

Results of the GISIN needs assessment survey

Introduction

The Global Invasive Species Information Network has convened an Invasive Species System Task Group as part of the Taxonomic Database Working Group. The task at hand is the development of a system for the exchange of invasive species information over the Internet. A critical step in the process is to determine what is required by the eventual users of this protocol, including organizations that will be providing data (providers), organizations that will be consuming data (consumers), the end-users, and stakeholders. Below are the results of a needs assessment survey to obtain this information, which ran from 15 December 2006 through 15 February 2007.

Summary of Results

There were 137 respondents from 41 countries, the vast majority (80%) identifying themselves as providers of invasive species data and consumers of data through web services. Most data providers (77%) potentially offer spatial/temporal information, profiles/species pages (65%), and checklist information (59%).

Data consumers seek all three kinds of information (spatial/temporal, profiles, and checklist). For spatial/temporal information, respondents seek date of observation (91%), name or type of location (83%), and spatial coordinates (75%). For species accounts or profiles, they seek a wide variety of data types, most popular being (in descending order) life history/ecology, habitat descriptions, impacts, geographic range, pathway information, images, management information, descriptions, references, and requirements or tolerances. For checklists, respondents seek (in descending order) occurrence data, habitat affected, introduction and dispersal information, invasiveness, local impacts, date of introduction, and origin.

To filter or search for spatial/temporal information, species profiles, and checklist information, virtually all respondents (93-96%) use scientific name.

There were some areas of the questionnaire where the majority of the respondents were not knowledgeable: half do not know the level of web services their organization provides and/or uses; 80% do not know what existing protocols are appropriate for invasive species information management; and 75% do not know what schemas or grammars would be acceptable to copy or extend to implement the proposed system.

The overwhelming operating system preference (96%) is MS-Windows, followed by Linux (24%). Preferred Internet frameworks were ASP (57%), JSP, 51%, and PHP (39%). Preferred programming languages were Java (58%) and PHP (49%). The majority offer Microsoft Internet Information Services (69%) for hosting web services, followed by Apache HTTP Server (46%).

Concerning the respondents' needs/wants related to the toolkit for implementing the proposed system, 80% would like a software toolkit to install and configure on their own server.

For a summary of the comments that were submitted by respondents, see the appendix on page 8.

Acknowledgements

Thank you very much to those who responded to the survey. Special Thanks goes to those who made this survey possible, especially Michael Browne, Jim Graham, Hannu Saarenmaa, Elizabeth Sellers, Annie Simpson, and last but not least Jeremy Kranowitz of The Keystone Center Inc., who administered it.

Results of the GISIN needs assessment survey

Basic Terminology

Invasive species information can be broken into at least three information categories: Profile, Checklist, and Spatial/temporal.

Profile information is a species page or fact sheet information that describes the invasive species itself, and can include life history, images, management, expertise, and other secondary information.

Checklist information contains location-specific observations (e.g. alien or native, invasive or not), its date and mode of introduction, whether it is considered a problem, if has been controlled, etc.

Spatial/temporal information describes the position of an invasive species on the earth at a particular point in time and can be represented by a point, polygon, or reference to a geographic location such as a country, park, lake, etc.

NOTE: we are using the term "invasive species" in a very broad sense, i.e., including any non-native species that are potentially harmful to the environment, livelihoods, or human health. A section with additional technical definitions can be found at the end of the survey results.

Additional Technical Definitions that were provided to participants in this survey:

Web server: a software package that allows a computer to provide web pages and web services over the Internet. Examples: Microsoft Internet Information Server (IIS), Apache, and Tomcat.

Internet framework: an extension to, or a feature of, a web server that allows programming languages to be used to add functionality to a web solution. Frameworks are often identified by the file extension of the scripting languages they support, for example PHP (PHP: Hypertext Preprocessor), Cold Fusion (CFM), Active Server Pages (ASP), Java Server Pages (JSP), and Common Gateway Interface (CGI). Another Internet framework example is Apache Axis.

Markup language: a specific syntax and grammar that allows the visual representation of relatively complex, text-based, data. Examples: HTML (hypertext markup language) and XML (extensible markup language).

Programming language: a defined syntax and grammar for communicating a sequence of operations for a computer to complete. Programming languages include scripting languages (e.g. PHP, JavaScript, Visual Basic for Applications, Python) and compiled languages (e.g. C++, C#, Visual Basic). Java is a special case and is partially compiled.

Protocol: a defined sequence of messages structured with a specific syntax and grammar for communicating a certain type of information between two computers. Protocols are typically used in networking to pass packets of information between computers. Examples: HTTP (hypertext transfer protocol), FTP (file transfer protocol), SMTP (simple mail transfer protocol).

Web solution: a web-based, complete answer to a problem. Example: GISIN is proposing a web solution for the exchange of invasive species data.

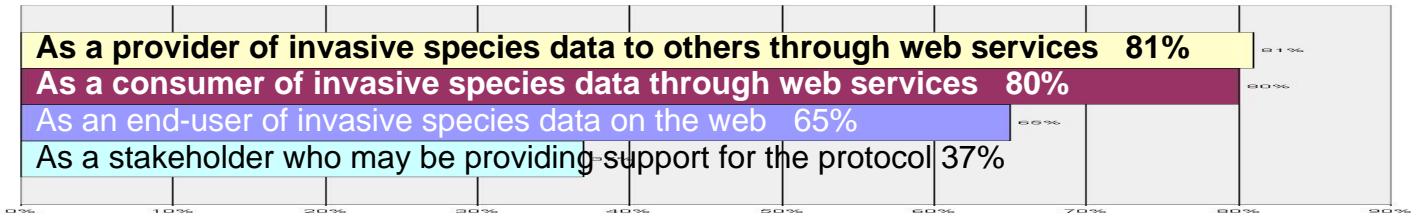
Compiled programming language: a programming language that is converted from a human readable form to a machine readable form before it is executed (run). Most operating systems and commercial applications are written in compiled languages because they typically execute faster and the source code is not available to end users. Examples: C++, C#, Visual Basic.

