Simpler Voice

A Key Message & Visual Description Generator System for Illiteracy

Minh N.B. Nguyen,[1] Samuel Thomas,[2] Anne E. Gattiker,[3] Sujatha Kashyap,[3] Kush R. Varshney[4]
University of Southern California[1], IBM Watson[2], IBM Research[3], IBM Research AI[4]
***Partner Organization*: Literacy Coalition of Central Texas**

Introduction

- **The social issues of illiteracy**
  - Nearly 1 in 5 adults worldwide cannot read this sentence.
  - Difficulty in comprehending complex texts.
- **Exceptional development in Data Science**
  - Nature language processing
  - Artificial intelligence
- **Proposed system SIMPLER VOICE:**
  - Decoding complex texts
  - Simple key messages
  - Object2Text, Text2Visual

Technical Challenges

- Semantic parsing / semantic analysis from unknown word to image queries
- Word-sense-disambiguation, image-sense disambiguation mechanism, optimal visual components
- Linking words synsets to big datasets of descriptive images [3]
- Ranking queries & evaluating metrics [2,3]

Case Study & Results

- **Case Study: Grocery shopping**
  - Identifying how to use a product.
  - Encouraging customers to try new products.
  - Other products info (eg. Warning, allergy, etc.)
- **System Demonstration**
  - *Input:*

    Grocery product barcode (eg. 3700098084)

  - *Generated Queries:*

    [Subject] + [Verb] + [Category]

    > Woman / Man washing with Dishwasher Detergent

    > Woman / Man washing with auto dish care

  - *Final Results:*

Category Tree of H-E-B Bakery Cookies by the Pound

Categories:
H-E-B: [['cookies', 'desserts', 'bakery', 'food-and-drinks']]
Parent: ['desserts']
Neighbor: ['brownies', 'pies']

(1)

Query of H-E-B Bakery Cookies by the Pound

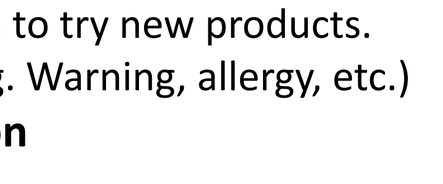
First Texts
Spanish: Mujer comiendo galletas

Level of Reading:

(2)

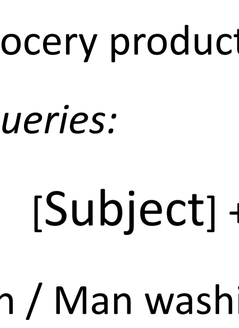
1. cookies
2. eating cookies
3. Woman eating cookies

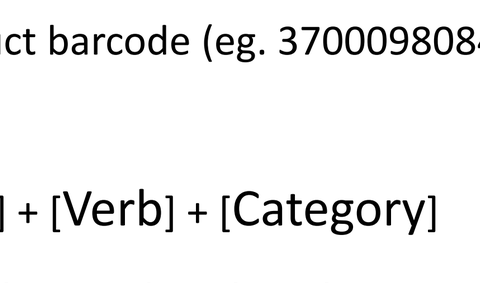
(3a)



(3b)

