A Food Logging System for iOS with Natural Spoken Language Meal Descriptions

Mandy Korpusik,[1] **Salima F. Taylor**,[2] Sai Das,[2] Cheryl Gilhooly,[2] Susan B. Roberts,[2] James Glass[1]

[1]MIT Computer Science and Artificial Intelligence Lab, Cambridge MA USA,

[2]Tufts University Jean Mayer USDA Human Nutrition Research Center on Aging, Boston MA USA

OBJECTIVE

Introduction: This study presents the design and implementation of *Coco: The Conversational Calorie Counter*, a spoken food logging application for iOS. The aim of this work is to reduce the burden on individuals wanting to monitor their food to support healthy eating, as well as individuals with obesity tracking food intake to manage their weight.

Problem: existing methods are too tedious.
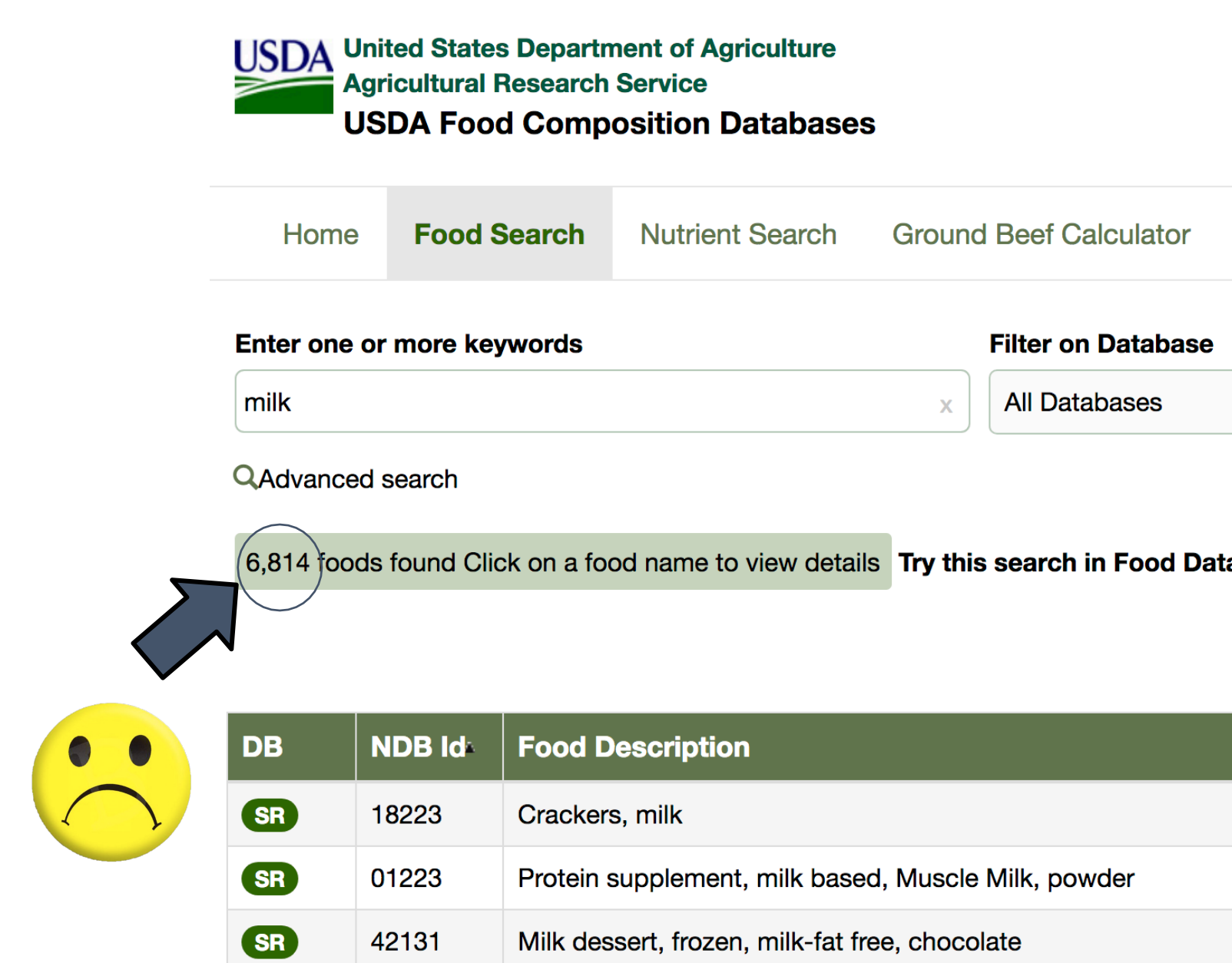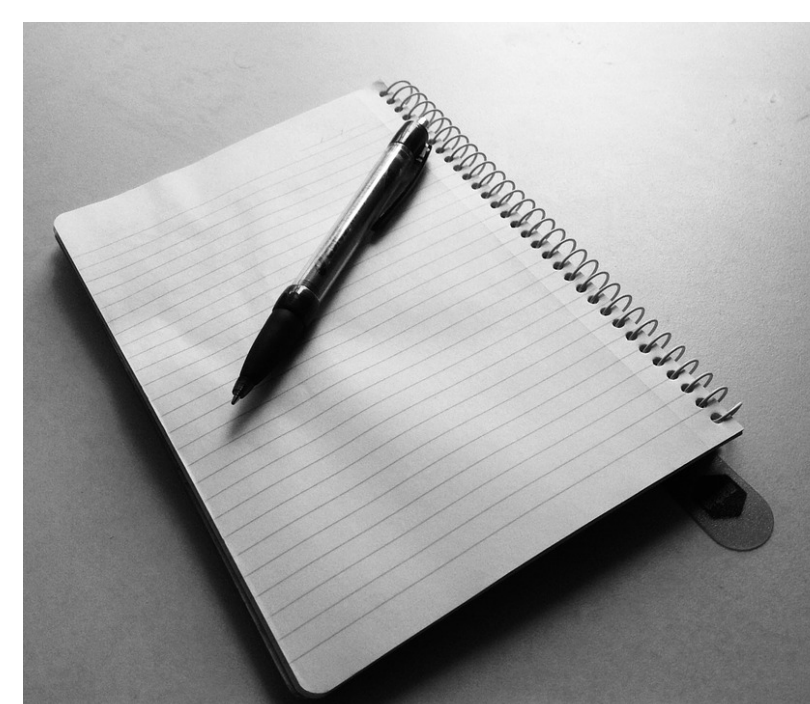
Goal: allow people to describe their meal using natural spoken language understanding technology.

