

Overview of the NTCIR-11 MobileClick Task

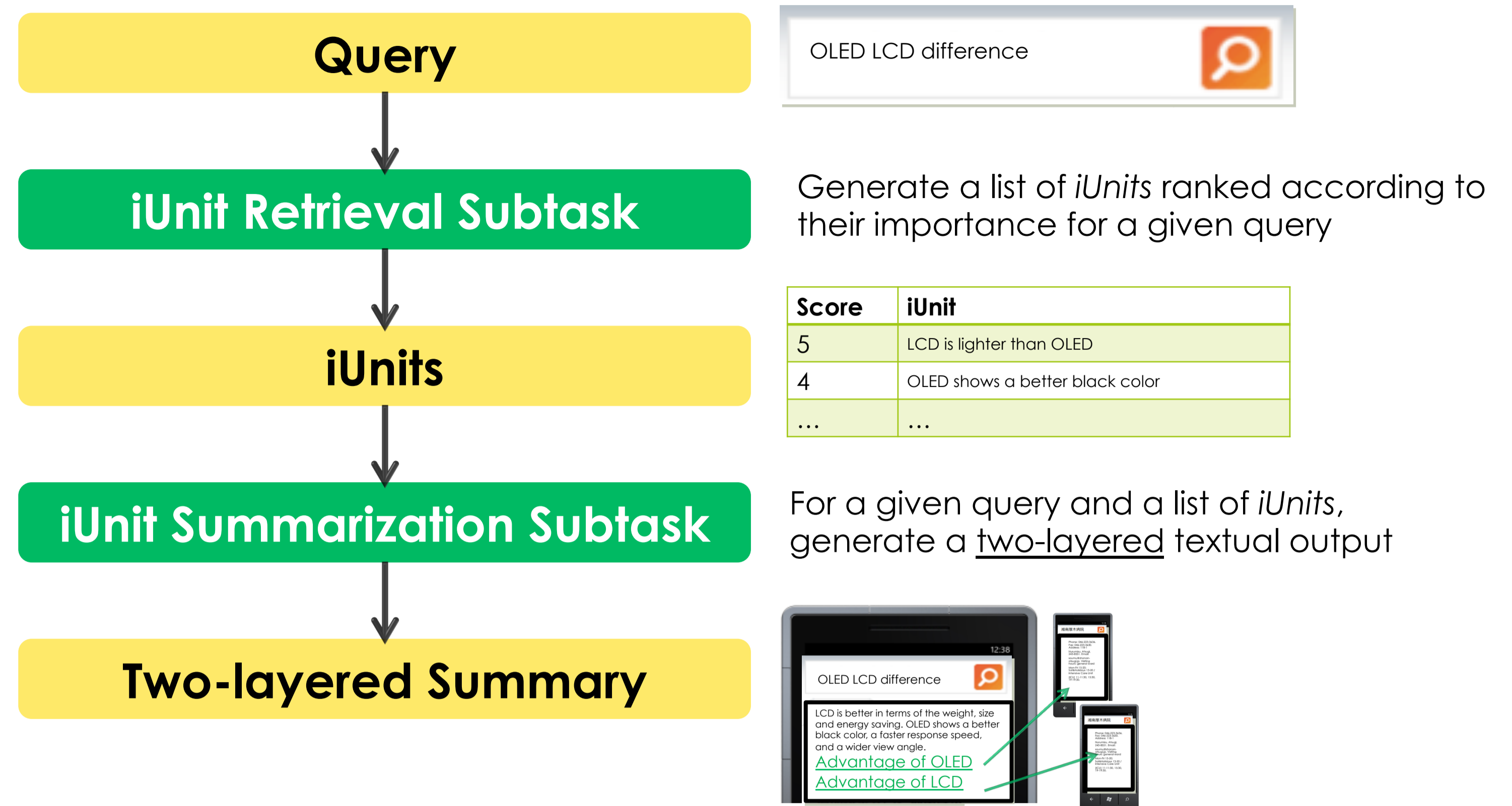
Makoto P. Kato (Kyoto University), Matthew Ekstrand-Abueg, Virgil Pavlu, (Northeastern University)
Tetsuya Sakai (Waseda University), Takehiro Yamamoto (Kyoto University), Mayu Iwata (KDDI Corporation)

