International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Determinator – A Generic DSS For Hazard Identification Of Species Or Other Physical Subjects

Michel Uiterwijk[1], Leo W.D. van Raamsdonk[2], Sander J.C. Janssen[1]
[1] *Alterra, Institute for Environmental Science, Wageningen, The Netherlands*
[2] *RIKILT, Institute for Food and Feed Safety, Wageningen, The Netherlands*
Michel.Uiterwijk@wur.nl

Abstract: Our living environment poses all kinds of obvious and perhaps less obvious hazards. As an example, Ragwort (*Senecio jacobaea*) and related species of the genus *Senecio* are the main source of toxic pyrollizidine alkaloids. The plants commonly grow in road verges, meadows and production fields and are equally toxic for humans and animals. Therefore, proper identification of such a hazardous object is vital. This paper describes Determinator DSS, a Decision Support System, that aims to cover problems of proper identification of a hazardous object, either in the field or in a lab, and in diverse domains. Determinator DSS is used in several practical applications for ensuring food and feed safety and supporting environmental issues. The paper highlights several software aspects and design features of the DSS that enable it use in diverse circumstances and for different users as a generic solution.

Keywords: *Determinator, DSS, identification, crowd-sourcing, environment.*

1 INTRODUCTION

Our living environment provides all sorts of hazards, which form putative risks for human beings or animals. These hazards can be based on the presence or overpresence of either substances or animal or plant species. Tracking and tracing are intensively used to tackle these hazards, ultimately aimed at ensuring food safety and safety of the environment. In order to collect information on the occurrence and frequency of these hazards, Early Warning Systems, or Emerging Risk Management Systems are implemented (Marvin, et al., 2009). Most of these systems have a human centred approach, and are based on data provided by expert networks.

A specific task that occurs in hazard analysis is the identification of the hazardous substance or species vis-à-vis other substances or species. In the situation of expert networks, the network itself approves the reliability of the identifications. However, in environmental projects this identification often has to be carried out by a user in the field, which is often not a highly trained expert without sufficient expert knowledge to distinguish between all the different substances or species into hazardous and non-hazardous. This imperfect knowledge might lead to miss, or ignore valuable observations.

For example, Ragwort is an exotic plant in the Netherlands, leading to hepatic veno-occlusive disease (HVOD), necrosis of the liver and ultimately death. At lower doses they may lead to liver cirrhosis (EFSA 2011). Plant inspectors or other professionals are often out in the field, which should allow them to identify suspicious plants or animals by using a mobile version of a DSS. Correct identification of Ragwort requires some expert knowledge as it resembles other native plants found in Europe. Also, insect plagues are becoming more and more frequent in the Netherlands, also in trees, with the ‘Oak Processionary’, a caterpillar

(*Thaumetopoea processionea*), whose poisonous setae (hairs) cause skin irritation and asthma. For these trees it is important to timely identify the infected oaks to inform hikers and others of the danger of the caterpillars.

In these examples, identification of the frequency and occurrence of a hazard can be implemented by means of crowd-sourcing, thus getting as many people as possible for doing (correct) observations. There are a lot of observation networks with experts, such as for the Insect Pests Monitoring program in the Netherlands [Moraal et al, 2011]. To make this observation network larger and include non-experts as well, expert knowledge should be disclosed to guarantee the quality of observations, which can be supported by the application of (mobile) DSSs.

This paper describes an expert system for identification of hazards, the Determinator Decision Support System. Determinator DSS is argued to be a generic tool for hazard identification, and this is demonstrated in this paper by explaining IT features of the Determinator DSS, that aid its generic nature. Furthermore, the framework for hazard identification as conceptualized for Determinator DSS and the methods used in developing Determinator DSS are discussed and the value of a DSS in the framework of crowd-sourcing is presented.

2 DEVELOPMENT OF THE DETERMINATOR DSS PLATFORM

2.1 Datamodel

The approach of Determinator is that all its specific applications share a generic datamodel structure. This datamodel structure contains the following elements: *Targets, Features, FeatureGroups, TargetMatches, Labels and TreeItems.*

Targets are the concrete entities (objects), such as plants or animals. Targets can also be categorized by means of adding *Labels* to it, such as a family name or region(s) of occurrence.

