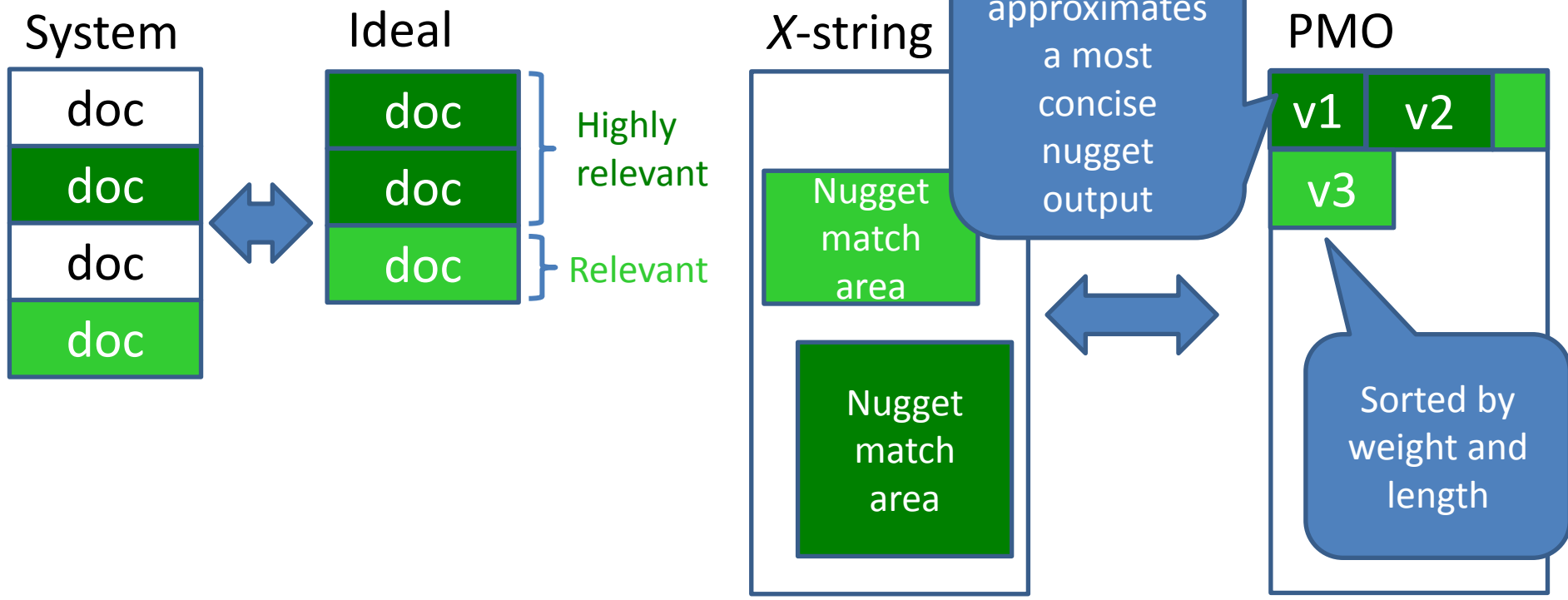
1CLICK encourages abstractive summarisation
by paraphrasing etc. to save space ->
automatic matching probably not sufficient

Pseudo-Minimal Output (PMO)

nDCG normalises system performance with **ideal ranked list**

Problem:
Nugget lengths vary!

