Computers and Electronics in Agriculture 78 (2011) 150–161

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

# Machine-to-machine communication for agricultural systems: An XML-based auxiliary language to enhance semantic interoperability

Edmund W. Schuster[a], Hyoung-Gon Lee[b], Reza Ehsani[c,*], Stuart J. Allen[d], J. Steven Rogers[e]

[a] Laboratory for Manufacturing and Productivity, Massachusetts Institute of Technology, 77 Massachusetts Av., 35-135A, Cambridge, MA 02139, United States
[b] Oracle America, Inc., 500 Oracle Parkway, 3op645, Redwood Shores, CA 94065, United States
[c] Citrus Research and Education Center, University of Florida, Lake Alfred, FL 33850, United States
[d] 7904 60th Drive NE, Marysville, WA 98270, United States
[e] Ecostat, Inc., Post Office Box 237, Highland City, FL 33846, United States

ARTICLE INFO

*Article history:*
Received 6 August 2010
Received in revised form 22 June 2011
Accepted 7 July 2011

*Keywords:*
Agricultural computation
M Language
Data interoperability
Precision agriculture
Machine-to-machine communication
Internet

ABSTRACT

This paper puts forth an Internet-based architecture for machine-to-machine communication and computation that enhances bio-productivity in agriculture. The approach utilizes an auxiliary language to enable data interoperability in a synthetic computing environment and to make connections between data and mathematical models. The approach also includes some aspects of cloud and context aware computing. At the prototype level, a practical application from the Florida citrus industry demonstrates the concept. In general, future agricultural systems will be Internet-based thus reducing cost and increasing capability. Standards organizations are certain to play an important role in this development, which might continue for the next decade or longer.



"Vague and insignificant forms of speech, and abuse of language, have so long passed for mysteries of science; and hard or misapplied words with little or no meaning have, by prescription, such a right to be mistaken for deep learning and height of speculation, that it will not be easy to persuade either those who speak or those who hear them, that they are but the covers of ignorance and hindrance of true knowledge".

John Locke (Hayakawa, 1964)

## 1. Introduction

This paper focuses on connecting Internet entities together through machine-to-machine communication. The core of this approach is an auxiliary language (termed "M") that describes concepts[1] with precision using English words and phrases. The outcome is a synthetic computing environment for interoperability between data and mathematical models. At the prototype level of development, the system is low cost and easy to use. A case study involving the Florida citrus industry demonstrates the system through a real-world example.

The direction of this research is consistent with the observations of others within the agricultural community. For example, Nikkila et al. (2010) write, "The (agricultural) information system must be able to manage various data formats, both standardized as well as proprietary and be able to exchange data with services that provide computation for precision agriculture".

The outgrowth of an early project in semantic and contextual computing, the research presented in this paper has three goals:

(1) Solve the issue of semantics and syntax for XML using an online dictionary, initially populated with words and relationships from Wordnet (Fellbaum, 1998).
(2) Create an auxiliary language capable of interoperating data and mathematical models at low cost.
(3) Apply to a practical problem in agriculture, demonstrating the ability to scale.

In the spirit of innovation, many ideas that are widespread industrial practice have legitimate application for agriculture. To make these ideas work requires a multi-disciplinary approach, creative thinking, and free-form adaptation. Starting with a technology that had its genesis in manufacturing, supply chain management, and the defense industry, the vision of this paper is

* Corresponding author.
*E-mail addresses:* edmund_w@mit.edu (E.W. Schuster), ken.hg.lee@oracle.com (H.-G. Lee), ehsani@ufl.edu (R. Ehsani), stuart99allen@yahoo.com (S.J. Allen).

[1] A concept is a cognitive unit of meaning, http://en.wikipedia.org/wiki/Concept (2-15-11).

0168-1699/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.compag.2011.07.002

to cross the traditional boundaries between industries and academic disciplines to create a robust information infrastructure for agriculture.

## 2. Philosophy

Before exploring the specifics it is important to appreciate the underpinnings for constructs of the M Language (Brock et al., 2005, 2006a; Schuster et al., 2007). At its core, the M Language is a way of creating uniform semantics and syntax for XML without using the standards approach of specification. In this paper, XML serves as an important means to store and integrate data along with being a descriptor for abstractions (mathematical models). One of the major purposes of the M Language is to enhance the effectiveness of XML in practice. Beyond agricultural systems, enhanced XML has broad financial implications for industry in terms of improving data interoperability.

### 2.1. The quandary of XML

Developed in the late 1990s, XML is a flexible approach for machine-to-machine communication. It usually involves the transfer of data. The intent is to replace all forms of flat-file computer communication, primarily its predecessor electronic data exchange (EDI). Though EDI is a major innovation, the rigidity of the standard causes difficulty for applications beyond highly controlled, point-to-point communication.

For all of its flexibility, XML has yet to completely displace EDI in many business applications. There are several significant limitations. First, prior agreement must exist on semantics and syntax for two parties to exchange XML data. This agreement often becomes a negotiation process that can take years to resolve, especially between various standards organizations.

Second, the relative ease of creating a specific XML standard almost always results in an explosion of the number available for use in industry. This is also potentially true for agricultural systems as machine-to-machine communication evolves and precision agriculture gains prominence.

None of the XML standards are interoperable. In many cases, multiple standards exist for an industrial segment along with multiple versions within a single standard, causing an enormous amount of systems integration effort and extra cost in practice. Essentially, the XML standard has an excessive amount of flexibility in comparison to EDI, which has no flexibility whatsoever.

### 2.2. The auxiliary language

The driving philosophy of this paper is the creation of an auxiliary language[2] as a means to interoperate data between different existing XML standards. Such a language is capable of integrating data and mathematical models across the Internet. The M Language, initially a defense industry research project dating to 2003,[3] serves the role of an auxiliary language (see Fig. 1).

The details of the M Dictionary appear in Section 4. What is important to know from a broad perspective is that data undergoes translation into M-XML at the edge of computing systems, forming a synthetic, interoperable environment for the Internet.

Combined with a standard for describing and connecting mathematical models, the auxiliary language is a way to deal with the many-to-many problem common to any type of communication network. The number of connections (or translations) between different XML standards is polynomial as a function of the total number in place. Given that there are hundreds of XML standards used in commerce, the number of from-to connections is immense. Each time data from dissimilar XML standards is integrated to form a set; individual computer programs must be written to reconcile the semantics and syntax.

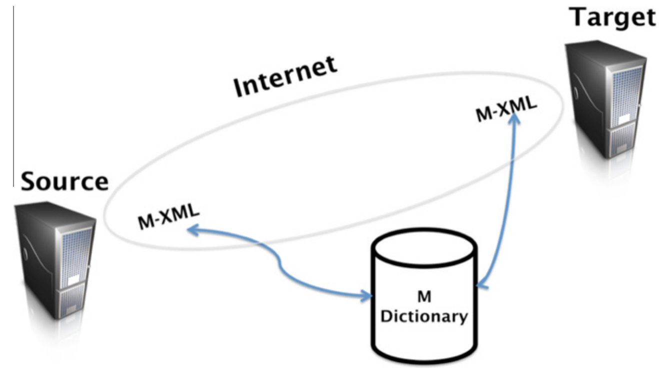



**Fig. 1.** The auxiliary language.

With a single, automatic translation at the edge using the M Language, data becomes interoperable in terms of semantics. This approach requires much less effort in systems integration. In addition, the auxiliary language allows for the re-use of mathematical models through a machine understandable method of description within a network. For precision agriculture, model re-use offers many opportunities to analyze field data at the farm level.

In summary, an auxiliary language specific to precision agriculture is a valid research activity given the need for data interoperability. Other authors have similar, supporting viewpoints such as the following from Nikkila et al. (2010):

> "Since some or most of the calculation related to precision agriculture occurs in external services, the data stored on the Farm Management Information System (FMIS) needs to be transferred to these services. Unfortunately, few if any external services can be expected to share a common format for data and none can be expected to support the internal presentation of data used by the FMIS".