Features are the visible entity characteristics (like plant height or leaf shape) of a target, which can be organised in *Feature Groups*. Determinator also allows two different types of features: Continuous and Discrete. Discrete features are described in three states (like none/few/many), whereas continuous features have a numerical value for minimum and maximum. Every value between these minimum and maximum values can be used to describe the target.

Decision rules are described in *TargetMatches* and contain the mapping of Targets and Features, by specifying valid feature states for each target. The decision rules in TargetMatches allow to use the Booleans AND and OR, e.g. IF a AND b THEN p OR q. This rule exemplifies the situations that feature states “a” and “b” both apply to either target p or target q.

Finally, *TreeItems* contain the representation of a determinate decision tree, where each node can link to either a Target or another node. Basically, every lemma in a tree is based on the simple rule IF m THEN p ELSE q. The (combined) feature state “m” can combine more than one simple state, e.g. “a” and “b”.

2.2 Tools

Tools that were developed for the Determinator DSS platform are a *Player* tool for identification by the end user and a *Developer* tool for the modelling expert.

The *Player* tool is intended for visual on-the-spot identification by the end user and its functionality can be divided into two categories: explorative and systematic identification.

Explorative identification is possible through the following options in the menu (Figure 1, left screen): *Browse, Filter, Select, Search, Overview and Compare*.

Browse (Figure 1, middle screen) lists all the targets in the datamodel and shows description, labels and one or more images of each target. Keyboard shortcuts are available for quick browsing. The *Filter* option is used to list all targets that have a specific feature value, like showing plants that are flowering in a specific month. *Select* does the same as *Filter* but filtering is now done with labels instead of feature values. The *Search* option provides text-based filtering in name and description fields and *Overview* shows all images of all targets in one window.

Compare functionality can be used to compare two targets by feature, by showing all feature values and providing them with a match colour: green means a full match, red means no match and yellow means a partial match on that feature.
A comparison table of Ragwort and Oxford Ragwort is shown (Figure 1, right screen), where can be seen that the feature of incisions has no match (red), so we can distinguish these two targets by this feature alone (feature value is double vs. simple). Partial matches (yellow) can be seen for flowering month, plant height and leaf location. The rest of the displayed features indicate a full match (green).

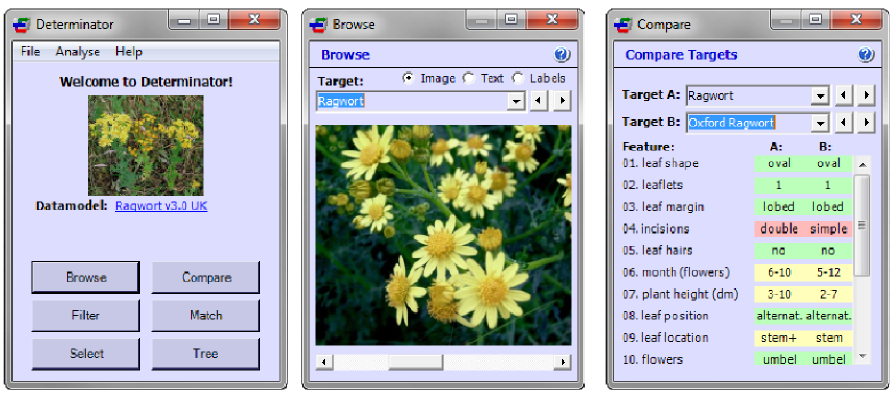

Figure 1 - Player screens for Menu, Browse and Compare

Systematic identification is done through options *Match* or *Tree* in the menu.

Match functionality first asks the user to select the feature groups of interest. All features in those selected groups are given as questions for the user to answer, by selecting one of the options at the bottom of the window (Figure 2, left screen). In case no definitive choice can be made by the user, there is a N/A option. Answering questions gives a matching percentage for each target. During Match identification, an indicator is shown in the screen header how many targets are still full matches (100% Targets). The user can click the Results button to get a complete list of targets and their matching percentage (Figure 2, middle screen).