Scripting programming language: a programming language that is converted from human readable form to machine readable form on a line by line basis. Scripts are easier to create but take longer to execute than compiled programming languages. Examples: PHP, JavaScript, Visual Basic for Applications, Python.

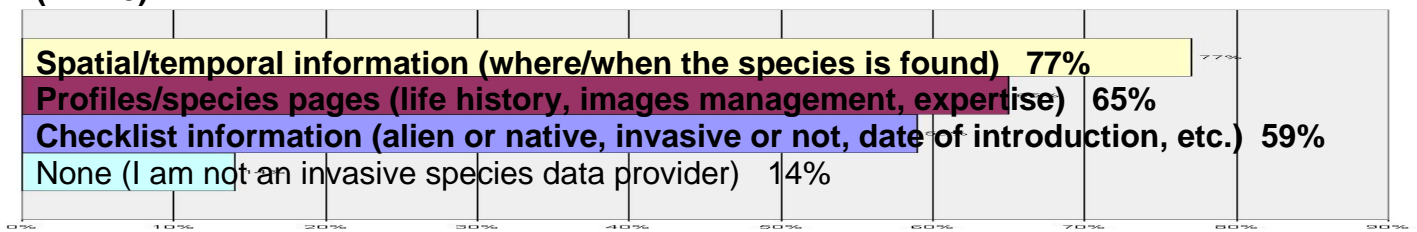
Results of the GISIN needs assessment survey

Basic Questionnaire Results (Additional comments submitted by participants in this survey are summarized in the Appendix on page 8.)

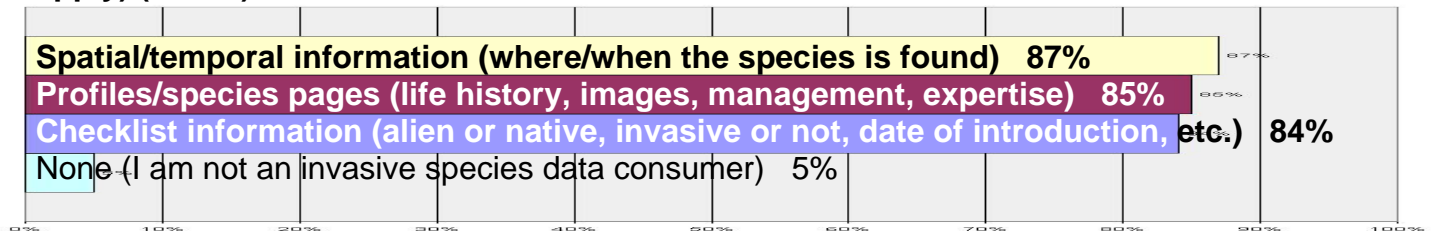
How will you potentially interact with invasive species data? (Check all that apply) (n=136)



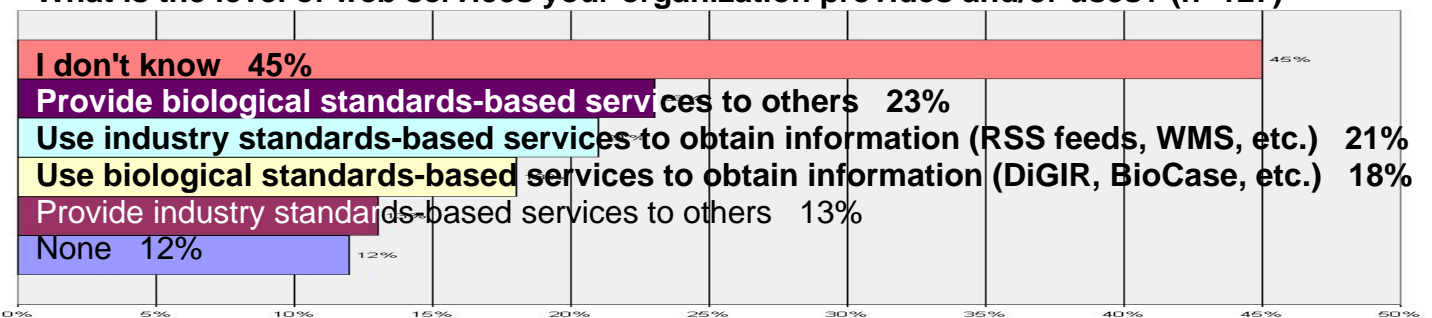
If you are a data provider, what type of data can you potentially offer? (check all that apply) (n=128)



If you are a data consumer, what is the general type of data you wish to have? (check all that apply) (n=132)

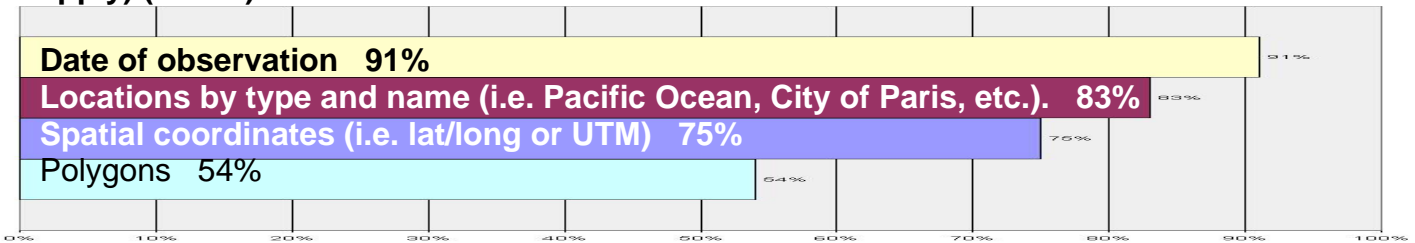


What is the level of web services your organization provides and/or uses? (n=127)