Task: Return a concise summary,
Satisfy information need of mobile users



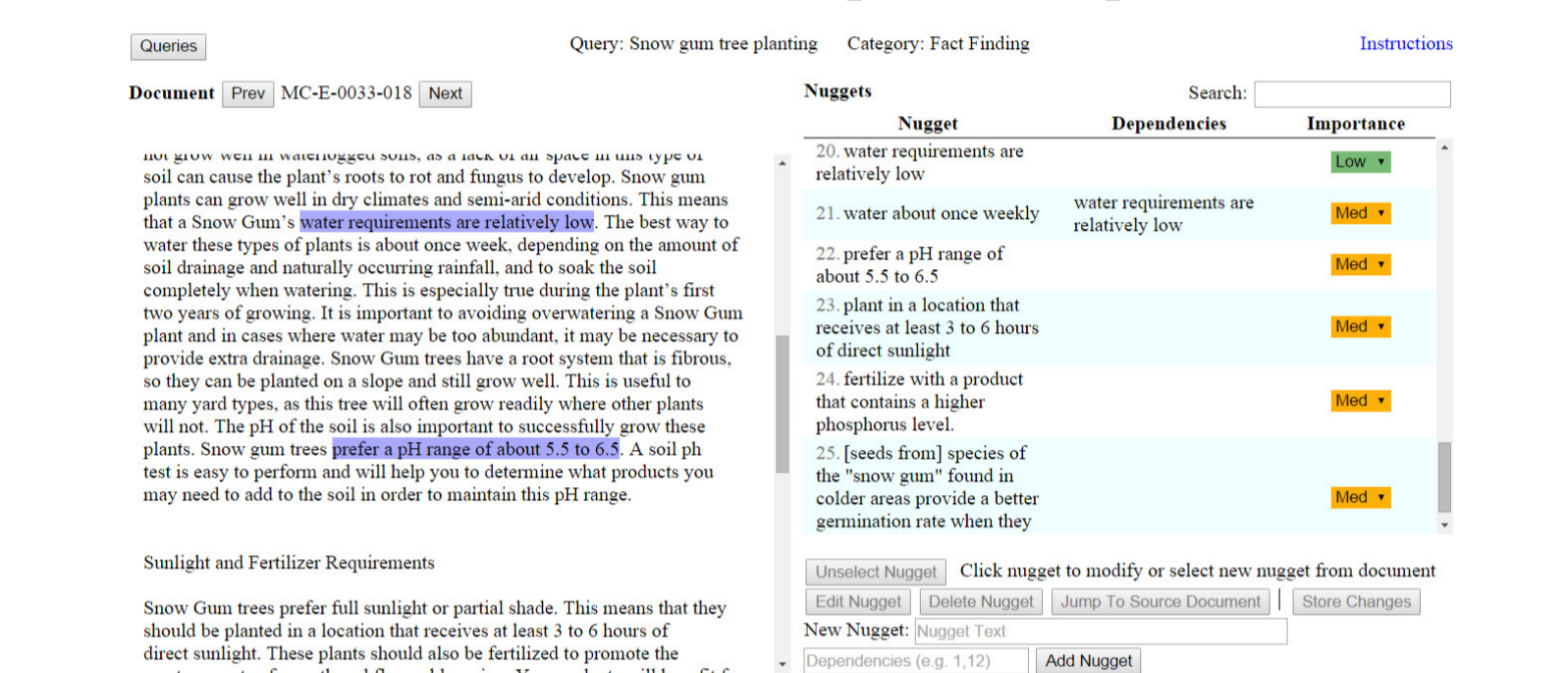
Task: Given a search query, return a **two-layered** textual output