The authors go on to state in the same paragraph:

> "This requires the FMIS to manage several transformations of data in order to achieve interoperability with all relevant systems and services. Currently, no standards or de facto standards exist for this communication, though some common formats, such as agroXML (Schmitz et al., 2008) and FODM (Anon, 2002) have been proposed".

### 2.3. Antecedents

The general idea of machine-to-machine communication has long been a goal of artificial intelligence (AI). In the study of AI, researchers employ various computer-based techniques to infer the meaning of concepts with the goal of enabling machine understanding on a large scale.

These approaches have largely failed in practice because the meaning of a word, a phrase, or a sentence depends on the semantics of each word, the syntax or order of usage, and the particular context. The unique properties of the human mind can determine contextual relevance, and then figure out how several relevant variables are associated (Deacon, 1997). However, to date, the sub discipline of AI called Natural Language Processing has not duplicated this mental property (Chandler, 2009).

[2] An auxiliary language is generally defined as an organized artificial means of communication created for situations where there is not a single shared language.

[3] See Raytheon, http://www.raytheon.com/technology_today/2010_i1/feature_15.html (2/15/11).

In 1999, the W3C[4] began an ambitious effort to create automated ways of understanding the meaning of concepts via the Semantic Web. The focus involves applying knowledge management techniques to the entire Internet with the intent of machine understanding for the information contained in each web page.

The standards for the Semantic Web, including Resource Description Framework (RDF) and Web Ontology Language OWL (see Allemang and Hendler, 2008), led to many efforts to create ontologies, which are deterministic connections between concepts expressed as natural language. The theory of ontologies involves the grouping of various concepts that make up a body of knowledge to form hierarchies based on similarities and differences. By one early definition, these are some of the criteria for human intelligence (Feigenbaum and Feldman, 1963).

A few examples include Simple HTML Ontology Language (SHOE), DARPA Agent Markup Language–Ontology (DAML–ONT), Unified Problem-Solving Method (UPML) (all from Fensel et al. (2003)), and many others (see Storey et al., 2008 for a robust example using DAML). Massive in size, scope, and funding, these efforts intend to establish the Semantic Web as the next stage of the Internet.

However, after twelve years the outcome of this work remains a topic of research and many of the standards are not part of common industrial practice or pervasive for web pages. In other associated areas, recent social networking technologies like Twitter and Facebook, along with mobile computing like the iPhone OS and Android mobile operating system have surpassed the Semantic Web by wide margins in general adoption and use. These technologies provide a different perspective in establishing connections between people, the sharing of information, and the re-use of low cost software.

In 2003 research began on the M Language, an early context aware approach viewed by the authors as an alternative to the Semantic Web. The effort uses some of the standards from the W3C but concentrates on an Internet-based dictionary for exact definitions and relationships rather than solely on ontologies, the approach of the Semantic Web community. In addition, the M Language does not involve natural language processing for web pages or complex reasoning through networks of concepts formalized by an ontology such as OWL. Rather, the goal is to provide unambiguous definitions for words, compound words, and noun phrases used in XML.

Some of the general motivations for creating the M Language originate with Schuster and Allen (1998, 1999). An initial source of prior inspiration is from Jones (1996). Another major influence is Geoffrion (1987, 1989), who put forth the concept of Structured Modeling and includes an early focus on semantics. In addition, Geoffrion (1976) discusses the use of auxiliary models to enhance insight. Schuster and Allen (1998) apply these early ideas about auxiliary models to complex financial and operational planning involving an agricultural cooperative and to plant level execution systems for process manufacturing common to the food industry (Schuster et al., 2000). In all of these cases, the research occurred prior to ample bandwidth, the widespread use of the Internet, the existence of the Semantic Web, and before other standards developed by the W3C.

### 2.4. Cognate literature

Given the practical data and modeling needs for precision agriculture, some researchers draw attention to the substantial challenges that exist in research, development, and implementation. In a paper by Nikkila et al. (2010), the authors make this point by stating "Research is... somewhat lacking on information systems for precision agriculture with no large-scale prototype in existence and only a few publications on the possible architectures for an FMIS for precision agriculture".

In consideration of this need, several authors make important research contributions that further define and advance knowledge for agricultural information systems. Murakami et al. (2007) propose sophisticated information systems for agriculture comprised of "a reference architecture; a standard language for data exchange between systems, named PAML – Precision Agriculture Markup Language; and a service bus, AgriBUS – a message-oriented middleware for connection of Web Services". PAML includes a standardized vocabulary based on early work done by the United Nations Food and Agriculture Organization. Of note, this prototype provides the ability to handle data from outside of the traditional agricultural community. In many ways, the research is synergistic and parallels some of the ideas of this paper.

Nash et al. (2009a,b) make a substantial contribution in mapping the complexity of data associated with precision agriculture. They discuss agroXML, a data standard for exchange between farmers and other supply chain partners and explore the need to integrate data from many different sources into a set. In addition, they examine the intricacies of geospatial data and issues of interoperability.

With both Nikkila et al. (2010) and Nash et al. (2009a,b) the approach is to gain consensus on a single standard and then seek adoption among all industry participants. Brock et al. (2005) point out that this approach is extremely challenging because by definition standards and ontologies are "rigid and inflexible, and assume one absolute definition exists for each knowledge element". Hence, it is difficult to incorporate anything new or unanticipated that exists outside of the existing standard or ontology.

The M Language overcomes some of these disadvantages and offers an alternative to the traditional standards approach. The auxiliary language method used by M reduces the dependence on formulating a single standard agreeable to all entities. It acts as a form of glue between standards thus enhancing interoperability.

### 2.5. Semantic web compared to M Language

While the M Language uses some of the standards developed by the W3C, namely XML, XSLT, and Web Services, there are specifics in design and architecture that differ from the Semantic Web. The core of the M Language is a central online dictionary that can be queried repeatedly along with the idea of an auxiliary language adapted from the study of linguistics. It also uses a unique way of building ontologies for use with XML using concepts with explicit definitions (Lee and Schuster, 2009).

This contrasts with W3C standards, which have few robust means of handling semantics for tools such as OWL beyond the use of triples. Current practice is to treat words and text as symbols connected via an ontology. Reasoning occurs through the established tree of relationships in RDF or OWL. This assumes that no semantic ambiguity exists.

In addition, formal ontologies assume that concept definitions do not change over time. Hence from the W3C viewpoint, the definition of a word (concept) is implicit and in relation to its ontological connections. Of course the definition of a concept can change by modifying its relationships to other concepts. However, for ontology languages like DAML that have established libraries available online, the connections between concepts are intended to remain static for the large part.

Yet this is hardly true in the real world. For example, the word "apple" could mean a fruit or a brand of computers, depending on the context. There is no way to know for sure the exact meaning of "apple" or how the word fits into existing ontologies for fruit or

[4] www.W3C.org (2/15/11).

computers. Few ontologies are absolute and all encompassing because about 10% of English words commonly used in business have multiple meanings based on an analysis from the MIT Data Center Program. W3C standards include no provision for an online dictionary using explicit definitions with the possible exception of the comment property in RDF.

Further, the unit of meaning for a concept included in an ontology can change over time. For example, the word "Chad" has existed for 14 centuries as the name of a saint who lived in the most northern county of England (Howard, 2005). During the US Presidential election of 2000, the word suddenly took on new meaning where a hanging chad referred to an improperly punched ballot. Within a single ontology, it is extremely difficult to re-assign a concept from one class to another (for example from "name" to "voting ballot") while maintaining consistency for all other relationships between concepts. In some cases, multiple ontologies exist for different knowledge domains making the situation even more complex. In general, W3C standards have no systematic provision to capture a change in the meaning of a concept.

Another important difference is that the M Language uses an advanced database (Oracle 11g) to store and manipulate XML. In this regard the M Language, which currently uses the Microsoft .NET platform, is akin to a cloud computing application where uniform semantics enhance the interoperability of XML as a carrier for data. The information technology architecture for the M Language is compatible with emerging cloud-computing applications such as Azure. In general, cloud computing represents a major growth area in information technology (Veverka, 2010).