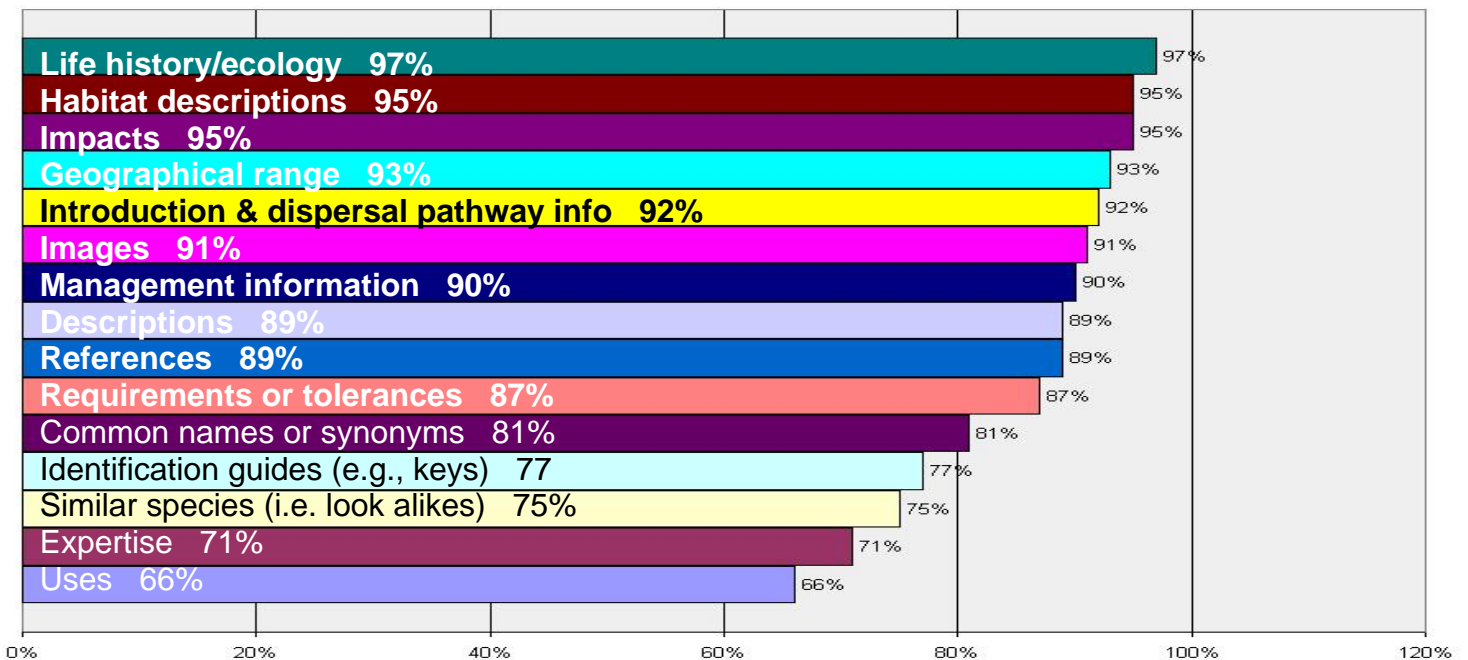


Results of the GISIN needs assessment survey

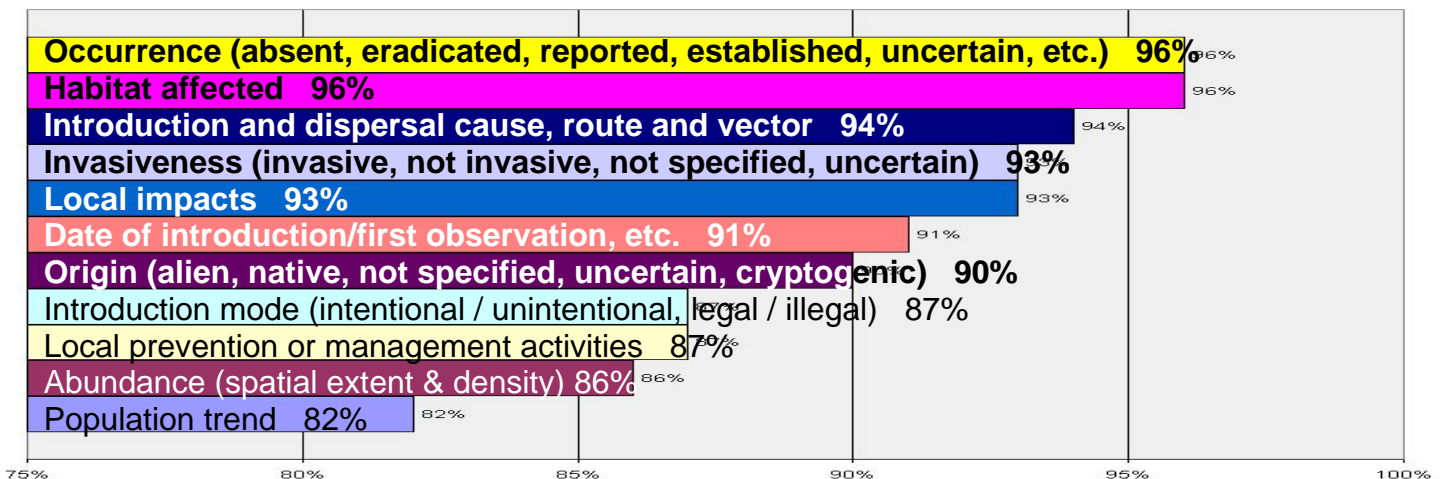
Which types of data are you interested in for spatial/temporal information? (check all that apply) (n=131)



Which types of data are you interested in for profiles/species pages? (check all that apply) (n=133)