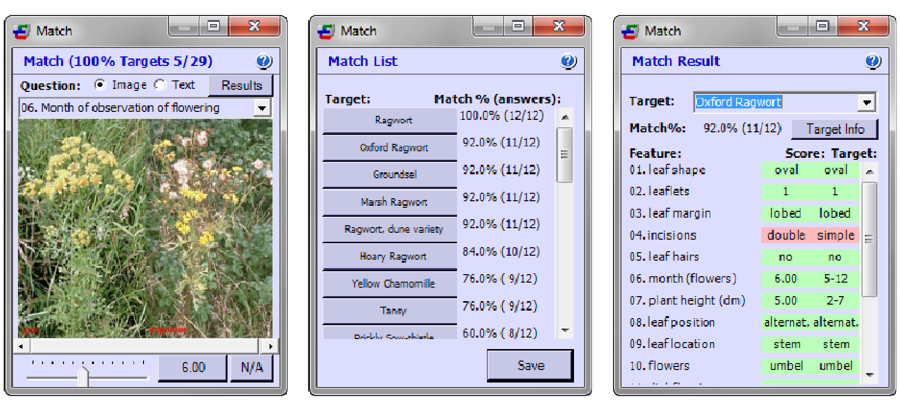

Figure 2 - Player screens for Match Questions, List and Result

Finally, *Tree* is walking along a determinate decision tree (of *TreeItems*) by choosing one out of two possible options, resulting eventually in a target (Figure 2, right screen). The user can also navigate back one step or directly to the root of the Tree to start over (Figure 3).

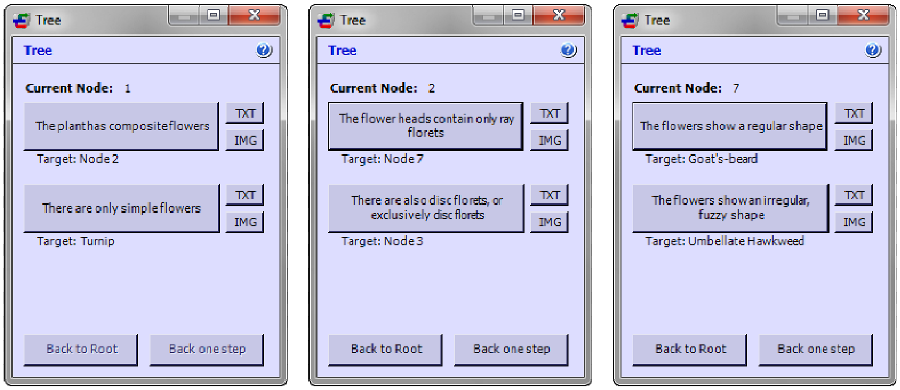

Figure 3 - Player screens for Tree (Nodes 1, 2 and 7)

The *Developer* tool can be used to create new or edit existing datamodels and all functionality is presented in different tabs (Figure 4).