S-measure normalises system performance with **PMO**



S-measure =

Nuggets found in X -string

Position of nugget in X -string

$$\sum_{m \in M} w(m) \max(0, L - \text{offset}(m))$$

$$\sum_{n \in N} w(n) \max(0, L - \text{offset}^*(v(n)))$$

All nuggets

Position of vital string in PMO

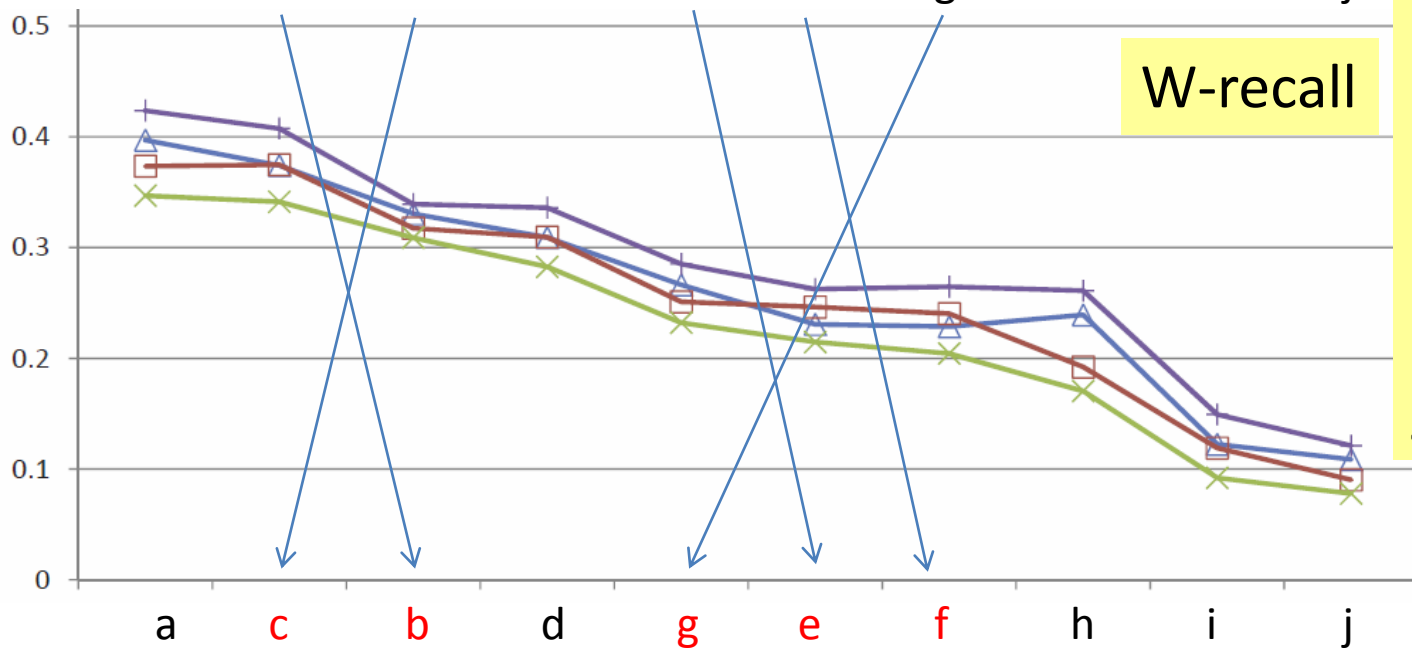
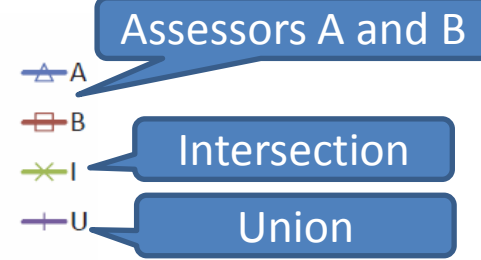
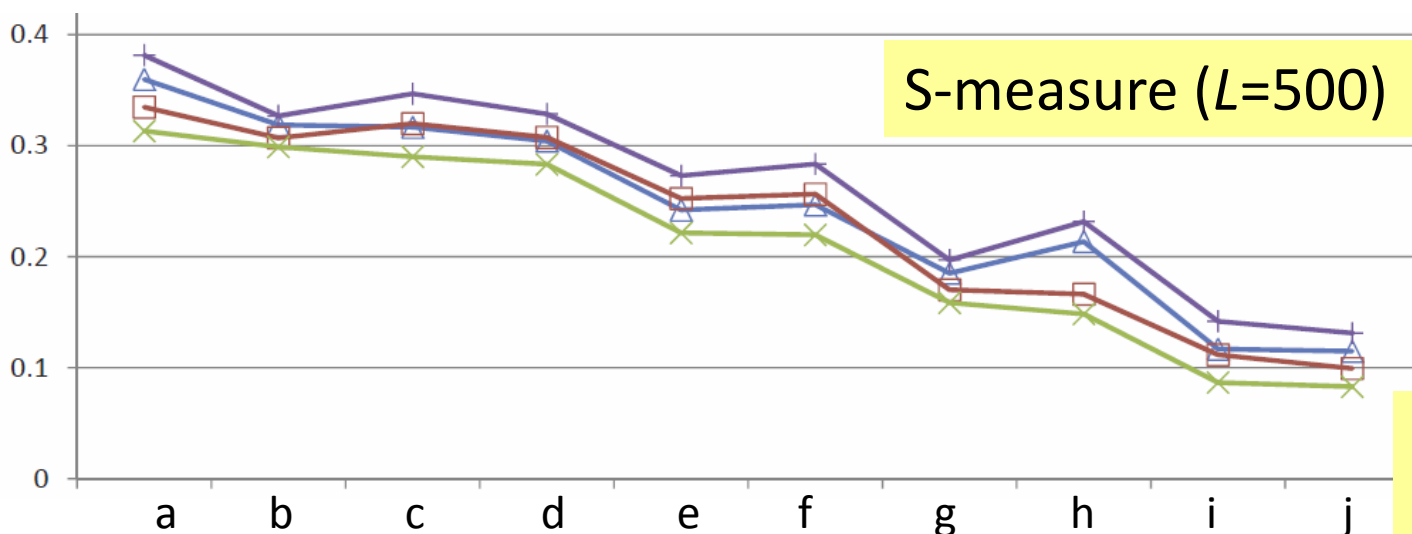
- Present important nuggets first!
- Minimise the amount of text the user reads!
- Reduces to Weighted recall (W-recall) when $L = \infty$
- For Japanese, $L = 500$ characters \doteq one minute