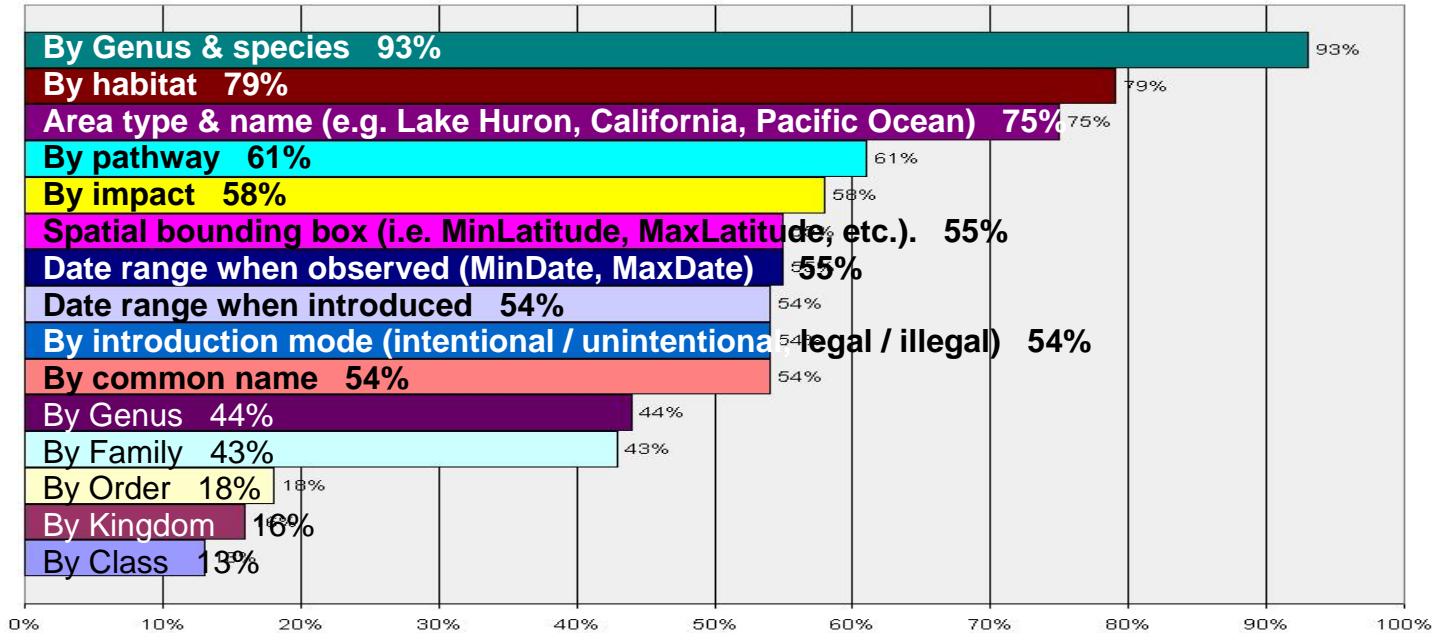


Which types of data are you interested in for checklist information about invasive species? (check all that apply) (n=134)

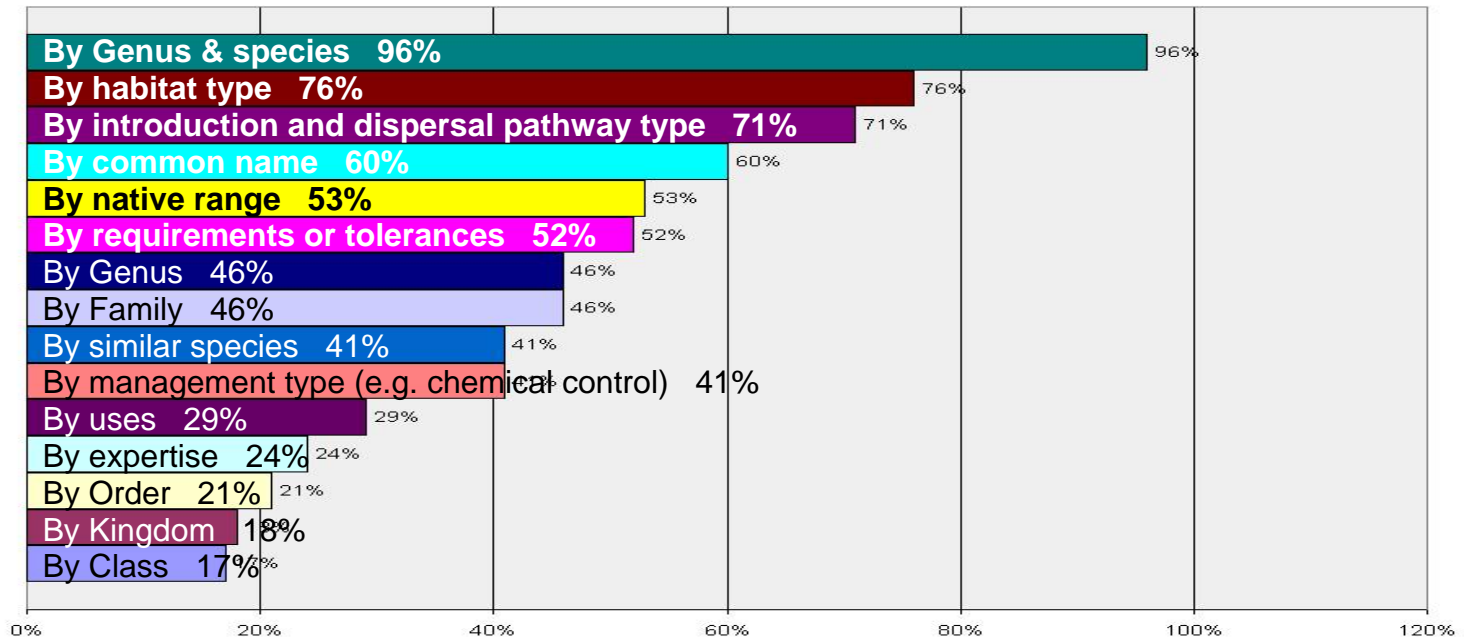


Results of the GISIN needs assessment survey

Which of the following would you use to filter spatial/temporal information? (check all that apply) (n=134)