A final difference is that the M Language is perhaps an early version of context aware computing for language in that an exact definition attaches to a concept. This is done through a simple computer expression for words and phrases that is symbolic and machine understandable while also being human-readable.

The next three sections of this paper continue the logical exposition of M as an auxiliary language for machine-to-machine communication. Section 3 addresses the web-based software needed to make an M Language application for agriculture. Section 4 provides the details of the dictionary component of the M Language. The initial population of words and relationships for the dictionary comes from WordNet.[5] Following the details of the M Language, Section 5 offers a specific, practical application involving the Florida citrus industry.

## 3. Web-based systems for agriculture

As defined by the World Wide Web Consortium, a Web Service is "a software system designed to support interoperable machine-to-machine interaction over a network" (Booth et al., 2004). The Web Services approach as a way to expose software located on a specific server to the Internet is now a widely accepted practice. It is becoming the standard for communication between different business computing applications (Vinoski, 2003).

Given enough bandwidth at the farm level, it is a reasonable extension that essentially all agricultural systems will eventually become an Internet-based service rather than a dedicated software application hosted on local computers at the farm level. This means that the Internet has potential to become a powerful vehicle to match a wide range of individual mathematical models to local data for a variety of real-world problems experienced in industry (Lee and Schuster, 2007).

Establishing Internet-based agricultural models, however, requires the resolution of several research issues in computer science and information technology. These issues relate to the connection of data and mathematical models, which both have complex hierarchies and semantics. Creating open system architecture for near universal access to mathematical models via the Internet is the innovation needed to expand Web Services as a fundamental way of delivering computational capabilities.

The problem of creating an open system for the agricultural models, and mathematical models in general, has two elements: (1) a simple user interface, and (2) a means to connect the interface to computer code located on remote servers across Internet. The central idea is to host an agricultural model, written in a structured computer language like Java or C++, on a single server with an interface that can be downloaded onto any computer using the Internet. The interface then connects to the central server when running the model. Such a system allows users located anywhere in the world to access a particular agricultural model with little implementation and at no cost.

For the Florida citrus industry example detailed in Section 5, which involves data integration and connection to a model, an Excel spreadsheet serves as the end-user interface. Spreadsheets are easy to understand and many farmers and researchers do agricultural modeling using custom spreadsheet approaches developed internally. Enhancements in Excel 2003 and 2007 allow for direct interaction with a remote server where computer code for a specific agricultural model resides. The approach outlined in Section 5 is not limited to spreadsheets. An Internet browser could also be the interface.

Creating an open system for modeling using Excel spreadsheets also requires a robust way to treat semantics. The example in Section 5 uses the M Dictionary to provide consistent, machine understandable semantics for words and noun phrases contained in the spreadsheet interface that are elements of data and mathematical models. In this way, there is no ambiguity regarding the definition of terms used to describe the data fields and Internet connections between the Excel spreadsheet and the remote server.

Section 4 contains a description of the important functional aspects of the M Dictionary in relation to the open system.

## 4. The M Dictionary

While Web Services are crucial to providing Internet connections for data and mathematical models, the M Dictionary is the essence that binds the entire synthetic environment together. From a general perspective, a noted language columnist (Angwin, 2009) sums up the basic requirements for machine understandable semantics:

> "We need a dictionary that is as dynamic as our use of the language. And although Google is doing a pretty good job aggregating meanings, I would prefer some human experts to give authority and heft to a new database of meaning".

The following sections provide the details of the M Dictionary. The specific aspects of the Web Services interface are found at mlanguage.mit.edu.

### 4.1. Dealing with semantic ambiguity

The words used in the M Dictionary are slightly different from English words. In M, every word has only one definition. This is an extremely important characteristic because computers that communicate using M do not need to understand the context or usage of a word to know its meaning. This is in contrast to artificial intelligence-based methods that take a deductive approach to infer the meaning of a concept based on surrounding words and phrases. It is the human readable unique identification of a concept along

[5] George A. Miller, Princeton University, http://wordnet.princeton.edu/ (2/15/11).


**Fig. 2.** An example from the M Dictionary.

with an exact definition that sets the M Language apart from traditional Semantic Web approaches.

English words are ambiguous. For example, the word "cell" might mean "cellular phone", "biological cell", or "fuel cell". Without some idea of the context, it is impossible to know the meaning of the word "cell".

To overcome this issue, the M Dictionary includes a numeric extension to denote individual words. For example, a specific meaning of "cell" becomes cell.1.

To account for several definitions of the same word, the M Language allows multiple numeric extensions, one for each definition. Thus, cell.1 is a word in M and cell.2 is a different word. With this method, every word has only one meaning.

Fig. 2 shows a screen shot of the M Dictionary (mlanguage.mit.edu) for the word "tractor". In this example, there are two different definitions.

### 4.2. Compound words

Depending on usage, some compound words are also part of the M Dictionary. An example is "operations research", represented as operations_research.1.

In practice, many descriptors for data and mathematical models are compound words. For even more complex descriptors, the M Language has the capability to form machine understandable noun phrases. This usually means a noun modified by one or two adjectives.

The upper limit of processing capability for the M Language is noun phrases. Descriptors that use simple or complex sentences are not part of M. There is no ability to do natural language processing or to identify clauses contained within a particular sentence. As such, M does not enter the realm of word, phrase, or language inference.

### 4.3. Relations, formats, and language translations

In addition to the definition, the dictionary entry also contains three other pieces of important information. These include: (1) word relations, (2) data format, and (3) language translations. All of this information is available through Web Service interfaces. In comparison to Semantic Web standards, the functions discussed below are similar to SKOS,[6] the difference being that the relationships for the M Language are contained in a database rather than URL.

Word relations are simply the connections between words. These relationships include synonyms, antonyms, types, and parts. Synonyms and antonyms are the same as in English.

Types refer to word generalizations. For example, automobile.1 is a *type of* motor_vehicle.1

Parts are words that are components of another word. This is often the case when thinking about physical objects, although this could also be the case with abstractions. For example, a wing.4 is a *part of* airplane1

Both types and parts establish a hierarchy within the dictionary through making connections between entries. These word connections are valuable in a number of different ways, forming an ontology based on explicit meaning (definitions for each word).

Data format provides guidance concerning the forms and patterns of data values that are associated with a particular word. In many situations, computer-to-computer communication might contain a word such as first_name.1 that has an associated data value such as "John". Other common situations include words like telephone_number.1, account_balance.1, or postal_code.1. In all of these cases, a word in the M Dictionary has a particular format or pattern for associated data. Regular expressions, contained in the M Dictionary, allow for a means of checking if data matches a pre-specified pattern. This capability is helpful in establishing data quality.

Finally, the language translation portion is simply the representation of the word in the M Dictionary as a word (or phrase) in another human language besides English (Brock et al., 2006b). In most situations, computer-based language translation is extremely difficult because a lack of context exists for the specific communication. Since in the M Dictionary each word has only one definition, the word cell.1 (biological), for example, cannot be confused with cell.2 (telephone). Words with a single definition allow users to specify exact meaning independent of context. This eliminates ambiguity in translation.

### 4.4. Wiki

The M dictionary uses the wiki approach with several important modifications including improved security through user registration, maintenance of the integrity of word relations, a monitoring function to reduce the chances of near identical definitions, and administrative controls to ensure accuracy. Of particular importance is the process to create a new definition for a concept (and numeric extension) where appropriate. To date, this process has relied on human judgment and tests of reasonability, recognizing that different combinations of word choice and order can result in the same meaning.

The M Dictionary has various statistical features that measure usage. This is helpful in understanding the overall characteristics of the lexicon (i.e. the number of concepts with more than one definition). Currently, there are 210,000 words in the M Dictionary and over 700,000 ontological relationships.