Subtasks



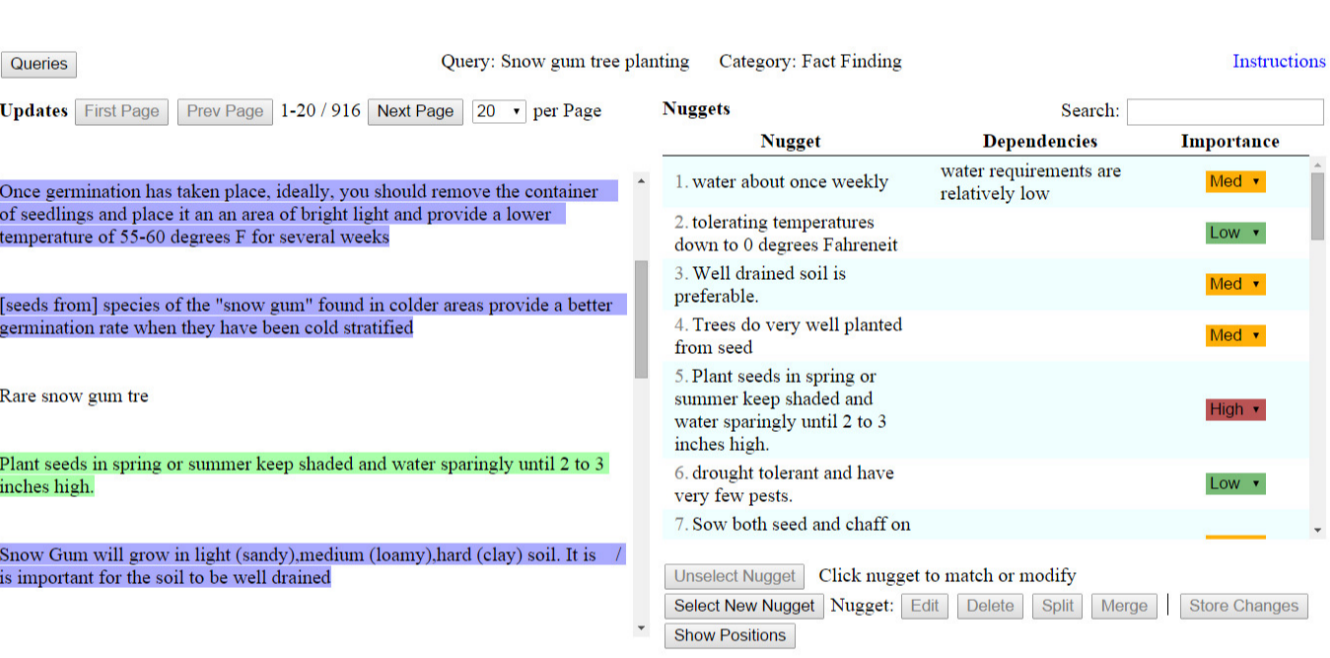
Task Flowchart:

Gold Standard *iUnit* (GiUnit) Extraction

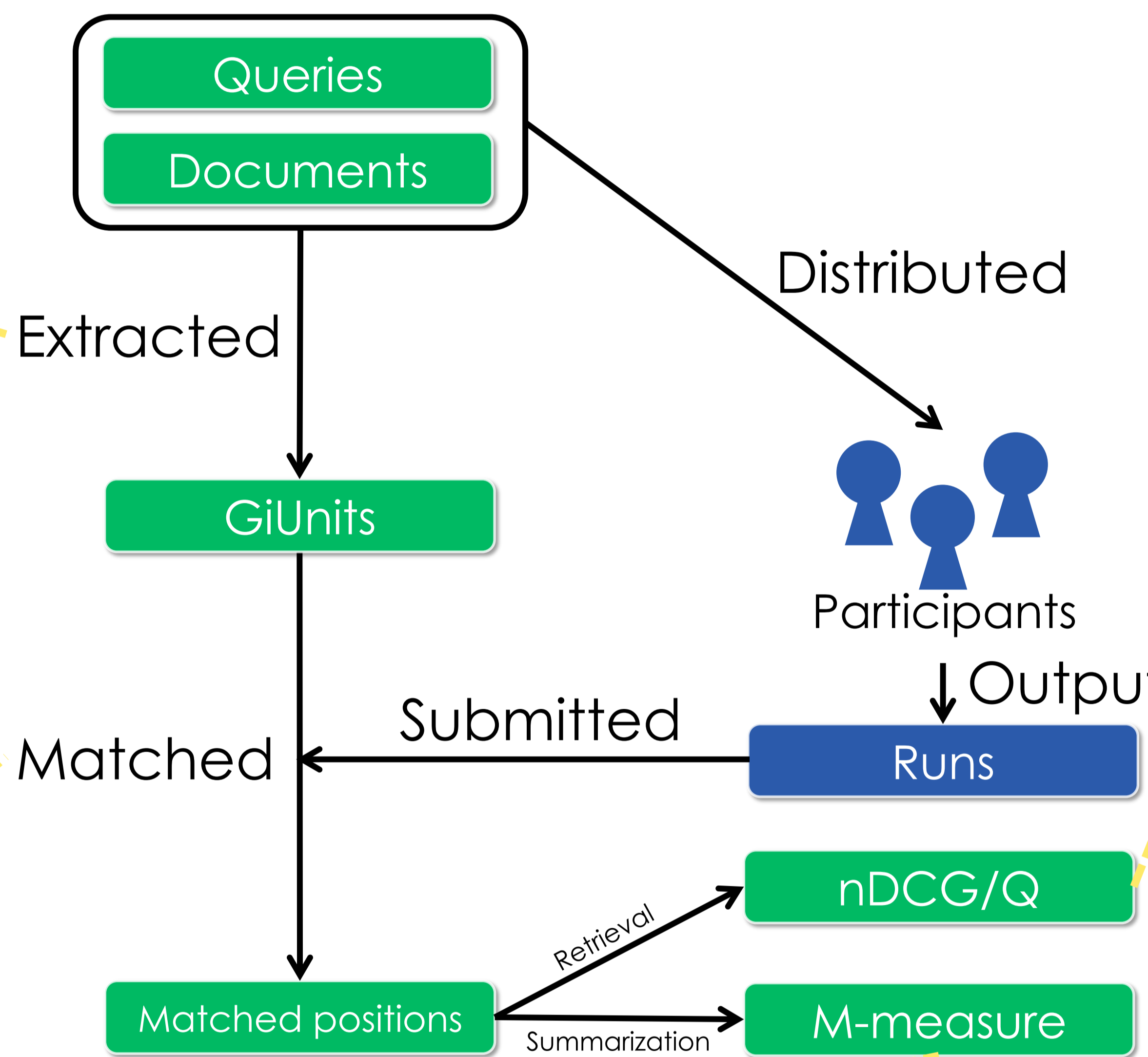


***iUnits* are atomic** (cannot be broken down further), **relevant** (useful to satisfy the information need), and **dependent** (may depend on others) info. pieces