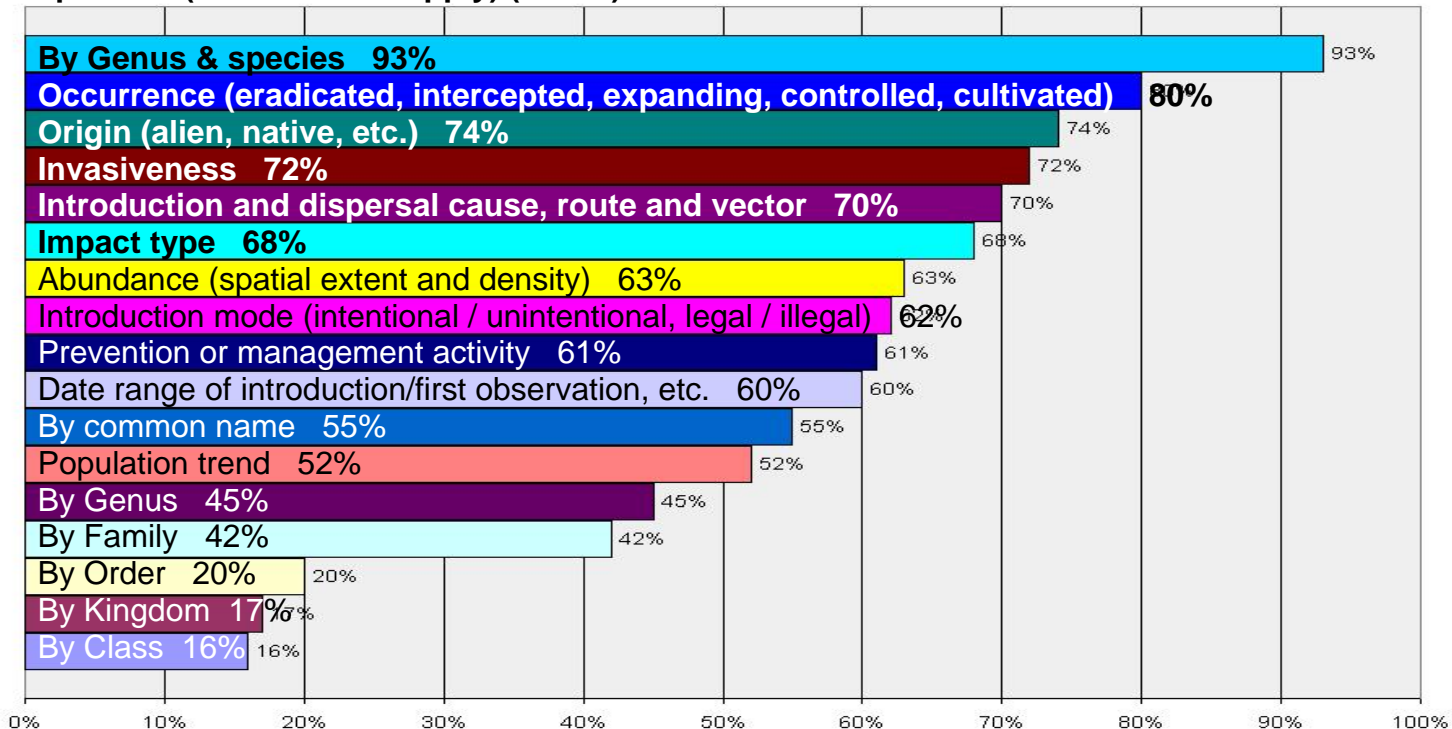


Which of the following would you use for filtering profile/species page information? (check all that apply) (n=134)

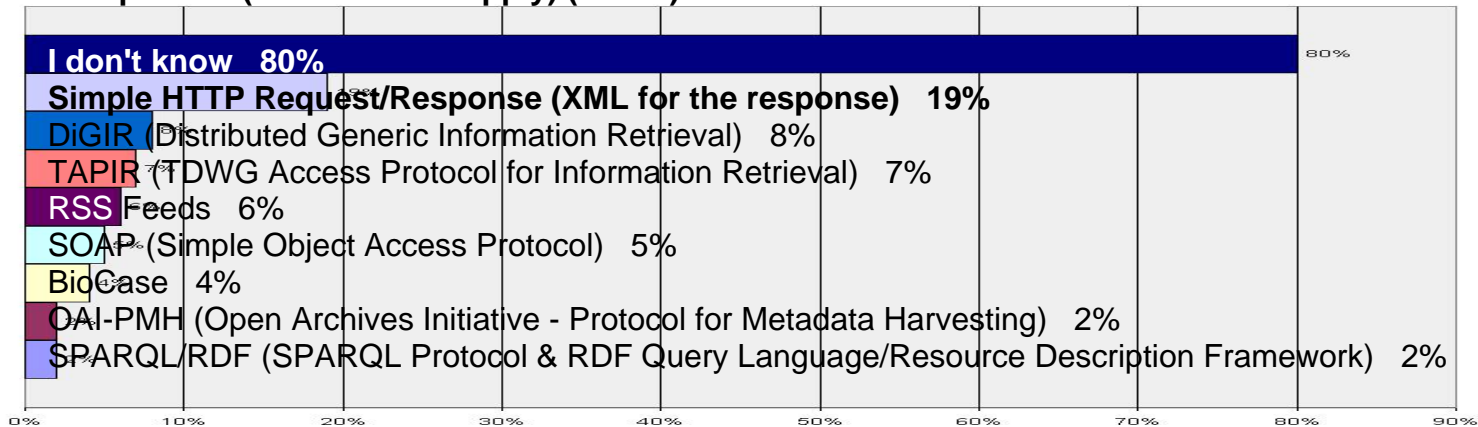


Results of the GISIN needs assessment survey

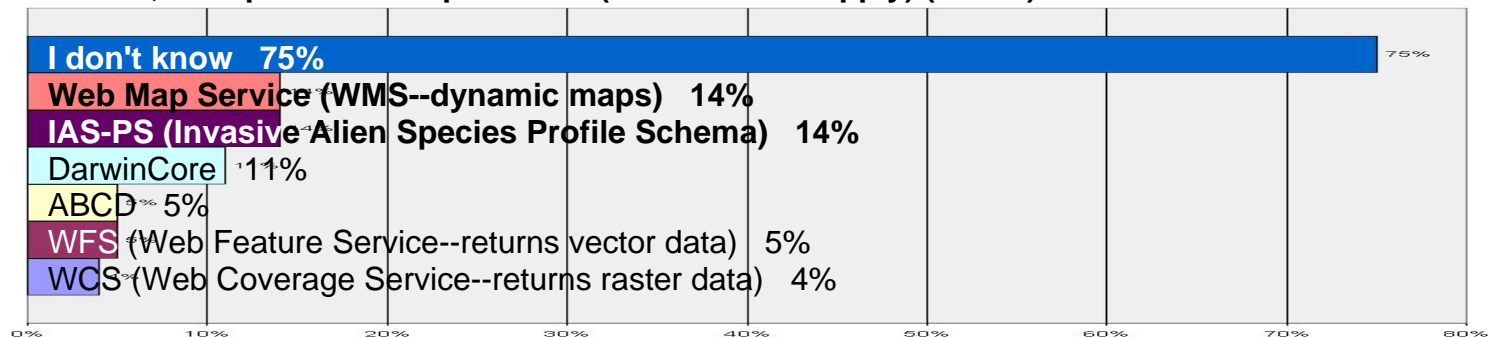
Which of the following would you use for filtering checklist information about invasive species? (check all that apply) (n=134)