[6] Simple Knowledge Organization System, see http://en.wikipedia.org/wiki/Simple_Knowledge_Organization_System (6-11-11).

The wiki process has emerged as an innovative application of Internet technology to knowledge management and consensus building. A wiki is a type of website that allows users to add and edit content and is especially suited to collaborative authoring. It is remarkably accurate (Associated Press, 2005) and several companies have begun to use the Wiki process internally (Wessel, 2005). Since 2001, Wikipedia has become the largest encyclopedia ever created with over 3.5 million articles in English.

### 4.5. Summary

The current XML standard allows the ability to use any word, regardless of its definition, for a tag. This becomes a drawback, as users might unknowingly use a semantically ambiguous word or phrase that could be misinterpreted.

With the data contained in WordNet as a starting point, the M Dictionary innovates in two general areas: (a) one definition per word, and (b) an ontology for word relationships based on exact definitions. The Web Services connections to the M Dictionary such as GetWord and TestRelation allow for query and reasoning of the definitions and ontology contained in the dictionary. With these connections, the M Dictionary can interact (through the construction of additional tools) with various Internet-based XML data as a means of providing precise definitions for words used in XML. Depending on the scope of the industrial domain, multiple M Dictionaries might be appropriate with each focusing on a particular segment of terminology. Initial reviews of prototypes by others involved in the transportation industry support the value of the M Language in solving "many of the problems of XML". (see Englert et al., 2010).

As a concluding remark, the design of the M Dictionary relies on the principle of centralization using a database (SQL) with access through Web Services. In this way all concept definitions, relationships, and regular expressions are in one location. The design facilitates the loading of Wordnet as an initial seed of definitions and relationships and the operation of the Wiki process.

It is possible to form alternative architectures using an URL approach.[7] In this case a URL, for example something named http://mlanguage.mit.edu/website/tractor.1, could represent a concept (tractor.1), located on a particular server and linked to a document that represents an instance, which in turn contains the concept definition.

This approach will work for Internet-based concept definitions and relationships. However, in many cases URLs do not have consistent, human understandable meaning. In addition, it becomes difficult to consistently assign URL names across distributed entities unless recognized sets of criteria exist. In this situation, namespace control might be lost. One of the evolving functions of standards organizations is to establish a methodology for consistent URL naming that is human understandable yet distributed across the Internet.

## 5. Application

Envisioning the early promise of the Internet and computing, Steve Rogers states in 2001 that, for agriculture in general, "... there is incredible untapped potential in unused data. Many of our most important research and marketing issues could be resolved today if existing agricultural information was digitally organized, interoperable, and immediately accessible to people who could make a difference".[8]

[7] See http://www.w3.org/2002/11/dbooth-names/dboothnames_clean.htm#UsingURLsToIdentifyThings (2/10/11).

[8] Rogers et al. Internet-Based Survey Systems (2001).

Perhaps one of the most data-intensive problems faced by those in agriculture is the real-time identification of disease and insect outbreaks in the field. Spatio-temporal detection is important information for timing chemical applications to halt such outbreaks, thereby preserving the quality and yield of a crop.

The initial motivation for this application of the M Language comes from a meeting held in Arlington, VA during April 2007 involving research priorities for the specialty crops industry (berries and brambles, ornamentals, wine and grape, tree fruit and nuts, and citrus).[9] During the group discussions, a grower mentioned that insect threshold calculation for a specific field represents a good area for agricultural research. Specifically, the grower wanted to know when to apply various agricultural chemicals based on the observed concentration of insects per square meter. If the insect concentration were below a pre-calculated threshold, then the chances of rapid growth would be small. Otherwise, application of chemicals should happen as soon as practicable to avoid a rapid multiplication of insect counts and crop damage. As such, this situation represents a fundamental problem in diffusion, modeling, and spatial analysis.

In many cases, detailed data about the location and intensity of disease or pest outbreaks does not exist. Where data does exist through commercial scouting services, there are frequently significant time lags before various data processing steps can take place to make the information available for action.

Currently, there is little if any capability among the various commercial scouting services to rapidly consolidate data from other sources such as the National Oceanographic and Atmospheric Administration (NOAA). Because of this lack of spatial and temporal information, farmers frequently apply pesticides and other chemicals at predefined intervals and at uniform rates intended to deliver a lethal dose over the entire field. This is in spite of the fact that disease is generally not uniformly widespread. Rather, clusters tend to appear in specific areas of a field.

### 5.1. Florida citrus

A typical case involves two diseases, citrus canker (*Xanthomonas citri* subsp. *citri*) and huanglongbing (*Candidatus* Liberibacter asiaticus) (also known as HLB or greening), which together could destroy the Florida citrus industry within about ten years if not controlled. The source of great anxiety among growers, these bacterial diseases spread via insect vector (HLB) as well as through the movement of cells blown from grove to grove by wind-driven rain (canker) (Rogers, 2000). To ensure that canker is under control, growers usually apply a solution of copper up to 9 times per year in many groves. The copper solution is the only recommended chemical method to control citrus canker in Florida (Dewdney and Graham, 2011), while HLB is managed through control of its insect vector, Asian citrus psyllid (*Diaphorina citri*) (Rogers et al., 2011). The long-term environmental effects on ground water quality are undetermined and the numerous applications also represent a significant production cost.

Given a lack of automatic ways for sensing, Florida growers employ human scouts to survey the fields in search of these diseases and insects. Using pencil and paper, most scouts write down various observations and make an estimate of the location.

A recent innovation involves the use of mobile devices such as 3G cell phones to record disease observations and insect outbreaks (Rogers et al., 2006). The advantage is the elimination of paper records as the devices contain specific computer programs that prompt the recording of observations with the aid of GPS

[9] See Engineering Solutions for Specialty Crop Challenges, http://www.csrees.usda.gov/nea/ag_systems/pdfs/specialty_crops_engineering.pdf (2/15/11).

coordinates for exact location. Upon completion of a field audit, the disease observations are uploaded via communication satellite to a data center that processes the information. Same day, the grower receives a complete digital map of the field via an Internet system and decides a course of action to deal with various outbreaks.

### 5.2. The data and model

While the mobile device scouting system is a significant advancement in making detailed agricultural data available to the grower, an important aspect for success remains missing. To adequately project the rate of growth, there needs to be a correlation between the state of the outbreak, daily temperature and other weather parameters. Logistic functions, commonly referred to as an "*S*" curve, exist that can give reasonable projections of pest growth given temperature and other inputs. The general form of the *S* curve equation is:

$$P(t) = 1/(1 + e^{-t}), \tag{1}$$

where *P* is the population and *t* is time (days). Graph 1 shows a plot of this equation.

When plotted, this equation gives the following:

Expanding on Eq. (1), Rogers (unpublished) made an adaptation:

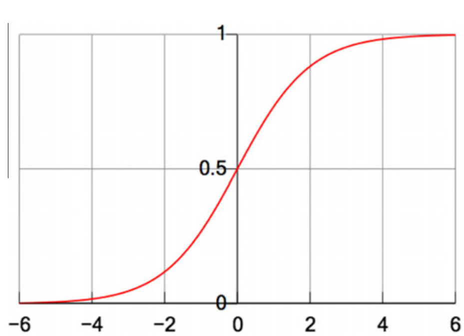
$$P(t) = K/(1 + (B * \mathrm{Exp}(-(\mathrm{GR}) * t))), \tag{2}$$

where: *K*, maximum environmental capacity; GR, the modified Allen growth rate; *B*, initial pest density at $t = 0$.