COCO *the conversational calorie counter*

for dinner I had a bowl of **chili** over **rice** and an **apple**

chili with beans canned
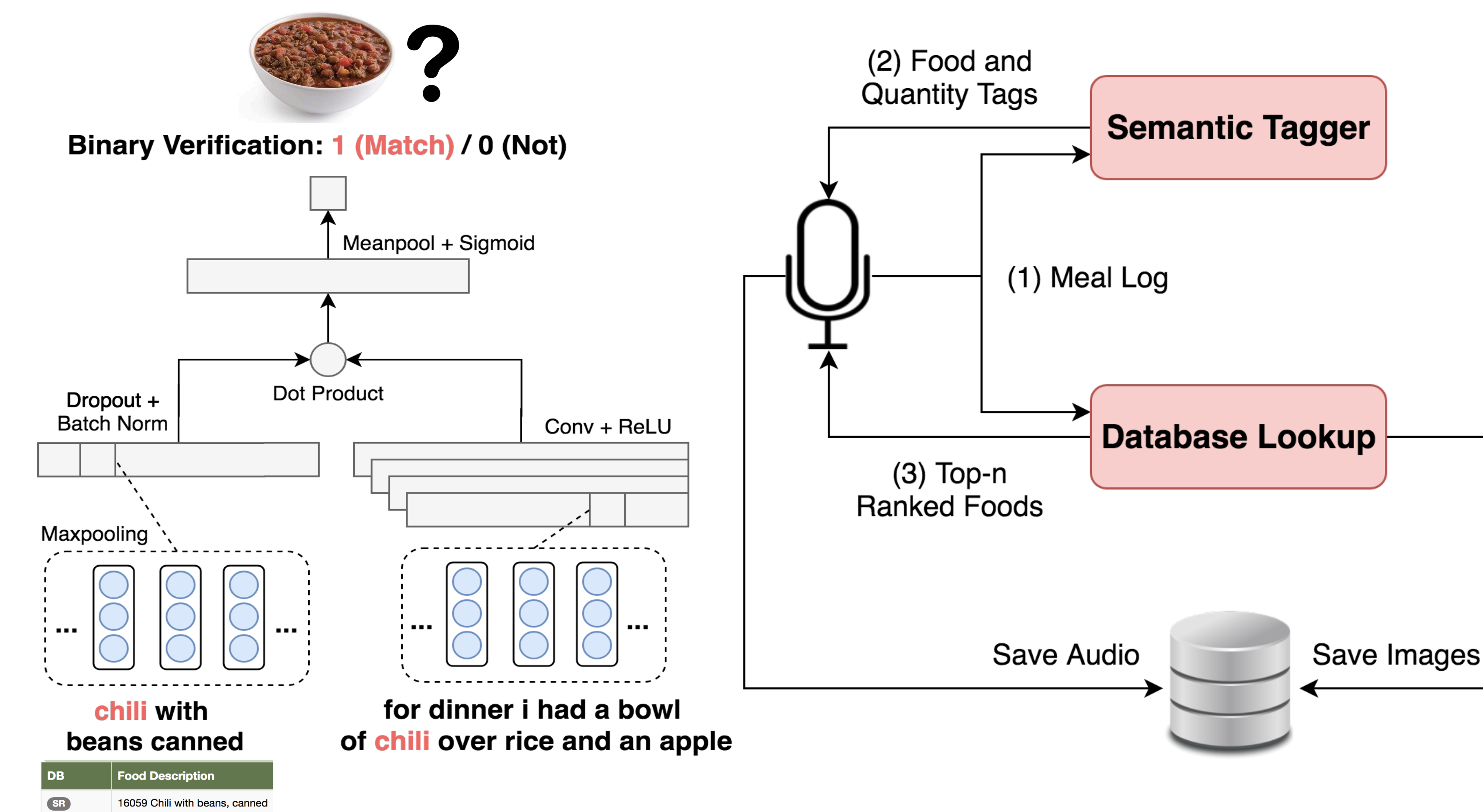
rice white short-grain cooked

apples raw with skin

METHODS

Approach: a convolutional neural network tags foods and quantities in a user's natural language meal description, and selects matching foods in the most recent USDA database.

Binary Verification: 1 (Match) / 0 (Not)