If we were to base the protocol on an existing protocol, which of the following would be acceptable? (check all that apply) (n=134)

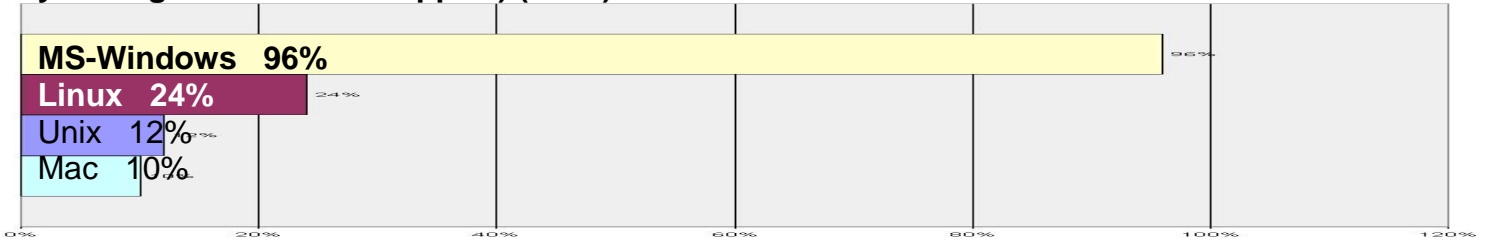


Which of the following types of schemas/grammars would be acceptable for us to copy or extend, to implement this protocol? (check all that apply) (n=132)

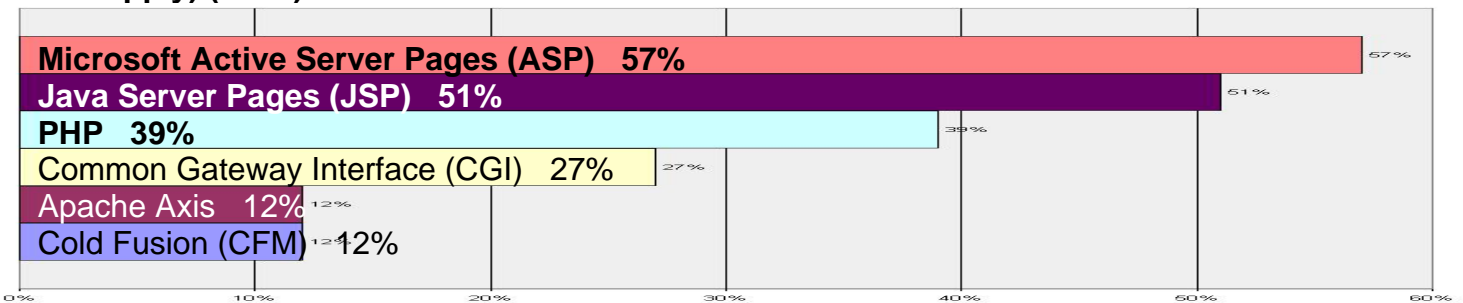


Results of the GISIN needs assessment survey

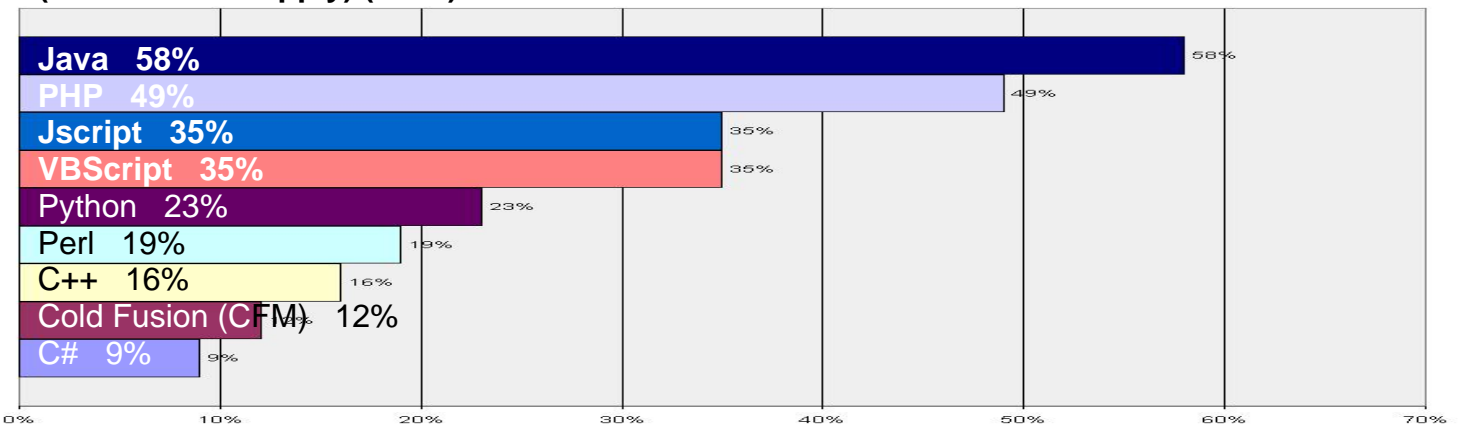
Which of the following operating systems are acceptable for a toolkit to support? (check all your organization can support) (n=84)