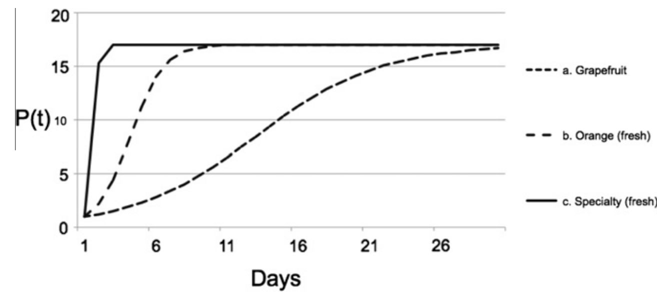
*K* is a constant derived from experimental results. With historical and forecast temperatures as inputs, a separate equation calculates GR. A computer program runs iterations of Eq. (2) in daily steps from $t = 0$, recording *P* (the projected density of insects) over time, which follows the *S* curve pattern. This gives an estimate of how many days until reaching the threshold concentration of insects requiring the application of agricultural chemicals.

Graph 2 shows several examples involving grapefruit and oranges using real data from groves. This highlights that the shape of the logistics function can vary based on initial parameters. In time series *c. Specialty Fresh* the growth rate shows a rapid projected increase in insect density. In contrast, the growth rate for time series *a. grapefruit* is gradual. This helps the grower to plan the timing of pesticide application. Estimates of GR are initially based on experimental observation. One valuable feature of the technologies introduced in this paper is that the accuracy and precision of GR can be improved iteratively over time as more data are contributed to the system.

Overall experimental evidence shows that Eq. (2) gives accurate projections for the growth of various classes of insects, including citrus rust mite (Rogers and McCoy, 1996) and Caribbean Fruit Fly (Rogers et al., 2006). Unfortunately many growers do not have weather stations on their farms or sophisticated sensing systems to




**Graph 1.** The logistics function (from http://en.wikipedia.org/wiki/Logistic_function (2/15/11)).




**Graph 2.** Projected insect density vs. time.

record spatial temperature readings for an orange grove, which the supporting equations need for the calculation of GR.

NOAA does record a history of detailed weather observations. However, this data is in a completely different format that is not interoperable with the disease observation data obtained via the mobile device. Likewise, NOAA publishes daily weather forecast data in XML, extending several days into the future. As in the case with historical weather data, the XML semantics and syntax are different compared to the mobile device observational data set of disease location.

### 5.3. Problem definition

Though surface observations of diseases and insects (mobile device) and weather data (NOAA: historical and forecast) are available through the Internet, the formats are not interoperable. This limits the ability to rapidly combine the data for analysis. Further, no systematic way exists to connect the logistics function to the data via the Internet.

Two computer programs address the problem.[10] The first is architecture to integrate data. The second provides a means for connecting the logistics function to the combined data set. Both computer programs use the M Dictionary for machine understandable semantics.

### 5.4. The process of combining data

The computing architecture for bringing data together appears in Fig. 3. The core of the architecture is Oracle 11g, a native database capable of holding data in XML and relational form. The entire architecture uses Microsoft.NET as a platform. External to the architecture are the two data streams in XML, surface observation data and NOAA weather history, along with a web service connection to the M Dictionary for semantics.

An important aspect of the system is the flow of XML data into 11g, which is a continuous process after some initial set-up. At this step, XML goes through a semantic translation into M-XML using the *M Converter Factory*. This employs XSLT along with other computer programs to replace existing XML tags with words from the M Dictionary. The advantage is that after conversion, M-XML has consistent semantics, creating a level interoperability. In addition, M words have an exact definition contained in the M Dictionary accessible through Web Services. There is no semantic ambiguity

Fig. 4 shows the conversion of NOAA data into M-XML. As an example, the raw NOAA XML contains the tag <MaxTemp>. The M Converter Factory automatically replaces MaxTemp with the M word maximum_temperature.1. This M word references an exact definition in the M Dictionary that specifies such important things as units of measure. It also is in human-readable form. This

[10] Lee-Schuster Semantic Enterprise Architecture (MIT Case #13754) and Kratulos (MIT Case #13752).

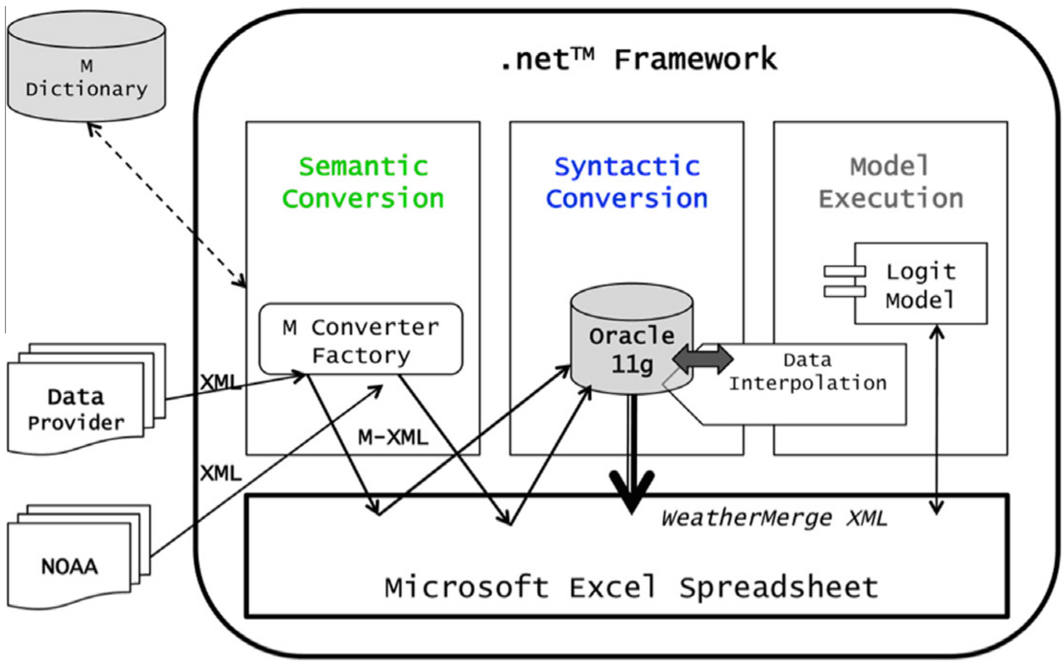



Fig. 3. Architecture for data and model interoperability.

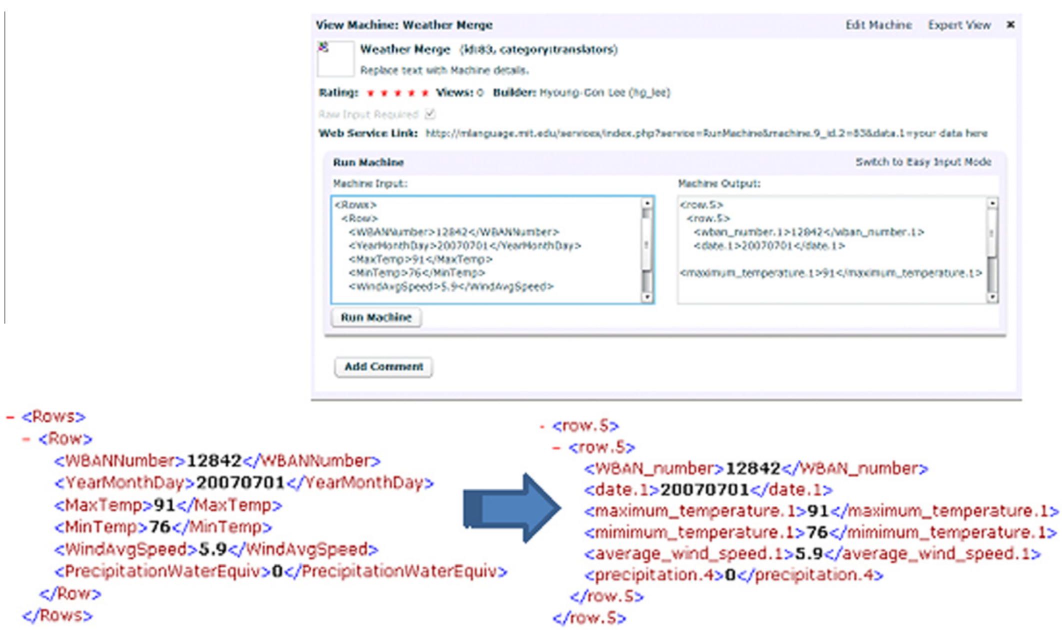



Fig. 4. Conversion of XML tags.

is an advantage because XML tags usually consist of abbreviations and camel case words that are nearly impossible to decipher at first glance.