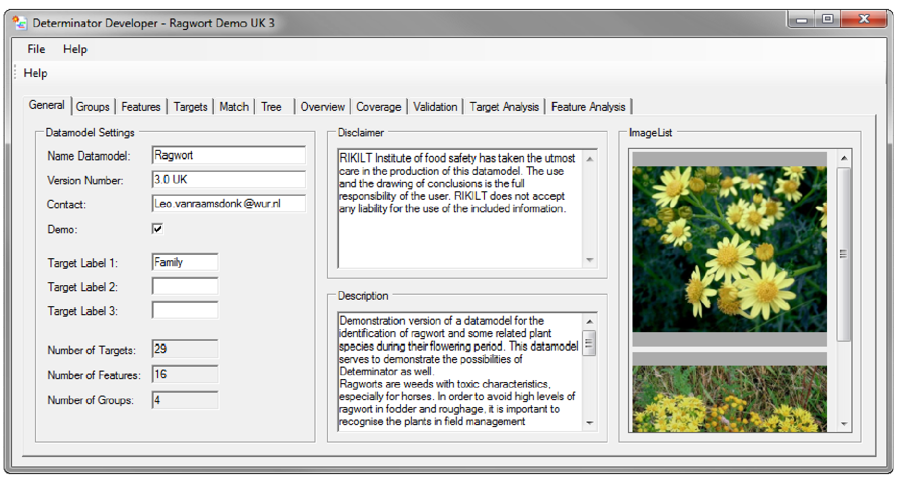

Figure 3 - Developer screen showing general information of a datamodel

The strength of the Developer tool is not only the ease of creating a new datamodel, but also the analysis and validation functionality. Target Analysis (Figure 5) provides a table displaying overlap between targets. A cell with a green colour indicates that the two targets have at least 1 differentiating feature and can be distinguished by this feature.

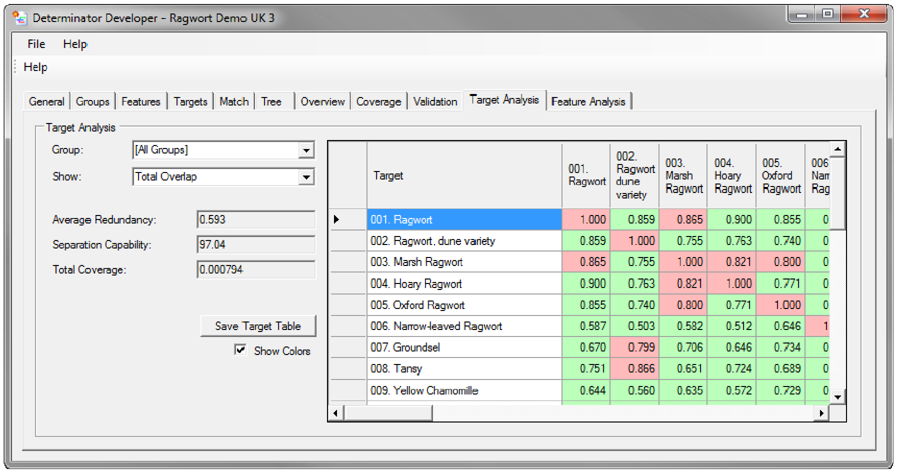

Figure 4 - Developer screen showing Target Analysis

It is vital that a datamodel is adequate and effective. For more information about these issues, see [Raamsdonk et al (in press)].

Both tools are built in Microsoft C#.NET and work in all versions of Windows, but the Player also works on a Windows Mobile device. All knowledge for a datamodel is stored in an encrypted xml and images are stored in a separate folder.

Currently, a mobile web application is being built in HTML5/jQuery to support multiple mobile platforms (like iOS, Android and Windows Phone) by running in a HTML5 compatible web browser. Also there are plans to link identification with observation networks by sending identification results over the internet. To guarantee the quality of the observation it can be documented with photographs and GPS coordinates, because today's smartphones have those capabilities.

3 DISCUSSION

The successful application of Determinator, or in more general sense of decision support systems, depends on the domain of application. As mentioned in the Introduction, the domains of environmental studies and of feed and food safety will be discussed.

Projects for environmental monitoring studies, such as the Insect Pests Monitoring Network [Moraal et al., 2011], and monitoring programs for the effect of climate change on biodiversity (Battisti et al., 2010) are partly based on a reliable identification of the target species. Also in the framework of the relationship between environmental changes and health issues, a proper identification of the hazards is necessary. An example is the recent legally based monitoring of seeds of *Ambrosia* species in bird feed, which are able to cause severe allergy problems (Frick et al., 2011). These and other projects (see e.g. Key to Nature: Nimis and Lebbe, 2010) can benefit from support systems for validation of identification of the target, i.e. the species. At the same time, several of these projects use crowd-sourcing as a process for gathering data on occurrence and frequency of putative hazards. In order to facilitate the identification, the following strategy could be implemented in a DSS for improving the quality of the uploaded data: *Experts* can report observations directly (without systematic identification) and *non-experts* have to follow systematic identification, either through a simple (decision tree) or detailed identification (match matrix) process. Observations by these non-experts, accompanied with proper documentation (photo and automatically generated GPS coordinates and description), can later be confirmed by an expert based on the uploaded data.