TALK OUTLINE

18

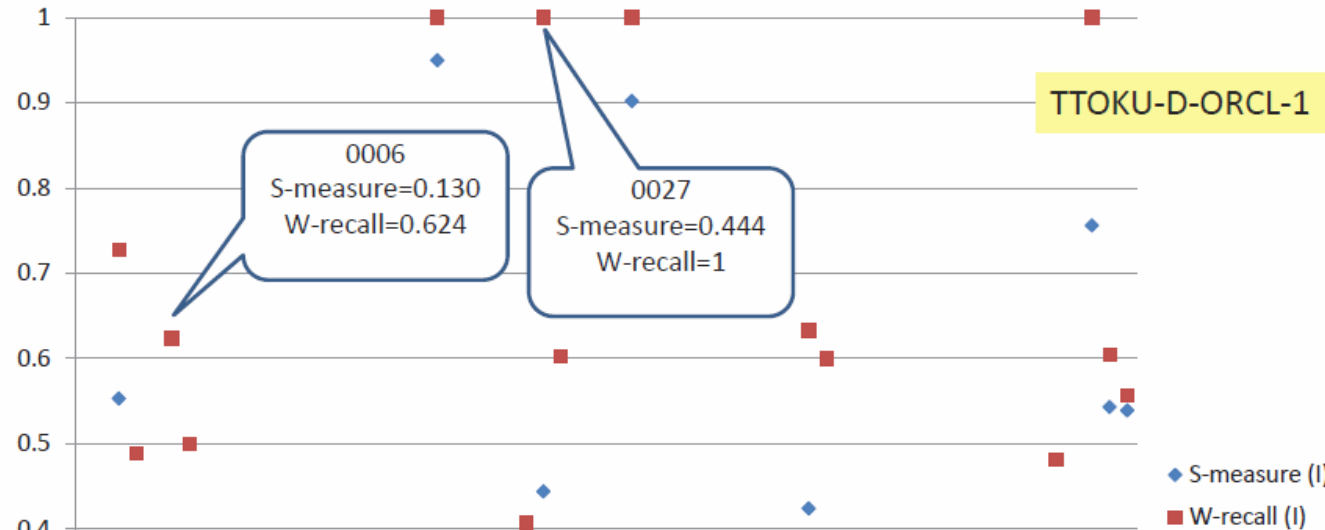
- MOTIVATION
- TASK
- METRICS
- **RESULTS**
- TENTATIVE CONCLUSIONS

System ranking (Mean over 60 queries) 19



- a: **KUIDL-D-OPEN-1**
- b: **MSRA1click-D-OPEN-2**
- c: **KUIDL-D-OPEN-2**
- d: **MSRA1click-D-OPEN-1**
- e: **KUIDL-M-OPEN-2**
- f: **KUIDL-M-OPEN-1**
- g: **TTOKU-D-ORCL-1**
- h: **TTOKU-D-ORCL-2**
- i: **TTOKU-M-ORCL-1**
- j: **TTOKU-M-ORCL-2**

Per-query disagreement: S vs W



Relevant text is at the very end of the X-string!