It is important to note that an individual tag like maximum_temperature.1 could be a tag for other XML code in an entirely different system, like temperature sensor readings, as long as the data meets the same definition in the M Dictionary. In this way, the M Dictionary establishes a uniform semantic for XML tags.

Once the semantic conversion is complete, the M-XML resides in 11g and is mapped to a relational table. Extraction of the data in XML format from 11g, with M words as tags, can be in any pre-specified syntax. This overcomes the difficulty of establishing a universal standard for XML syntax, something that has never been successful in practice. In general, a researcher notes "the failure of XML to ensure interoperability at a syntactic level" (McDonough, 2008).

Fig. 5 is a representation of merging two M-XML sets of data for surface observations (pests) and weather data (temperature). The merging function is part of 11g and requires some programing in the database language. The merged data retains the M words as tags.

Upon merging in the database, the M-XML can undergo extraction in a pre-specified syntax and then be automatically imported into a pre-prepared Excel spreadsheet. The spreadsheet and 11g

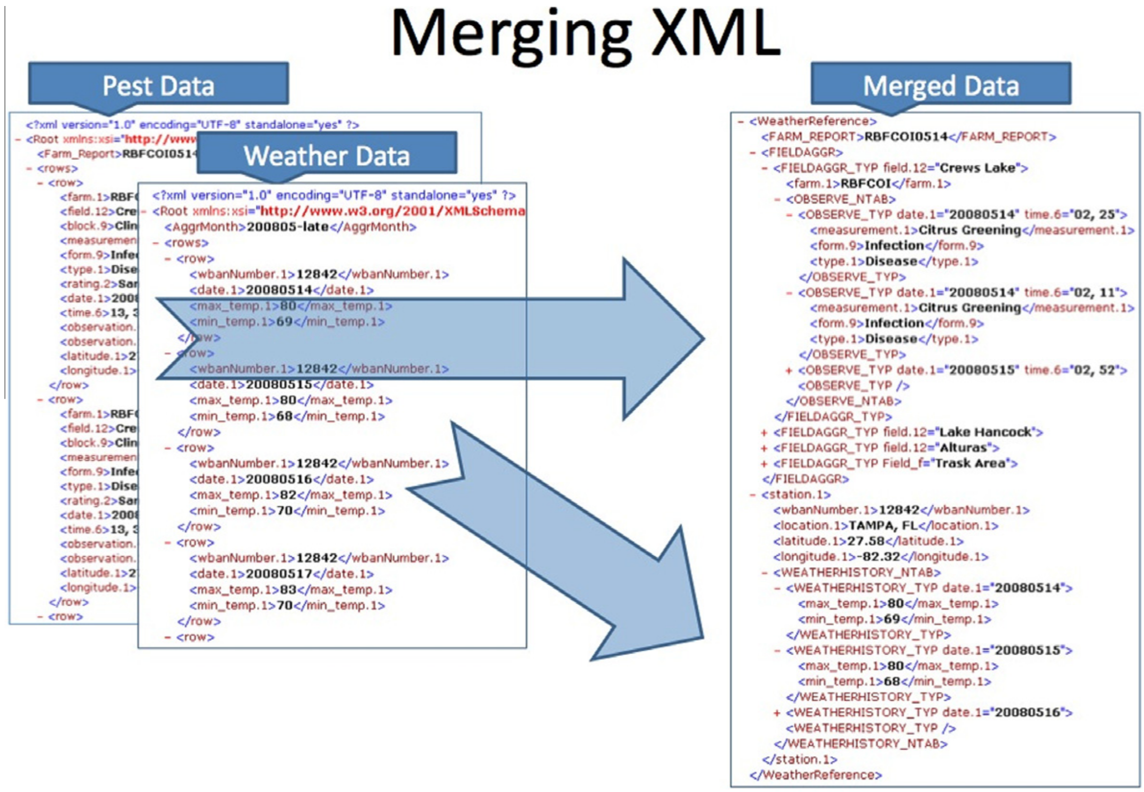
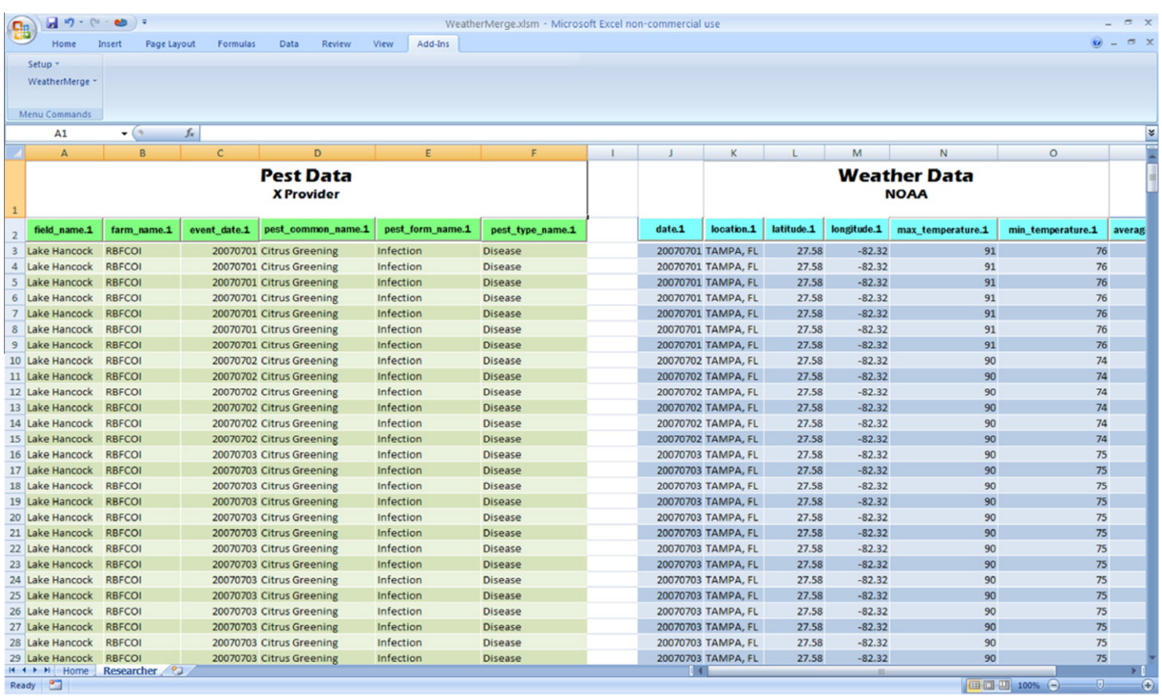



**Fig. 5.** Integration of M-XML.




| field_name.1 | farm_name.1 | event_date.1 | pest_common_name.1 | pest_form_name.1 | pest_type_name.1 | date.1 | location.1 | latitude.1 | longitude.1 | max_temperature.1 | min_temperature.1 | averag |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |

**Fig. 6.** Pest and weather data from 11g.

are independent. This means that a user with the appropriate spreadsheet, perhaps downloaded from a web site, could access the data from 11g anywhere depending only on a suitable connection to the Internet. The authors have successfully made a connection and execution of the spreadsheet in Korea, a country noted for its very good Internet access, to servers located at MIT. As such, the data center for processing agricultural data is location independent. Fig. 6 shows the spreadsheet with the imported data from 11g.