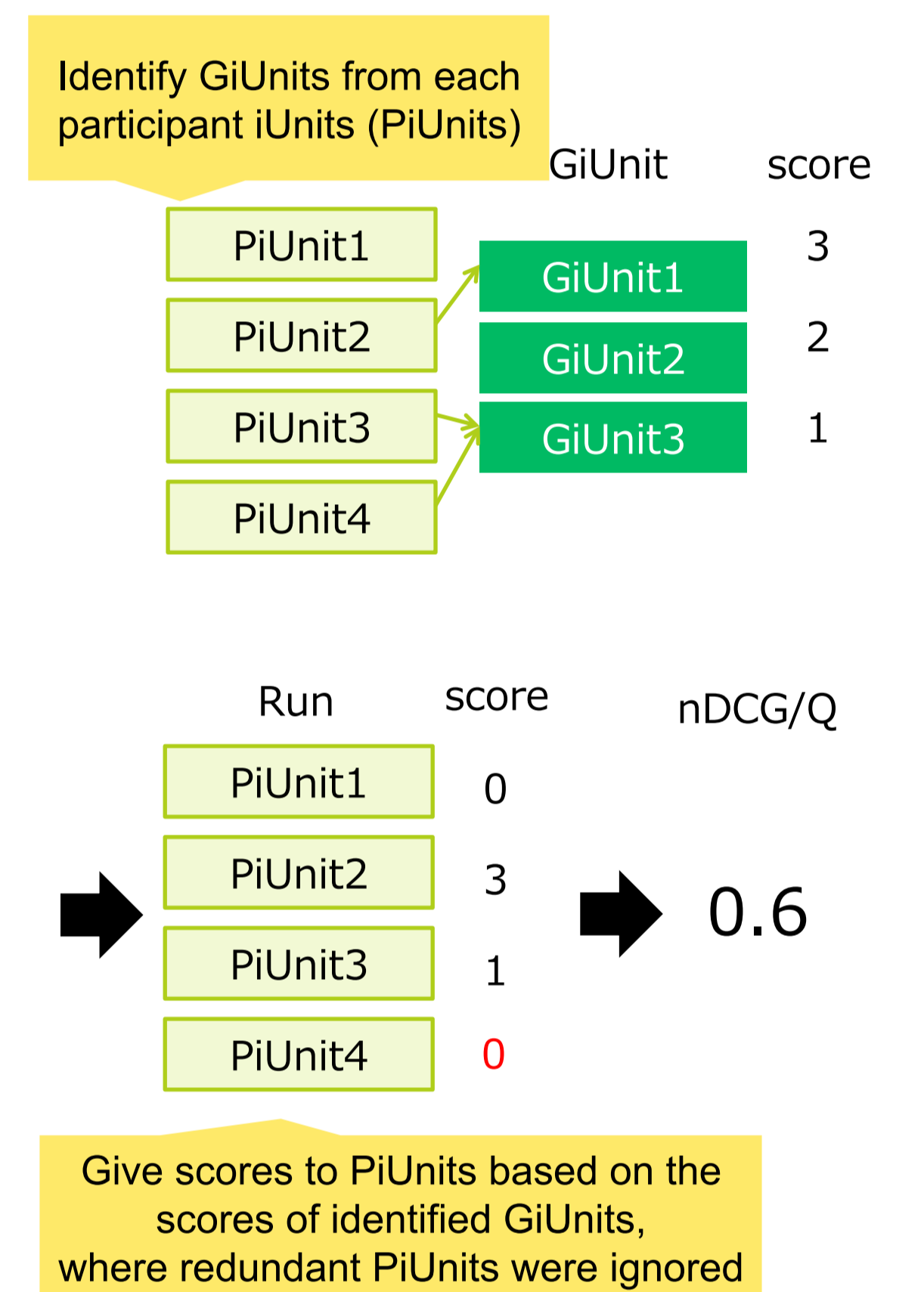
iUnit Matching Interface



Assessors can report which parts within the output match each *GiUnit* by text highlighting with a mouse



Evaluation of *iUnit* Retrieval



Evaluation of *iUnit* Summarization

M-measure = the expected utility of users who probabilistically click on links

$$M = \sum_{t \in T} P(t)U(t)$$

T: all possible trailtexts (text read by users)
U: expected utility of a trailtext
P: probability of users' reading a trailtext

Expected utility = **U-measure**

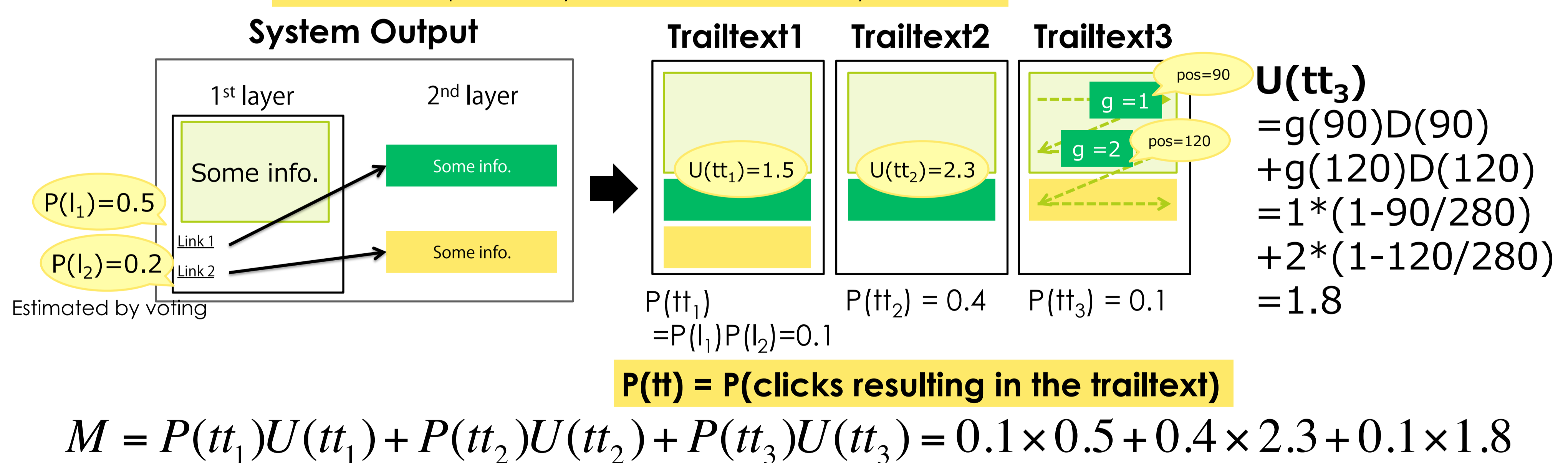
Idea: "More important *GiUnits* earlier" is better

$$U = \sum_{pos=1}^{l(t)} g(pos)D(pos)$$

D(*pos*) = $1 - \frac{pos}{L}$

- pos*: position in text
- g*(*pos*): importance of *GiUnit* at *pos*
- D*(*pos*): decay function = how much effort is required to reach *pos*
- L*: patience parameter (e.g. 280)