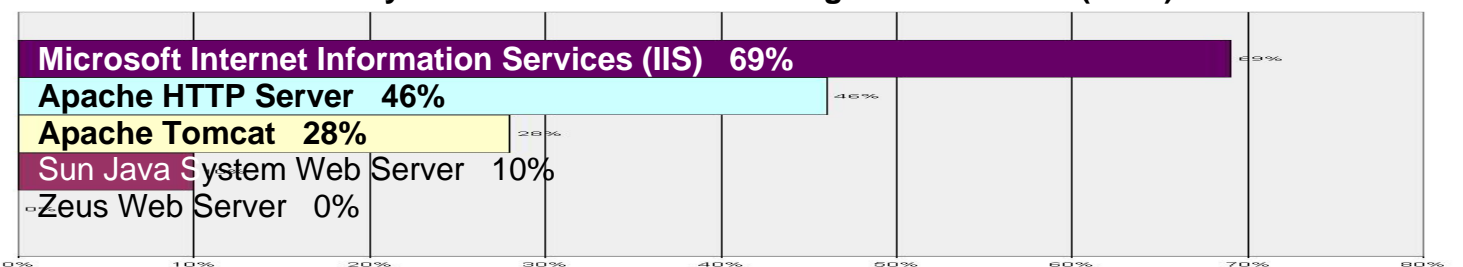
Which of the following Internet frameworks do you feel are acceptable for a toolkit? (check all that apply) (n=49)



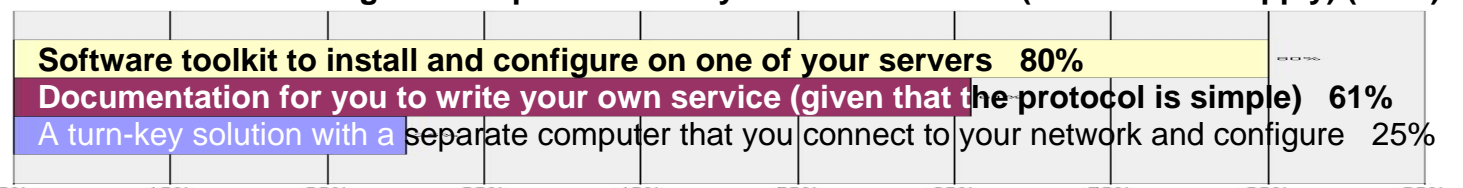
Which of the following programming languages do you feel are acceptable for a toolkit? (check all that apply) (n=43)



Which web servers do you have available for hosting web services? (n=39)

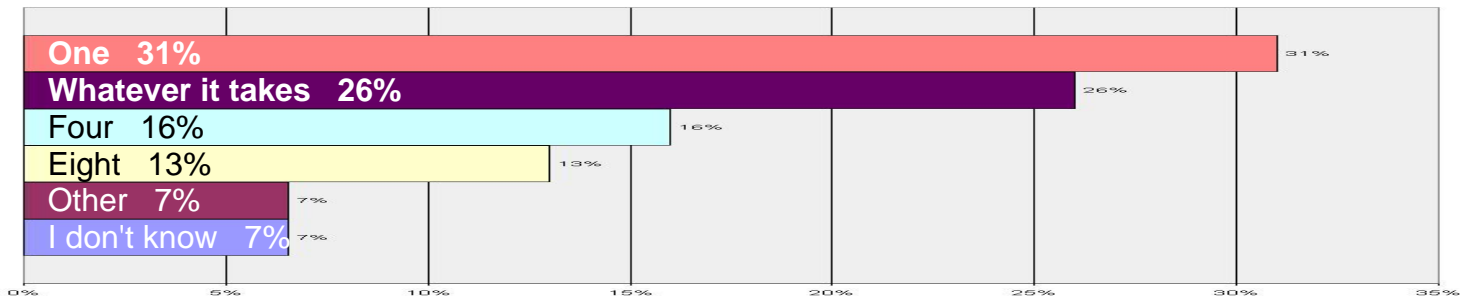


Which of the following toolkit options would you find desirable? (check all that apply) (n=51)