As a final example, the spreadsheet in Fig. 6 automatically connects with the M Dictionary via various Web Services. This allows for the definitions of each M word to be pulled into the

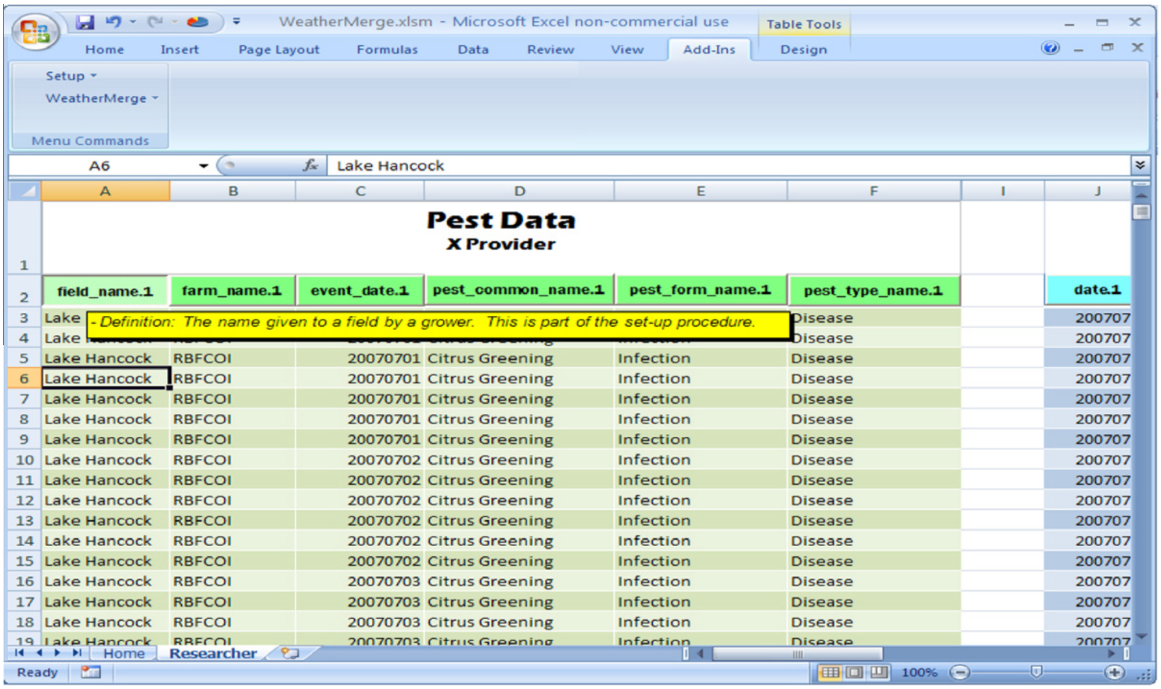



**Fig. 7.** Definition for an M word.

spreadsheet. In Fig. 7, the definition for field_name.1 appears as a drop down box by clicking on the M word. In this way, an exact definition appears for each word. No semantic ambiguity exists.

As an additional comment, the above prototype involves about 10,000 records of diseases and Caribbean Fruit Fly observations and weather data. The operation of the computer programs to integrate data and execute mathematical models via the Internet requires only seconds to accomplish. The capacity of 11g can handle records numbering in the millions.

### *5.5. Connecting to a model*

Upon consolidation of the data shown Fig. 7, a model located on a remote server can be connected directly to the spreadsheet. In this case, the spreadsheet acts as a data interface to the model.

For demonstration purposes, the logistics function is embedded directly into the Excel Spreadsheet. However in other, even more advanced applications, researchers have successfully used the spreadsheet as an interface to complex algorithms located on a server at MIT (Lee et al., 2010). The application is in the area of production planning though many of the concepts are the same as the examples above.

Regarding the concept of semantic description, the M Dictionary provides a robust capability to label the inputs, outputs, and attributes of a mathematical model using words, compound words, noun phrases, and relationships. This forms an intricate, unambiguous ontology. Lee et al. (2010) put forth a detailed example of model description used for manufacturing systems that is applicable to all mathematical models, including those in agriculture. This is helpful in matching a model to specific data.

## 6. Discussion

The prototype detailed in Section 5 demonstrates that a commercially viable system exists for rapid data and mathematical model integration in agriculture. It is a base in research and practice for the eventual evolution of sensor networks to monitor various agricultural variables and to recommend a course of action through automatic data analysis using mathematical models. While this represents a step forward for data-driven agriculture there remain limitations with the current Internet-based architecture. Through computing engineering most of the shortcomings can be overcome, however this will take significant time to achieve in research and practice.

The chief concern lies with the M Dictionary and the universal approach outlined above as a base for all industrial semantics. It soon became clear that this is impracticable.

A better idea is to establish an M Dictionary specifically for the agricultural industry, perhaps to include proper nouns such as the names of agricultural chemicals. Even with this approach, coming to agreement regarding content is a major, but necessary undertaking.

A second concern involves the issue of semantic distance. Since M word entries are Wiki enabled, a chance exists that two separate definitions for a single word could be very close in meaning. The process of refereeing between definitions is time consuming and requires consistency. However, once this process is complete a valuable set of definitions exists.

It is worth noting that other researchers express similar concerns. An early paper from the medical community notes that there are "variable semantic distances between concepts of a same class, and dynamic variations of these distances due to contextual representation" (Pellegrin et al., 1994). For rigid ontologies expressed in RDF or OWL, this observation based on a reasonable experiment presents a major issue in practical Semantic Web implementation. To date, the lack of precision for spoken and written English continues to represent the ultimate computer science challenge in established robust intelligent systems.

Perhaps the greatest concern lies in the business model needed to carry this research into practice. The current architecture is not expensive from a fixed cost perspective (prototype < $1000 in software costs), and with some development could be a totally open system. Costs do accumulate as scale expands. Cash outlays for programing and maintenance increase with the addition of new data streams and the need for complex spatial interpolation to extend the value of point estimates for such measurements as temperature. These types of programing require specialized skills that are not common in the computer industry. Likewise, this lack of a skilled and experienced base has been a major limitation to the growth of Semantic Web applications.

The key element for the entire architecture to work in practice is the choice of a capable standards organization that has experience in implementation and has a genuine desire to improve the flow and use of data for decision-making in day-to-day agricultural operations. Motivated standards organizations working with interested industrial partners play an enormous role. This is true for all of industry, including agriculture.

Perhaps the most important aspect of M Language research is the idea of an auxiliary language and the long-term capability to create a synthetic, interoperable computing environment to deal with the explosion of data input from both humans and sensors. This explosion of data from different sources presents an entirely new problem – namely, dealing with a large number of different standards meant to organize data for sharing and analysis (Schuster and Lee, 2010).

Broadly labeled "convergence", the problem is well known to end users and industrial developers alike, who often find the number of technologies and standards hard to comprehend let alone synthesize into meaningful and productive systems. The question becomes, "Could even more productivity arise from interoperability between various standards?"

Considering the significant amount of money spent by industry each year in systems integration projects, the potential for monetary savings is great. Beyond financial savings, the interoperability of standards should significantly speed the flow of data and information within the broader economy. Increasingly, commerce views the explosion of standards as a hindrance to greater productivity gains. This is a unique paradox because the value of standards in driving commercial growth is well proven by the size and effectiveness of such organizations as GS1,[11] which administers industrial standards for bar codes, radio frequency identification (RFID), and item master data systems, among other things. However, as the number of successful standards increases, it becomes much more difficult to interoperate these into robust systems.

## 7. Conclusion

Through the use of an auxiliary language, this paper puts forth a computer architecture that integrates data and mathematical models. A real world application proves that the system can work to solve practical problems at the farm level.

Computational agriculture holds promise for many new areas of research and practice. The immense increase in computing power and the capability of the Internet to connect with advanced databases such as 11g using XML as a data carrier will inevitably bring about a new era. To facilitate this potential, the industry needs to come to agreement on the standards and computer architectures needed to enable data-driven agriculture.

Future research and practice for the M Language will focus in several areas, including the application of design of experiments to an entire field (Willers et al., 2008) and an expansion of harvest risk analysis (Allen and Schuster, 2004). These two areas focus on increasing the productivity of agriculture through optimizing inputs such as agricultural chemicals and optimizing the risk of losing a portion of the crop because of poor weather conditions. Both areas are data intense and represent the frontier for computational agriculture.