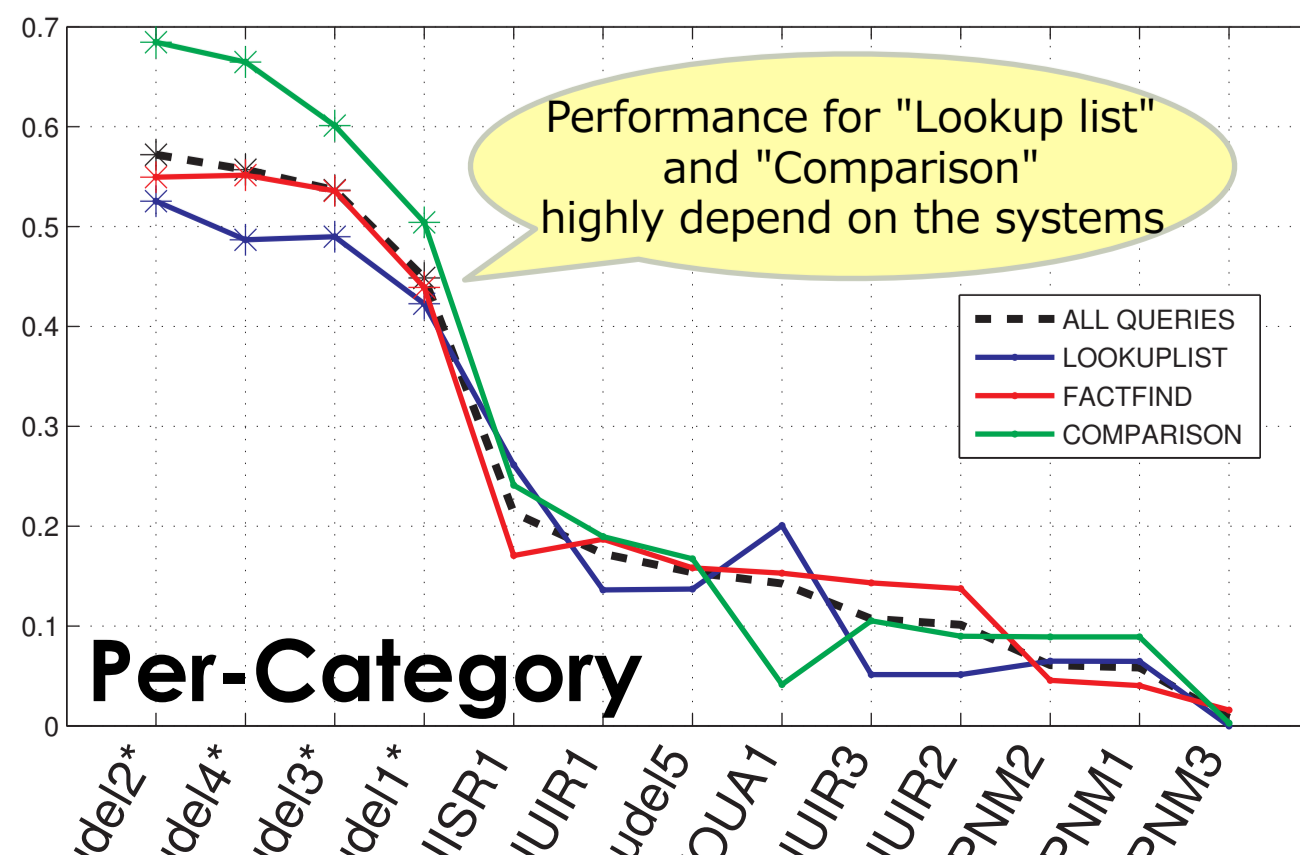
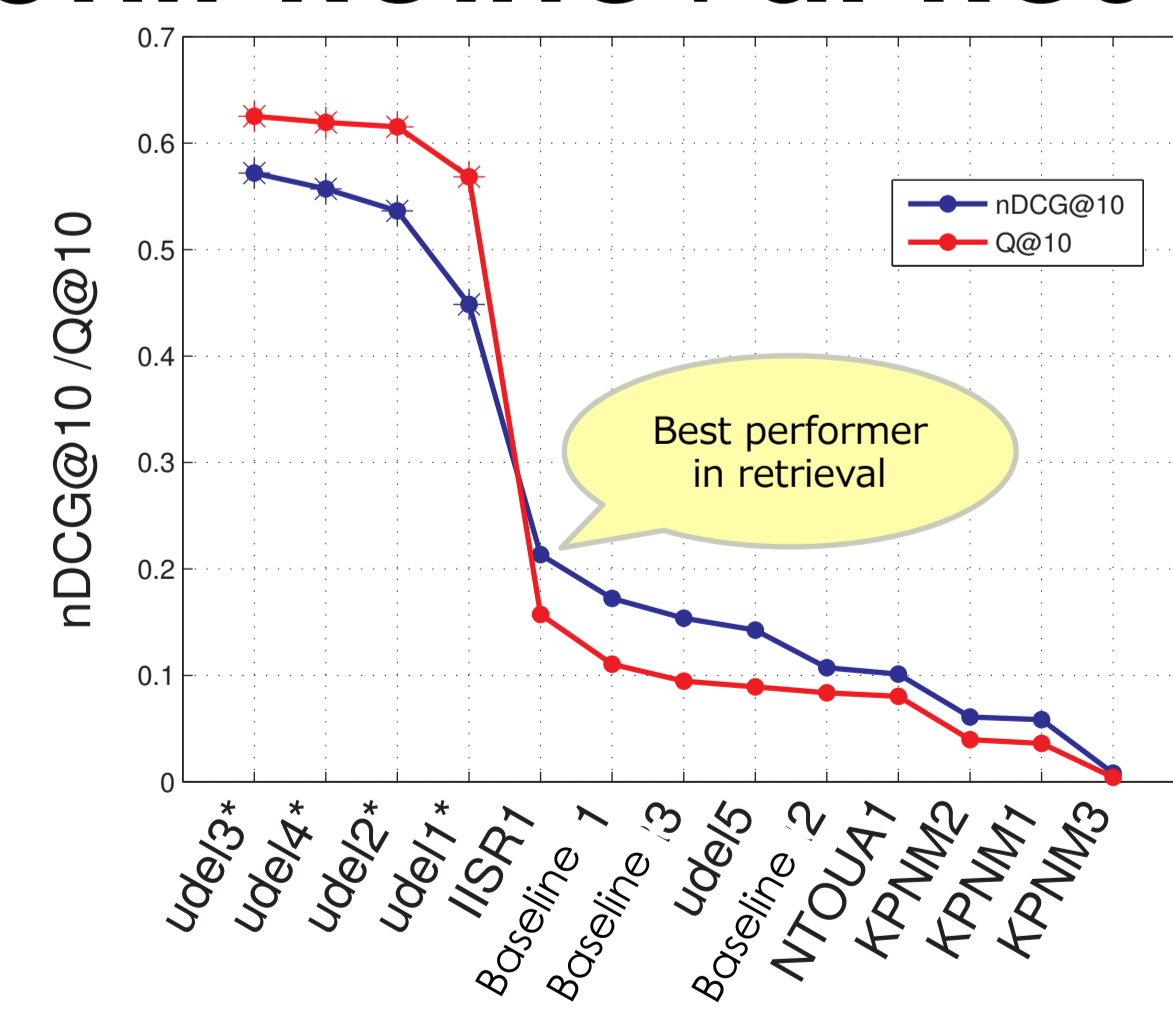
Consider possible reading paths for a two-layered summary, estimate their probability and evaluate their utility



Evaluation Results:

Team name	Organization/Method
KPNM	Hunan University of Science and Technology, China [Retrieval] Chain simple techniques based on statistical models and heuristic rules to extract significant text units
IISR	National Central University, Taiwan [Retrieval] Classify each query into eight query types and set the weights of the extraction methods accordingly
udel	University of Delaware, USA [Summarization] Simple re-ranking approach based on the cosine similarity between each <i>iUnit</i> and a dynamic "model" pseudo-document
NTOUA	National Taiwan Ocean University, Taiwan [Summarization] Grouping by longest leading substring.

iUnit Retrieval Results



iUnit Summarization Results

