

Allergen Guru: The App for Allergens

N. Jha, S. Agah, M.D. Chapman

Department of Computer Science, University of Virginia, Charlottesville, VA 22903, USA
Indoor Biotechnologies Inc., Charlottesville, United States



RATIONALE

Many existing web technologies have made the jump to mobile devices. Scientific resources, however, have been slow to follow. Current allergen databases are a powerful source of bioinformatics knowledge, but their utility is diminished by a lack of accessibility. Most productive science occurs at the lab bench, away from desktop computers but accessible to mobile devices. Our aim was to develop an Android application that could provide up to date information about allergens and be immediately accessible.

METHODS

A C++ program was written to download HTML content from Allergen.org. These HTML files were processed through the command-line tools grep and sed, as well as through a Python program. The entries were then validated and parsed into a SQLite database. Finally, a user interface was written in XML format with underlying logic written in Java. The source code is made freely available on [Github.com](https://github.com).

RESULTS

An Android application that will automatically update as new information is added to the WHO/IUIS allergen nomenclature database was successfully developed. This was made possible by constructing a web scraper that would periodically create a local, searchable database using the technologies outlined above. The app replicates functionality present in the WHO/IUIS website; allergens can be searched by name, taxonomy, source, or biochemical name. All information contained in the online database is stored in the application locally, so users are not required to maintain an internet connection - functionality that will never be present in the webpage-based implementation.