In the domain of feed and food safety a range of monitoring, detection or early warning systems is implemented (Marvin et al., 2009). As an example, the

European Union has started the Rapid Alert System for Food and Feed safety (RASFF), which is used in Europe for fast communication on deviations from EU legislation and to further improve food and feed safety (Kleter et al., 2009). In this system, hazards are identified by national competent authorities of EU member states, which are classified as experts. In cases of e.g. botanic impurities, notifications depend on both the level of expertise of the competent authority and on the presence of a monitoring system for botanic impurities (van Raamsdonk et al., 2009). The discrimination between rarely occurring hazardous botanic contaminations and legally applied food and feed ingredients is sometimes difficult to make. A DSS platform such as Determinator can aid to this identification process. The case of Ragwort is equally important for the domains of food and feed safety and for environmental monitoring.

In general, entities that are defined as (physical) objects can be described in models for DSSs which are focusing on pictorial information. These descriptions can support the data which is generated by the process of crowd sourcing. The combination of a recognised *hazard* and data on frequency and occurrence (from crowd-sourcing) can provide useful information for the assessment of the accompanying *risk*.

ACKNOWLEDGMENTS

This work was partly funded by the Dutch Ministry of Economic Affairs, Agriculture and Innovation.

REFERENCES

European Food Safety Authority (EFSA), 2011, Opinion of the Scientific Panel on Contaminants in the food chain on pyrrolizidine alkaloids in food and feed. EFSA J. 2406:1-134.

Battisti, A. et al., *BACCARA - Biodiversity And Climate Change, A Risk Analysis, Deliverable 2.1 - List of damage caused by pests and diseases in relation to climate change in major European forest types,* EU FP7 (226299), 2010

Frick G., Boschung H., Schulz-Schroeder G., Russ G., Ujčič-Vrhovnik I., Jakovac-Strajn B., Angetter D., John I., Jørgensen JS., 2011. Ragweed (*Ambrosia* sp.) seeds in bird feed, Biotechnologie, Agronomie, Société et Environnement 15 special volume 1 : 39-44.

Kleter GA, Prandini A, Filippi L, Marvin HJP. 2009b. Identification of potentially emerging food safety issues by analysis of reports published by the European Community's Rapid Alert System for Food and Feed (RASFF) during a four-year period. Food Chem Toxicol. 47(5):932–950.

Marvin, H.J.P.; Kleter, G.A.; Pradini, A.; Dekkers, S.; Bolton, D.J., Early identification systems for emerging foodborne hazards Food and Chemical Toxicology 47 (5). - p. 915 – 926, 2009.

Moraal, L.G., Jagers op Akkerhuis, G.A.J.M., Changing patterns in insect pests on trees in The Netherlands since 1946 in relation to human induced habitat changes and climate factors—An analysis of historical data, *Forest Ecology and Management*, Volume 261, Issue 1, 1 January 2011, Pages 50–61

Nimis, P.L., Lebbe R.V. (Eds.), 2010. Tools for Identifying Biodiversity: Progress and Problems, Proceedings of the International Congress Paris, France, September 20-22, 2010. http://dbiodbs1.units.it/bioidentify/index.html

Raamsdonk, L.W.D. van, VanCutsem, J. Jørgenson, J.-S., 2009. A survey on the presence of undesirable botanic substances in feed in the European Union. Biotechnol. Agron. Soc. Environ. 13(S): 33-38.

Raamsdonk, L.W.D. van, Mulder, P., Uiterwijk, M., Identification tools as part of Feedsafety research: the case of ragwort, *Tools for Identifying Biodiversity: Progress and Problems. Proceedings of the International Congress, Paris, September 20-22, 2010*, Trieste, Italy, 213-216, 2010.

Raamsdonk, L.W.D van, Uiterwijk M., InTech (submitted) Reliability and evaluation of identification models.