0006 LO 1 神戸市立中央図書館 *Kobe city central library*

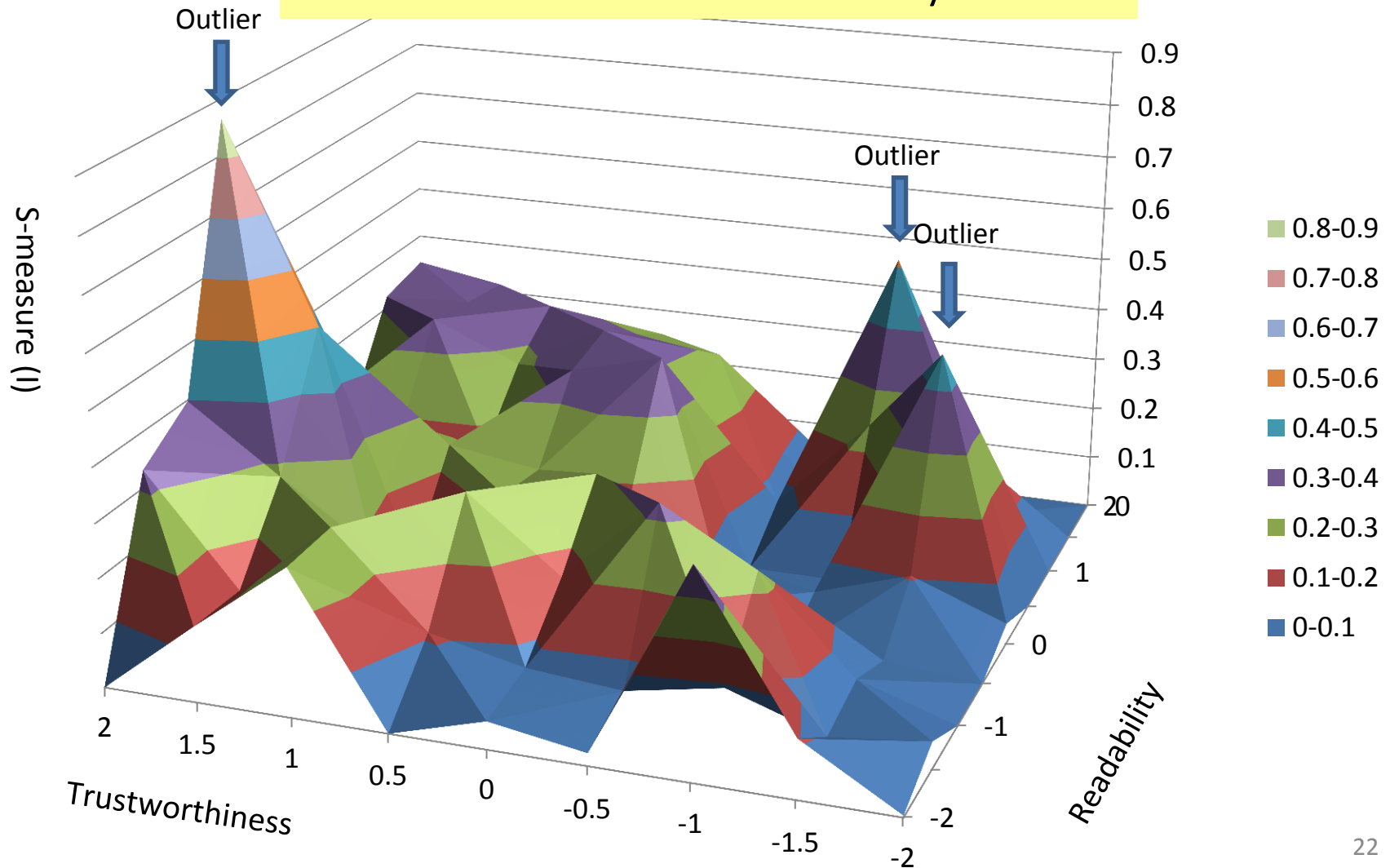
詳しい地図はこちらへ。中央図書館への道順・地図。館内整理日(年4回、2・6・10・12月の第3木曜)。サイトマップ神戸市ホームページのご利用案内プライバシーポリシー。トップページ利用案内施設案内・アクセス開館時間と休館日行事調べ物ガイド新聞・雑誌所蔵一覧神戸ふるさと文庫1.17文庫(震災関連資料)K117の本棚事業概要書燈(神戸市立図書館報)意見募集(パブリックコメント)図書館協議会アンケート東灘図書館移転整備子どものページおうちのかた・先生がたへえ小箱。ページの内容についてメールで問い合わせする。サイトマップお問い合わせよくある質問と回答。神戸市トップページへページの先頭へ。ページの作成責任者は、総務課長小林史明です。市政やくらし、イベント情報などのお問い合わせは、神戸市総合コールセンターまで。〒650-0017神戸市中央区楠町7丁目2-1中央体育館・文化ホール山側/大倉山公園内Tel:078-371-3351Fax:078-371-5046交通地下鉄大倉山駅北200m高速神戸駅北500mJR神戸駅北800m。毎週月曜(祝日・休日の場合は開館、直後の祝日・休日でない日を休館)。