Figure 1. Allergen ID

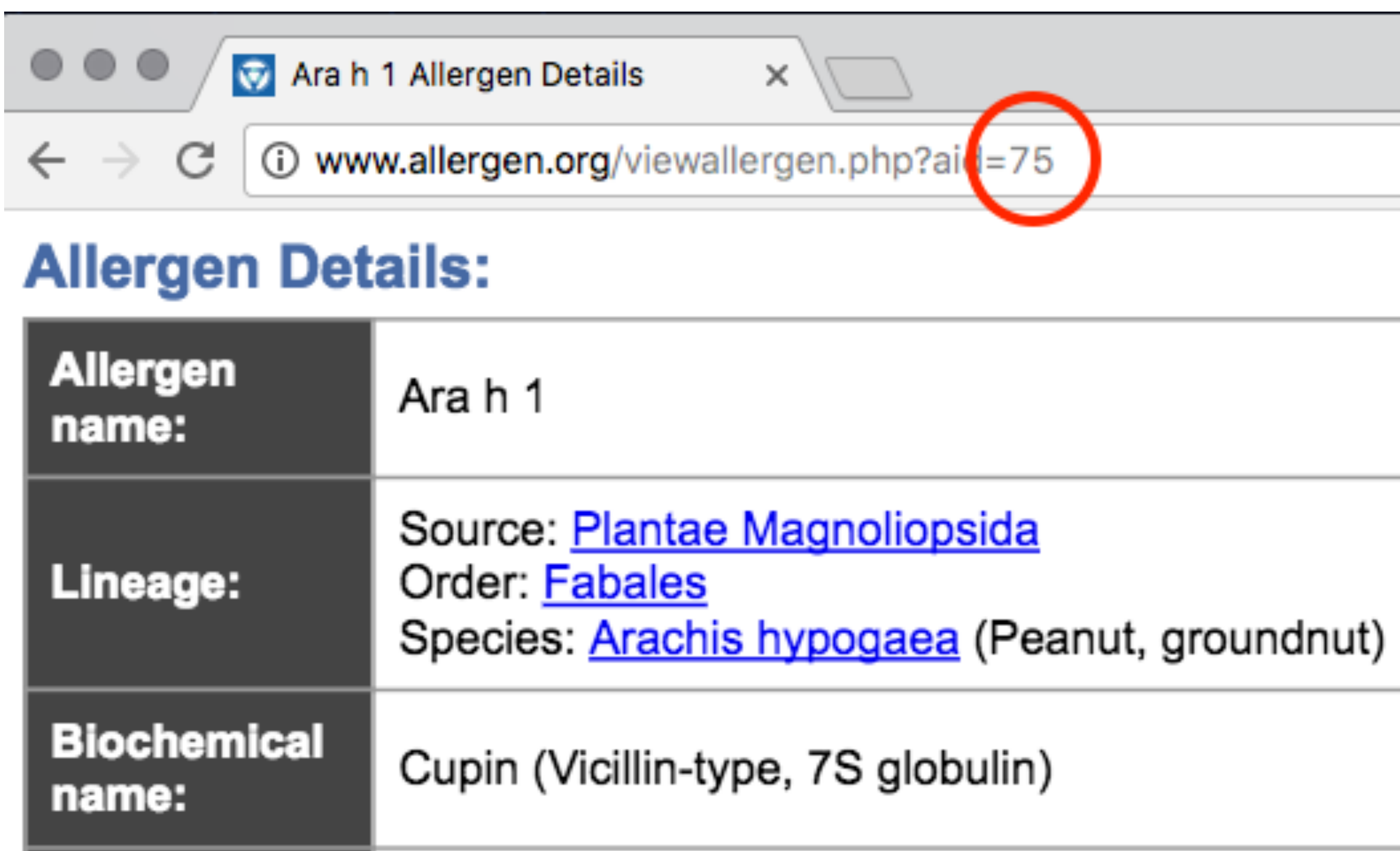


Figure 2. Data Extraction

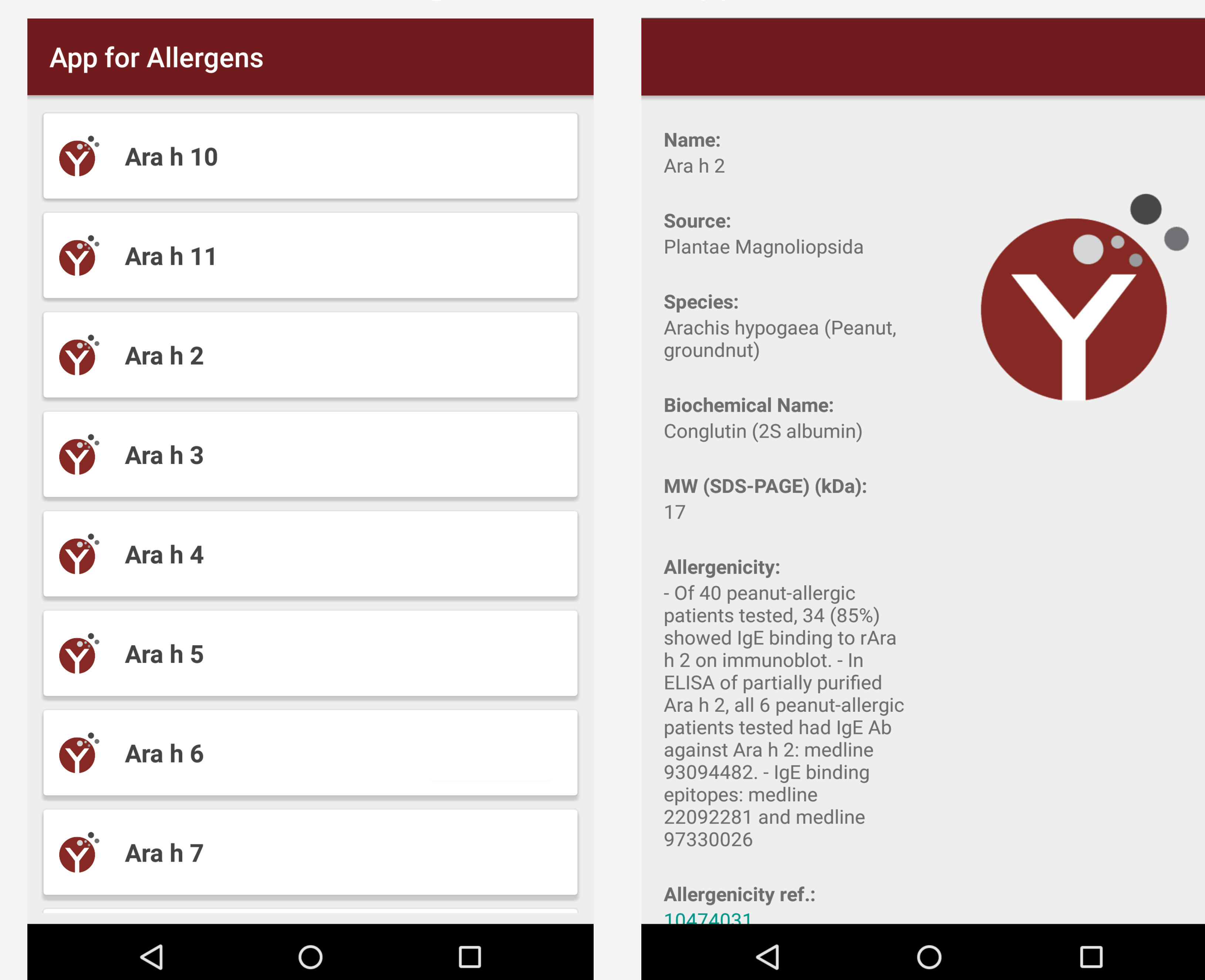
```
baseUrl = "http://www.allergen.org/viewallergen.php?aid=";  
for(int i = 1; i < 902; i++) {  
    ostringstream ossURL, ossFILE;  
    ossURL << baseUrl << i;  
    string queryURL = ossURL.str();  
    ossFILE << "allergen" << i << ".html";  
    string fileName = ossFILE.str();  
  
    string command = "phantomjs save_page.js " + queryURL  
    + " > " + fileName;  
    std::system(command.c_str());  
}
```



Figure 3. Data Representation

Name	Source	Order	Species	Biochemical Name	MW
Aca s 13	Animalia Arthropoda	Astigmata	Acarus siro (Storage mite)	Fatty acid-binding protein	15 kDa
Act c 10	Plantae Magnoliopsida	Ericales	Actinidia chinensis (Gold Kiwi fruit)	nsLTP1	10 kDa
Act c 5	Plantae Magnoliopsida	Ericales	Actinidia chinensis (Gold Kiwi fruit)	Kiwellin	28 kDa
Act c 8	Plantae Magnoliopsida	Ericales	Actinidia chinensis (Gold Kiwi fruit)	Pathogenesis-related protein	17 kDa
Act d 1	Plantae Magnoliopsida	Ericales	Actinidia deliciosa (Green Kiwi fruit) Obsolete name: Act c 1	Cysteine protease (actinidin)	30 kDa
Act d 10	Plantae Magnoliopsida	Ericales	Actinidia deliciosa (Green Kiwi fruit)	nsLTP1	10 kDa
Act d 11	Plantae Magnoliopsida	Ericales	Actinidia deliciosa (Green Kiwi fruit)	Major latex protein	17 kDa
Act d 2	Plantae Magnoliopsida	Ericales	Actinidia deliciosa (Green Kiwi fruit)	Thaumatin-like protein	24 kDa

Figure 4. Android Application



CONCLUSIONS

With the rise of mobile computing, scientists should expect their tools to accompany them wherever they go, whether it be the desk or the bench. The App for Allergens updates and improves a valuable bioinformatics resource, the WHO/IUIS allergen database, for allergy/immunology research. In addition, it provides an upgradeable, extendible platform that can quickly absorb changes in the database, as well as provide new features and research capabilities.

This work was supported in part by the Computer Science Department. of the University of Virginia and Indoor Biotechnologies.