- *Mobile App:*

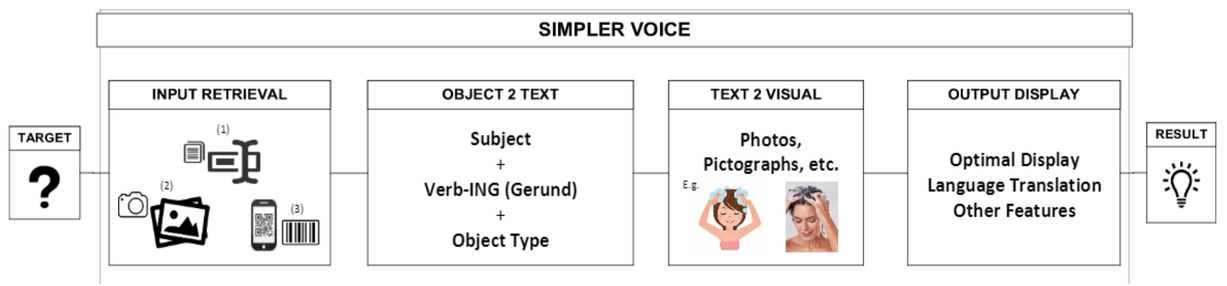




Related Research

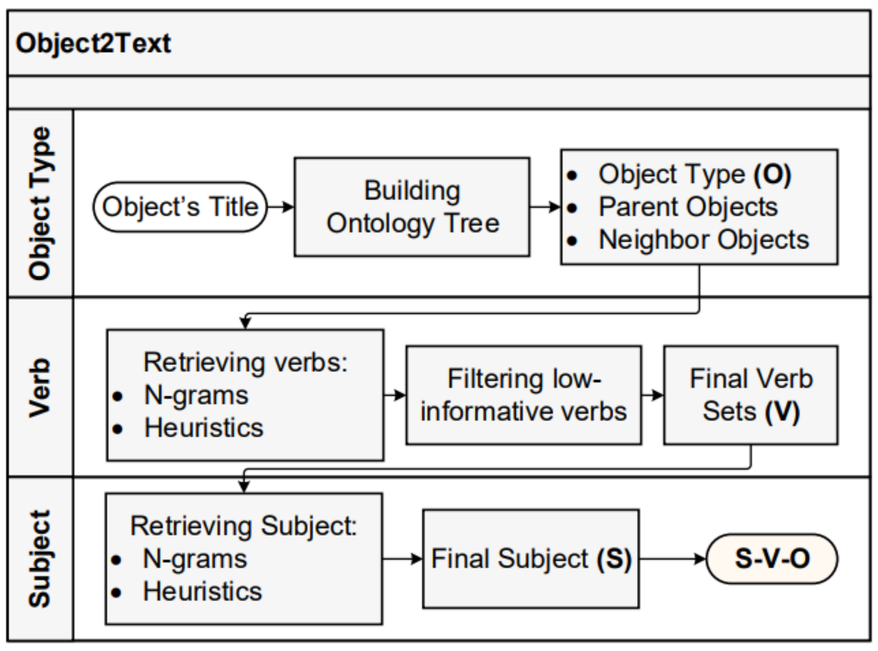
[1] Google books: https://books.google.com/ngrams/
[2] F. Schroff et al. "Harvesting Image Databases from the Web". IEEE transactions on pattern analysis and machine intelligence (2011)
[3] V. Vandeghinste et al. "Translating Text into Pictographs". Natural Language Engineering (2015)

System Architecture

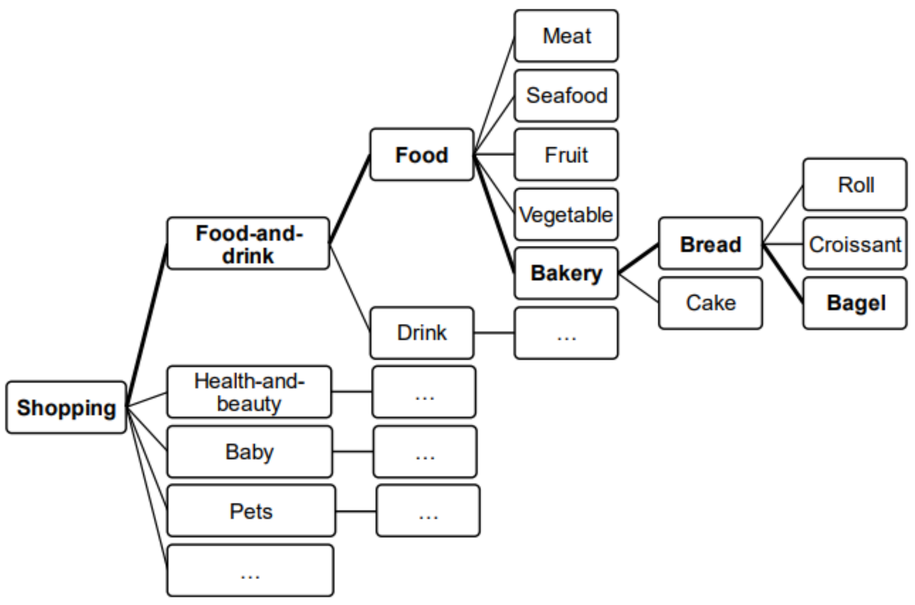


- **Input Retrieval**
  - Unknown words
  - Eg. A Product name, a barcode, etc.

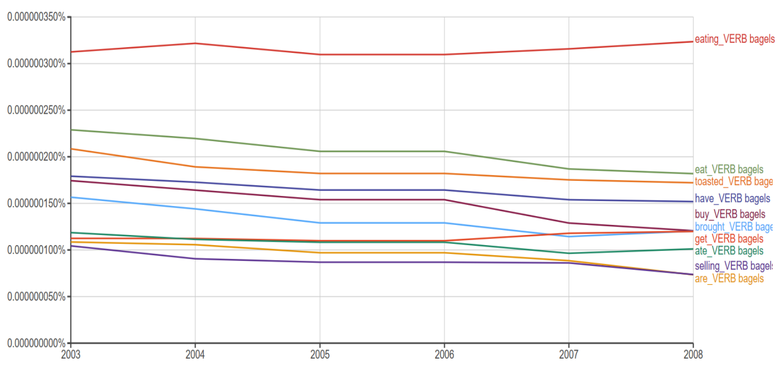
- **Object2Text** SUBJECT + VERB (-ing) + (With) + CATEGORIES

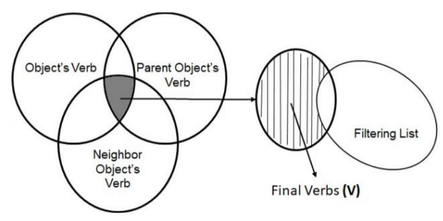


- *Objects:* Building products category tree – E.g.: "Bagel"

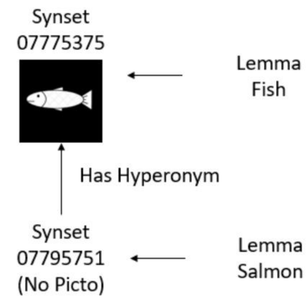


- *Verbs + Subjects:* n-grams[1], word-sense disambiguation, low-informative words filtering





- **Text2Visual**
  - Photorealistics image + Pictograph
    - Image sense ambiguity
    - Optimal visual component



IBM Research | LITERACY COALITION OF CENTRAL TEXAS BUILDING STRONGER COMMUNITIES | USC | WiML Women in Machine Learning

IBM Science for Social Good

WiML - NIPS 2017
December 2017, CA, USA