# YouUndestood.me

YUM is an online environment built around a search engine, which aims to make the search process valuable for children. Yum is not meant to be treated as a new, child-oriented search engine, since studies \cite{ref} show that children tend to prefer popular search engines, such as Google, to perform their online information-seeking tasks. Instead, YUM acts as an intermediate layer between the child and an existing search engine, in order to facilitate the interaction between the two of them. For doing so, YUM puts into practice strategies oriented to address issues children face when using popular search engines, as well as strategies that can enhance the search experience to foster learning. A description of the mentioned strategies is provided below.

## Search intent

Children are not always able to formulate succinct queries \cite{Bil11}. Studies show that they tend to write natural language queries, instead of short, keyword-based ones that search engines usually expect \cite{Dru09}. Unfortunately, the longer the query, the less likely it is for a search engine to retrieve relevant results in response to it, making children unable to successfully complete information seeking tasks\cite{Dru09}. In addition, children also tend to misspell words, but not necessarily in the same fashion as average users. For example, children commonly repeat letters in a word to emphasize it, such as in “amaaaaaaaaazing”, which can cause search engines to misunderstand the intended meaning of the word. In order to best satisfy children needs Yum takes advantage of QuIK\cite{Sven}, a search intent module designed specifically for children. QuIK addresses common patterns in queries written by children including, but not limited to: diminutives, emphasis, children trendy terms or children specific misspellings. In doing so, QuIK transforms an initial child-query into a new keyword-based query that captures the information expressed by the child in a way that can more easily be comprehended by search engines.

## Query suggestion

Even if the search intent module identifies the most likely user intent for each query, users have different interests and information needs, which is why when dealing with ambiguous queries, it is only each specific user who knows the purpose of his search. With this is mind, Yum takes advantage of ReQuIK \cite{requik}, a query recommender that is specifically tailored to children, in order to provide alternatives for the initial query, based on the goal of his respective information seeking task, to better inform the search process. ReQuIK is a multi-criteria recommendation system based on traits commonly associated with children that suggests queries that (i) are associated with children topics, (ii) lead to the retrieval of resources with levels of readability matching those of the K-12 audience, and (iii) are diverse enough to capture the different topics children can be interested in.

## Filtering by readability

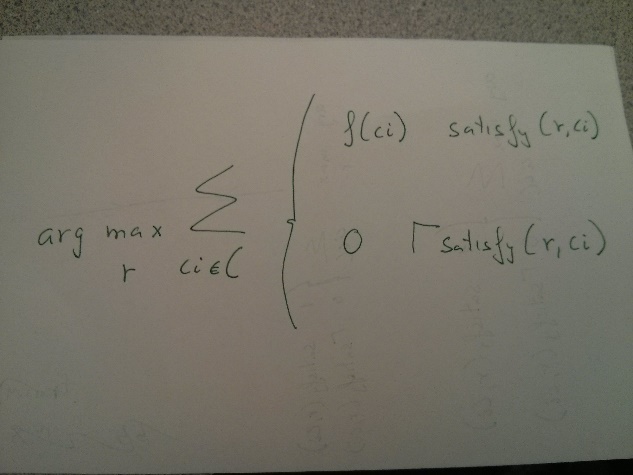
Even when the search engine has understood the intent of a child query and retrieves results that match the information needs expressed in the corresponding users, the suitability of the retrieved resources is still not assured. K-12 students tend to find difficult to understand documents containing complex or technical vocabulary. For example, in the case where a child is looking for information about chemistry, retrieving a scientific publication would not be adequate, while retrieving information from an elementary chemistry book would. If the retrieved documents are too complex, children may not succeed in completing their information discovery tasks. In order to avoid this situation, Yum incorporates as part of its environment a filtering strategy based on readability levels. This strategy ensures that the retrieved documents will match, to a degree, the reading ability of each individual user. Yum allows users to go through a one-time process where they can select their grade level. This grade level determines their readability level, which is matched against the one of documents retrieved by the search engine, eliminating the ones that are not within half a grade level above or below the grade of the corresponding student. For estimating the readability of retrieved resources, Yum uses the Flesh-Kincaid \ref{Fle48} (see formula XX) readability formula. We initially selected this formula given that it is widely used for measuring the readability levels or complexity of texts among educators nationwide and is considered as a standard by institutions\ref{porposal31\_38\_40} to measure readability levels. However, we are aware that it is only based on simple text-based features, which is why we plan to improve this formula in the future (see Future Work Section).

Placeholder for fleshkincaid

## Tracking

K-12 students have diverse reading abilities. Even in same grade class, students’ readability skills can differ. Furthermore, the reading skills of each individual progressively improves over time\cite{sh13}. Consequently, a one-size-fits-all strategy is not applicable for conducting successful information-seeking tasks that lead to the retrieval of resources individual users can read and understand. Instead, Yum employs an adaptive strategy that allows its users to provide optional feedback specifying whether the resource was too “*easy”*, “*OK”* or “too *complex”* for them. This feedback is used to determine and update the reading skills of each user so that the system can retrieve documents adequate to their current level of readability.