N003 467 休館日 毎週月曜(祝日・休日の場合は開館、直後の祝日・休日でない日を休館)。毎週月曜だけではNG
 N007 353 住所 〒650-0017神戸市中央区楠町7丁目2-1
 N008 388 phone 078-371-3351
 N009 403 fax 078-371-5046
 N010 417 アクセス 地下鉄大倉山駅 北200m
 N011 427 アクセス 高速神戸駅 北500m
 N012 437 アクセス JR神戸駅 北800m

Nuggets representing opening hours, address, phone/fax, access from nearest station etc.

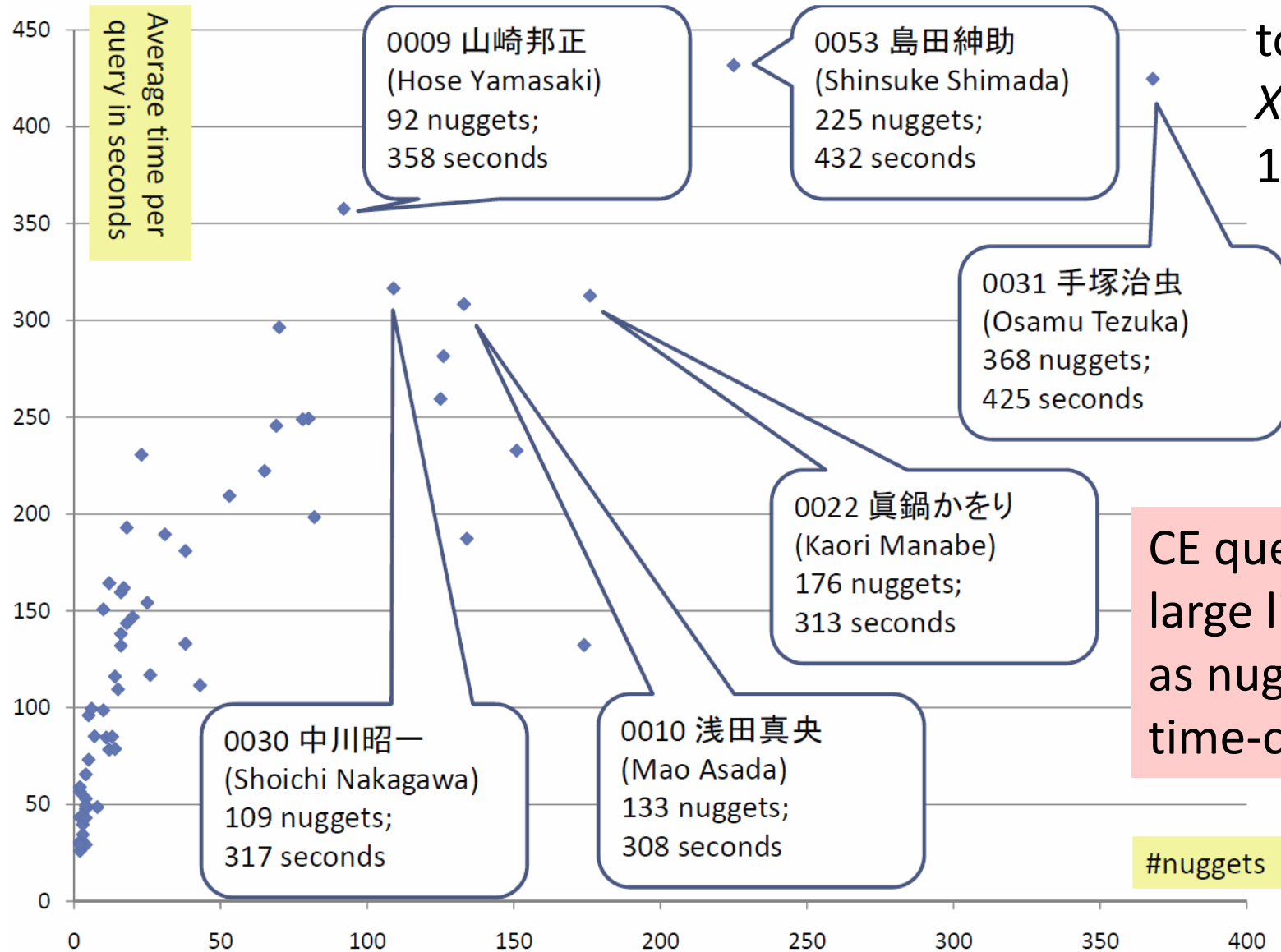
S vs readability vs trustworthiness

High trustworthiness AND readability -> high S
Low trustworthiness AND readability -> low S



Nugget match evaluation effort

Average time to handle one X-string:
151 secs



CE queries with large lists of titles as nuggets are time-consuming.

TALK OUTLINE

24

- MOTIVATION
- TASK
- METRICS
- RESULTS
- **TENTATIVE CONCLUSIONS**

Tentative Conclusions

- The 1CLICK evaluation framework is **useful!**
 - It can penalise systems that make the user read a lot of irrelevant text.
- The 1CLICK evaluation framework is **feasible!**
 - Average time to handle one *X*-string: **151** secs
 - The manual nugget match evaluation effort can be further reduced by carefully selecting queries and nugget types.

1CLICK@NTCIR-10?

- Japanese ->
Japanese, English, Korean (,Chinese)
- Separate training data and test data (esp. Yahoo! Chiebukuro QA queries)
- Select realistic queries and nugget types (e.g. avoid long lists of comic book titles)
- Refine nugget definition
- Refine nugget match criteria
- Handle query ambiguity?
- More query types
- Penalizing irrelevant information

Query Types

CELEBRITY User wants to gather various facts about a celebrity: date/place of birth, real name, blood type, height, hobbies, profession, personal history, awards, publications, discography, films, TV series, favourite baseball team, favourite food etc.

LOCAL User wants to contact or visit a facility (school, shop, office, amusement park, hotel, train station etc.). Hence (s)he wants facts such as postal and email addresses, phone and fax numbers, opening hours, how to access the facility by train/bus/car, nearest stations, time required for the travel, whether the facility has a car park and its opening hours etc.

DEFINITION User seeks the definition of a term.

QA User seeks a short answer to a question.

Readability and Trustworthiness

- **Readability** is to do with coherence and cohesiveness, and how easy it is for the user to read and understand the text. For example, garbled text and the lack of spaces between two unrelated contexts can hurt readability
- **Trustworthiness** means whether the user is likely to believe what it says in the *X*-string, as well as whether the user is likely to be misled. For example, if the *X*-string looks as if it was extracted from a source that is clearly not authoritative, it is not trustworthy. Moreover, if what is implied in the *X*-string is contrary to facts, it is not trustworthy.