[11] http://www.gs1.org/.

## Acknowledgements

We acknowledge that this work has been partially funded by the MIT Data Center Consortium and the USDA Technical Assistance for Specialty Crops program.

The authors would like to thank the Editor and the reviewers for the constructive and professional review of our manuscript. We also gratefully acknowledge the support of the Indian River Citrus League and the cooperative assistance of the Caribbean Fruit Fly TASC Research Group.

## References

Allemang, D., Hendler, J., 2008. Semantic Web for the Working Ontologist – Modeling in RDF, RDFS and OWL, first ed. Elsevier.
Allen, S.J., Schuster, E.W., 2004. Controlling the risk for an agricultural harvest. Manufacturing & Service Operations Management 6, 225–236.
Angwin, L., 2009. Are dictionaries becoming obsolete? The Wall Street Journal (September 7).
Associated Press, 2005. Science Journal Says Wikipedia is Fairly Accurate, December 15.
Booth, D., Haas, H., McCabe, F., Newcomer, E., Champion, M., Ferris, C., 2004. Web Services Architecture, W3C Working Group Note.
Brock, D.L., Schuster, E.W., Allen, S.J., Kar, P., 2005. An introduction to semantic modeling for logistical systems. Journal of Business Logistics 26, 97–117.
Brock, D.L., Schuster, E.W., Kutz Sr., T.J., 2006a. An Overview of the M Language. MIT, The Data Center.
Brock, D.L., Schuster, E.W., Thumrattranapruk, C., 2006b. Multi-Lingual Display of Business Documents. MIT, The Data Center.
Chandler, D.T., 2009. Rethinking Artificial Intelligence. MIT, News Office, December 8.
Deacon, T.W., 1997. The Symbolic Species, first ed. W.W. Norton & Company, December 8.
Dewdney, M.M., Graham, J.H., 2011. Florida Citrus Pest Management Guide: Citrus Canker. University of Florida, Institute of Food and Agricultural Sciences, pp. 129–134.
Englert, B., Lam, S., Chiou, C., 2010. XML Based Supply Chain Integration at the Los Angles and Long Beach Ports. METRANS: Transportation Center USC and CSULB. August 11.
Feigenbaum, E.A., Feldman, J., 1963. Computers and Thought, first ed. McGraw-Hill.
Fellbaum, F. (Ed.), 1998. WordNet: An Electronic Lixical Database (Language, Speech, and Communication), first ed. The MIT Press.
Fensel, D., Hendler, J., Lieberman, H., Wahlster, W., 2003. Spinning the Semantic Web, first ed. The MIT Press, p. 363.
Geoffrion, A.W., 1976. The purpose of mathematical programming is insight, not numbers. Interfaces 7, 81–92.
Geoffrion, A.W., 1987. An introduction to structured modeling. Management Science 33, 547–588.
Geoffrion, A.W., 1989. The formal aspects of structured modeling. Operations Research 37, 30–51.
Hayakawa, S.I., 1964. The Language in Thought and Action, fourth ed. Hardcourt Brace Jananvich, Inc.
Howard, P., 2005. A way with words. The Wall Street Journal (August 23).
Jones, C.V., 1996. Visualization and Optimization, first ed. Kluwer.
Lee, H.-G., Schuster, E.W., 2007. MRP-as-a-service: An Alternative Way to Execute Models for Business Processes. Business Technology Trends & Impacts Advisory Service, Cutter Consortium, November 29.
Lee, H.-G., Schuster, E.W., 2009. Intranet architecture for the semantic enterprise. Cutter IT Journal 22, 31–37.
Lee, H.-G., Schuster, E.W., Allen, S.J., Kar, P., 2010. The open system for master production scheduling: information technology for semantic connections between data and mathematical models. International Journal of Operations Research and Information Systems 1, 1–15.
McDonough, J., 2008. Structural metadata and the social limitation of Interoperability: a sociotechnical view of xml and digital library standards development. In: Proceedings of Balisage: The Markup Conference.
Murakami, E., Saraiva, A., Ribeiro Jr., Luiz C.M., Cugnasca, Carlos E., Hirakawa, Andre R., Correa, P.L.P., 2007. Computers and Electronics in Agriculture 58, 37–48.
Nash, E., Korduan, P., Bill, R., 2009a. Applications of open geospatial web services in precision agriculture: a review. Precision Agriculture 10, 546–560.
Nash, E., Dreger, F., Schwarz, J., Bill, R., Werner, Armin, 2009b. Development of a model of data-flows for precision agriculture based on a collaborative research project. Computers and Electronics in Agriculture 66, 25–37.
Nikkila, R., Seilonen, I., Koskinen, K., 2010. Software architecture for farm management information systems in precision agriculture. Computers and Electronics in Agriculture 70, 328–336.

Pellegrin, L., Bastien, C., Roux, M., 1994. Representation of medical concepts of the thyroid gland by physcians in anatomy and pathology. Methods in Medicine 33, 382–389.
Rogers, J.S., 2000. Citrus canker: science policy and new technologies for management. Citrus Industry 81, 51–54.
Rogers, J.S., McCoy, C.W., 1996. A laptop computer-based decision support tool for managing citrus rust mite populations. Proceedings of the International Society Citriculture 8, 625–631.
Rogers, J.S., Nigg, H.N., Clark, I., Hart, D.S., Ehsani, R., Brock, D.L., Schuster, E.W., 2006. Real-time Internet–PDA-based survey and modeling systems for agricultural insects and diseases: integrated fruit-fly, citrus canker, and greening surveys. Proceedings of the Florida State Horticultural Society 119, 107–119.
Rogers, M.E., Stansly, P.A., Stelinsky, L.L., 2011. Florida Citrus Pest Management Guide: Asian Citrus Psyllid and Citrus Leafminer. University of Florida, Institute of Food and Agricultural Sciences, Berlin, pp. 43–51.
Schmitz, M., Martini, D., Kunisch, M., Mösinger, H-J., 2008. agroXML enabling standardized, platform-independent internet data exchange in farm management information systems. In: Metadata and Semantics. Springer, US, pp. 463–468.
Schuster, E.W., Allen, S.J., 1998. Visualization and finite capacity planning models through computer animation. In: Proceedings of the APICS 41nd International Conference.
Schuster, E.W., Allen, S.J., 1998b. Raw material management at Welch's. Interfaces 28, 13–24.
Schuster, E.W., Allen, S.J., 1999. Finite planning models through computer animation. In: Proceedings of the APICS 42nd International Conference.
Schuster, E.W., Lee, H.-G., 2010. Mobile computing, spatial intelligence, and interoperability for agricultural field data. Cutter IT Journal 23, 28–34.
Schuster, E.W., Allen, S.J., D'Itri, M.P., 2000. Capacitated materials requirements planning and its application in the process industries. Journal of Business Logistics 40, 27–33.
Schuster, E.W., Allen, S.J., Brock, D.L., 2007. Global RFID: The Value of the EPCglobal Network for Supply Chain Management, first ed. Springer.
Storey, V.C., Burton-Jones, A., Sugumaran, V., Purao, S., 2008. CONQUER: a methodology for context-aware query processing on the world wide net. Information Systems Research 19, 3–25.
Veverka, M., 2010. Sky's the Limit. Barron's, January 4.
Vinoski, S., 2003. Integration with web services. IEEE Internet Computing 7, 75–77.
Wessel, R., 2005. Office technology opening up. The Wall Street Journal (October 24).
Willers, J.L., Milliken, G.A., Jenkins, J.N., O'Hara, C.G., Gerard, P.D., Reynolds, D.B., Boykin, D.L., Good, P.V., Hood, K.B., 2008. Defining the experimental unit for the design and analysis of site-specific experiments in commercial cotton fields. Agricultural Systems 96, 237–249.