The problem of predicting the current readability level of each student can be treated as a constraint satisfaction problem, where each feedback **f** provided by a student **s** generates a new constraint that needs to be satisfied by the readability of **s.** For example, a student **s** giving a feedback of “too hard” to a document of readability level 5 would generate the constraint readability(**s**) < 5 stating that **s**’ readability should be lower than 5. As showed in Equation XX, the predicted readability **r** is the one that maximizes the amount of constraints satisfied. The equation also considers the time when each constraint was created, giving more importance to constraints that were created more recently and discarding ones created outside the academic year. For doing so, f(Ci) is a function that starts at value 9 for new constraints and decreases by 1 for each month that has passed until 0. We selected 9 as the number of months to consider as this represents the average length of an academic year. In case of multiple maximum values, the one that has the biggest distance to its two (lower and upper) closest constraints is selected



where **r** is the readability value, **ci** is an individual constraint among the set C of constraints created based on the feedback provided on retrieved resources by a user **s**, and **f(ci)** refers to how new the constraint is, starting at 9 is the constraint was created at this month, and getting reduced by one if for every month, ending in 0 if the constraint was created outside the academic year.

At the beginning of each academic year or when the first time the environment is used, the constraints of every student’s accounts are initialized with two constraints based on the grade level of the student or the readability level prediction on the system for the previous academic year. Those constraints are representative of one grade of deviation from the current grade of the student: readability(**s**) < grade(**s**)+0.5 and readability(**s**) > grade(**s**) -0.5. These constraints give a starting point to Yum, that will eventually be better adjusted when the student start using the environment.

# Yum for teachers

Teachers can also benefit from using Yum within the class environment. Work setting standards have changed from a vertical structure, where only the top individuals of the pyramid had to think critically and the lower parts just followed directions, to an horizontal structure, where each individual is expected to collaborate with others and solve important problems using identification, searching, synthetizing, and communication skills \cite{leu13}. Given this change, education plans oriented to meet the new requirements of the current industry, such as the Common Core State Standards (CCSS) Initiative, have been developed. CCSS requests educators to make an emphasis on higher level thinking during reading and writing and focus on the acquisition of skills such as research and comprehension using digital tools, including search engines\cite{leu13}. Furthermore, educational studies \cite{kni15} showcase the benefit of in-class exercises such as exploratory talks, where students are asked to solve a problem in groups discussing information found on resources obtained using a search engine. Unfortunately, teachers might not be able to propose such a task to their students and lead critical discussions, if students have problems using search engines, whether that be struggling to find the right queries or not being able to understand the retrieved documents due to their complexity. YUm can help teachers overcome those issues so that they can instead focus on the discussion, rather than on the manner in which students should formulate queries or the type of results they access. Furthermore, YUm can serve as a monitoring tool that allows teachers to check students’ progress, based on the resources they have retrieved and their provided feedback in terms of complexity. We believe that YUm can not only facilitate learning when children use it for their information discovery assignments at home, but it will also help teachers within the classroom environment, by addressing the challenge of seamlessly integrating technology to perform everyday classroom activities\cite{Dan13,kni15}.

@article{sh13,

title={Exploring gains in reading and mathematics achievement among regular and exceptional students using growth curve modeling},

author={Shin, Tacksoo and Davison, Mark L and Long, Jeffrey D and Chan, Chi-Keung and Heistad, David},

journal={Learning and Individual Differences},

volume={23},

pages={92--100},

year={2013},

publisher={Elsevier}

}

@article{leu13,

title={The new literacies of online research and comprehension: Assessing and preparing students for the 21st century with Common Core State Standards},

author={Leu, Donald J and Forzani, Elena and Burlingame, Cheryl and Kulikowich, Jonna and Sedransk, Nell and Coiro, Julie and Kennedy, Clint},

journal={Quality reading instruction in the age of common core standards},

pages={219--236},

year={2013}

}

@article{kni15,

title={The role of exploratory talk in classroom search engine tasks},

author={Knight, Simon and Mercer, Neil},

journal={Technology, Pedagogy and Education},

volume={24},

number={3},

pages={303--319},

year={2015},

publisher={Taylor \& Francis}

}

@incollection{Bil11,

title={Evaluating leading Web search engines on children's queries},

author={Bilal, Dania and Ellis, Rebekah},

booktitle={Human-Computer Interaction. Users and Applications},

pages={549--558},

year={2011},

publisher={Springer}

}

@inproceedings{Dru09,

title={How children search the internet with keyword interfaces},

author={Druin, Allison and Foss, Elizabeth and Hatley, Leshell and Golub, Evan and Guha, Mona Leigh and Fails, Jerry and Hutchinson, Hilary},

booktitle={SIGCHI},

pages={89--96},

year={2009},

organization={ACM}

}

@inproceedings{Sven,

title={``Is Sven Seven?": A Search Intent Module for Children},

author={Dragovic, Nevena and Madrazo, Ion and Pera, Maria Soledad},

booktitle={ACM SIGIR},

pages={},

year={2016}

}

@article{Fle48,

title={A new readability yardstick.},

author={Flesch, Rudolph},

journal={Journal of Applied Psychology},

volume={32},

number={3},

pages={221},

year={1948},

publisher={American Psychological Association}

}