Meanpool + Sigmoid

Dropout + Batch Norm

Dot Product

Conv + ReLU

Maxpooling

chili with beans canned

for dinner i had a bowl of **chili** over rice and an apple

DB   Food Description

16059 Chili with beans, canned

(2) Food and Quantity Tags

Semantic Tagger

(1) Meal Log

Database Lookup

(3) Top-n Ranked Foods

Save Audio   Save Images

Evaluation: We measured the accuracy of our system via: 1) Amazon Mechanical Turk (shown below), 2) a pilot study with 14 participants, and 3) launching Coco to the App Store.
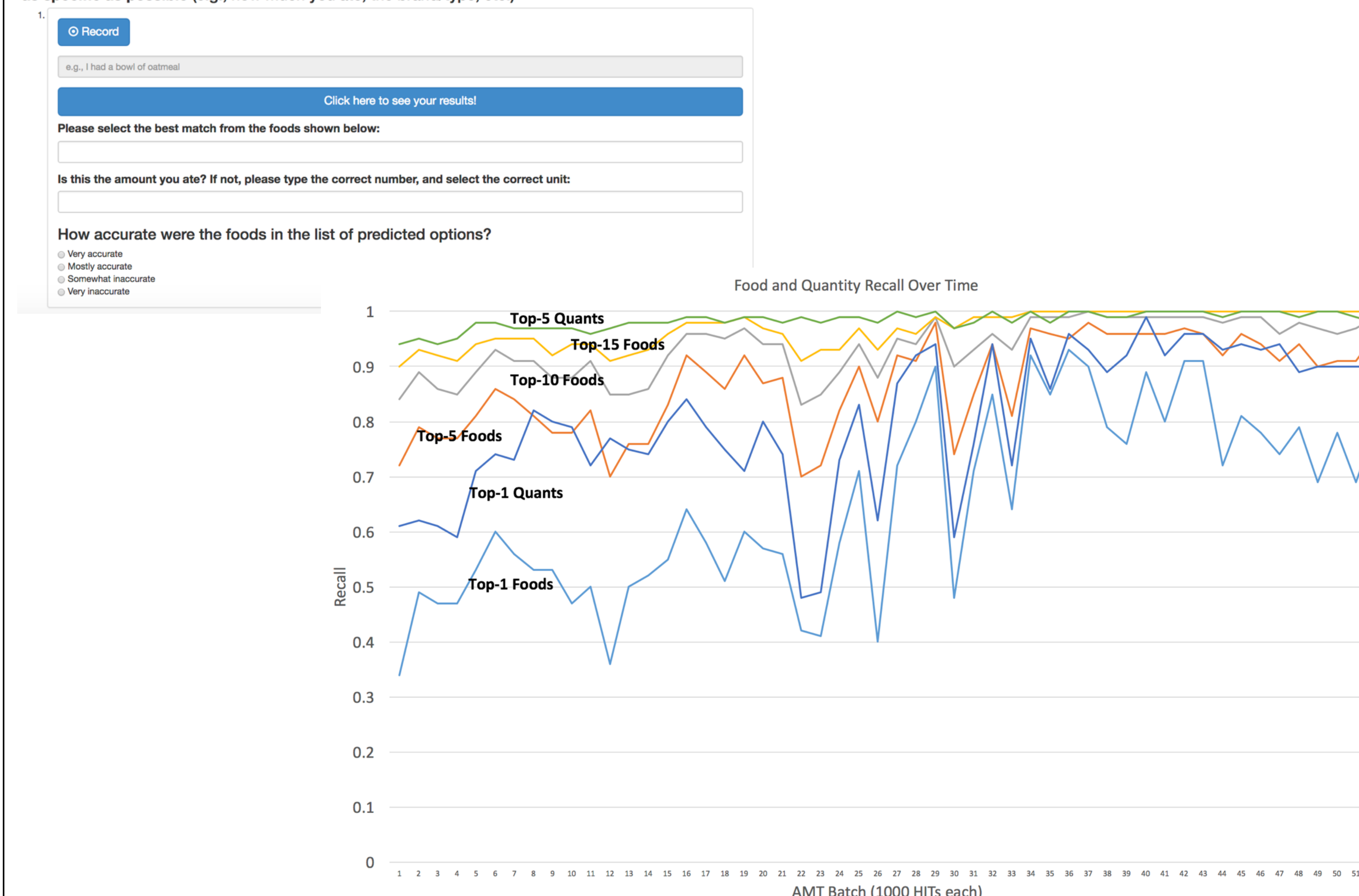
RESULTS

Table 1 Statistics from Coco App Store downloads.

Top-1 Food Recall	93.0%
Top-5 Food Recall	97.3%
Top-1 Quantity Recall	86.9%
Amount Prediction Accuracy	81.7%
Total Number Users	701
Mean Days Logged per User	2.64
Spoken vs. Written Logs	2,909 vs. 7,453
Foods Logged per Day	15.3 per person
Barcode Scans per Day	1.07 per person

Table 2 Pilot study survey results (1 best, 3 worst).

Perceived accuracy	1.86 ± 0.53
Personalization	1.79 ± 0.70
Appealing interface design	1.43 ± 0.65
Difficulty vs. other apps	1.67 ± 1.16
Personalized vs. other apps	1.33 ± 0.58
Fun vs. other apps	1.67 ± 0.58

CONCLUSION

This study demonstrates the acceptability of logging food intake with spoken natural language, with the potential to benefit individuals who find existing methods of tracking dietary intake too tedious for long-term use. In the future, we will use neural networks to predict quantity, and take photos of food.

FUNDING