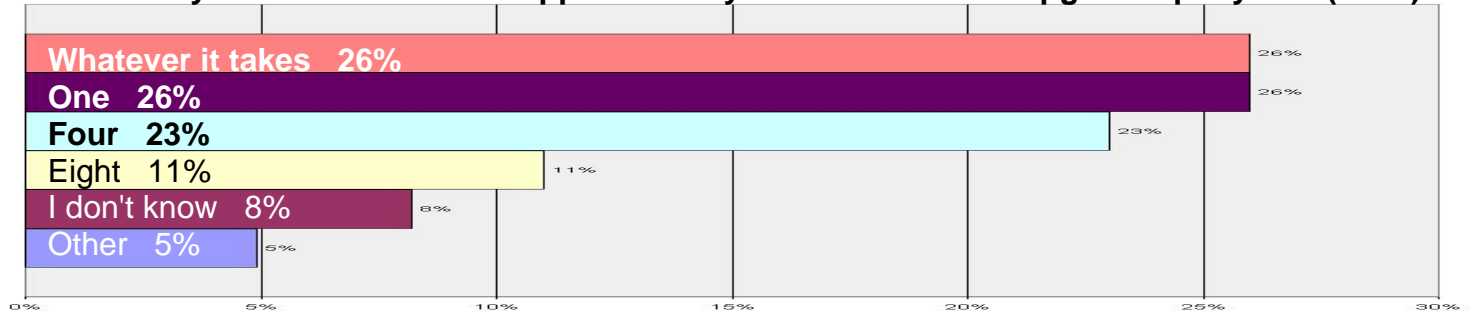


Results of the GISIN needs assessment survey

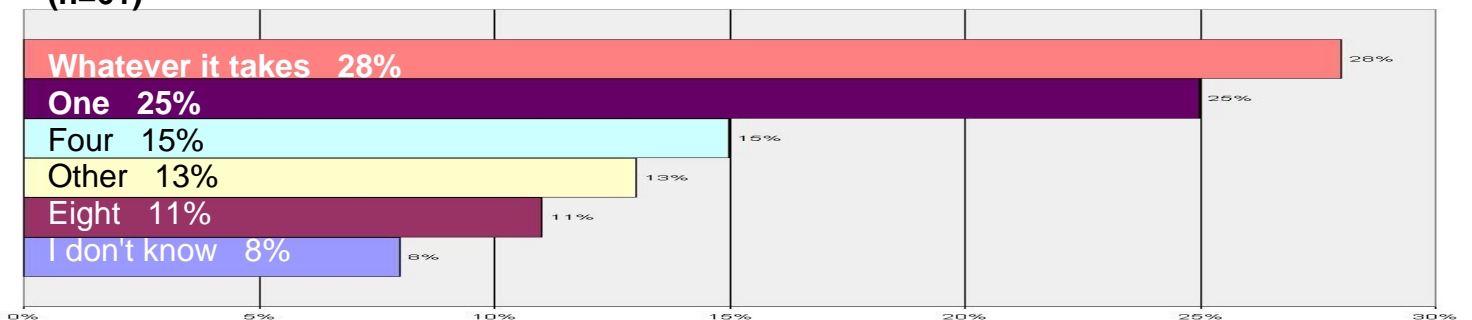
How many hours of technical support could you allot to install this toolkit? (n=61)



How many hours of technical support could you allot for toolkit upgrades per year? (n=61)



How many hours of development could you allot to refine the system we are developing? (n=61)



APPENDIX: Summary of submitted comments

Of the 136 respondents, 75 provided at least one detailed comment. A brief summary of some of the comments is provided here:

Please list how will you potentially interact with invasive species data.

- Provide free access to all IAS databases important to rice and rice based farming systems
- Working out where a species is likely to become a problem, based on where it is a problem now.
- To try and help us work out where to put resources in terms of which species we should be prioritising for future or potential management.
- Provider of aquatic invasive species data via publishing them in on-line journal Aquatic Invasions <http://www.aquaticinvasions.ru/>
- Teaching
- Use for pest risk analysis

Results of the GISIN needs assessment survey

- Provider of invasive species data through non-web applications.
- Involved in research into IAS (plants) and want to make my research more readily available.
- Hands-on efforts to control invasive species.
- Provide a way for field staff to collect, submit, and download invasive species data.
- Sharing the data and the source of data with others.
- To educate the general public and other staff members about invasive species threatening our area.
- Help disseminate info via my organization.
- Encouraging others to collect invasive species data.
- As a basis for further research.
- Providing financial support to eradicate, control and prevent invasive species.
- Site management and inventory level.
- Discussing monitoring protocol and risk assessment development.
- Exploring how we can share data on invasive plant species.
- Broker of information for other end users.
- As a collaborator.

What is the general type of data you wish to have?

- Metadata information may be important as well as primary contact information (data manager)
- References to the published literature on each species
- Many profiles/species pages exist already. Spatial data is hardest to get and the most difficult to upkeep.
- Quarantine actions/regulations in other countries for a pest species or group
- Particulars on trapping (sentinal) types, placement, lures, effective detection
- Links to any existing risk assessments that have been done (either on related pathways or for intentional introductions)
- Information on other programs tackling invasive species to allow for an exchange in techniques and experiences.
- Control/management history of invasive species sites, areas surveyed (or not surveyed)
- GIS datasets
- information how to manage/control/eradicate IAS
- Information on damage, preferably quantitative.
- Plus policy (use of bounties, native animal pests, role of commercial harvest) and social information - e.g. Codes of Practice for management (Animal Welfare considerations)
- impact of species if known
- Potential world and current distribution
- Habitat profiles for prediction modeling
- How likely is it that a species currently in low densities / not widely distributed, may become a significant problem
- For species that are assessed as being potential threats, we would be interested in main invasion pathways in other areas

What is the level of web services your organization provides and/or uses?

- Currently in the development phase of the project and have not moved to web distribution.
- Provide information on threats, biology, pathways and control for target species.
- Obtains and provides policy-relevant information to promote screening of live animal imports
- We are in the process of creating web services.

Results of the GISIN needs assessment survey

- Internet Information Server, Oracle webservices, Microsoft Access
- We would probably not maintain our own service but collaborate with service providers
- Use cold fusion for web site. Contributors can add own data and references monitored through a web manager.

Which types of data are you interested in for spatial/temporal information?

- Country presence/absence only with dates
- Availability of herbarium records or information about source data (eg. is the record verified, and if so how)
- Best is a grid map with a database behind
- Much of the historical information is not georeferenced, so names of places (toponyms) would be useful
- Historical information from herbarium specimens
- Location by type, for habitat and climate requirements analysis
- Species name, extent of infestation
- There is an URGENT need to know about IAS in all kinds of Protected Areas
- Species native range and origin
- Extent of threat
- Google kml files for easy graphical displays

Which types of data are you interested in for profiles/species pages?

- Differentiation of area of origin, area of secondary distribution; estimates of cost-benefit.
- Gene sequences for cryptic IAS.
- Cost of eradication/control.
- Altitudinal range.
- Economic losses due to the IS, costs of prevention/control, and socio-economic value of the IS as a useful species. Also invasiveness, impacts, reproduction, dispersal, establishment and prevention/control potential.
- Current research being done.
- Host data (for pathogens).
- Associated legislation.
- Some indication of who regulates it or who considers it invasive.
- Economic cost of prevention and control.
- Legislation (by geopolitical area)
- Quarantine, detection, and prevention.
- Public education efforts (and effectiveness thereof).
- Pathway information should include volume of trade and export/import countries.
- Case studies of successful and failed management (that can be used to learn from).
- Approved biological control agents available for release
- Cost of control per hectare (or other unit) in areas where it is being managed.
- Any costs associated with, for example, loss of production or loss of native species or human health impacts for any area affected by it where this assessment has been carried out by a local agency. (We use Weighted Average Gross Margin as a proxy, but every country is likely to have different ways of assessing things, so maybe this field could be most useful as a descriptive field using examples from anywhere an assessment like this has been undertaken.)
- Potential spread: This also might only be able to be done descriptively.

Results of the GISIN needs assessment survey

- RSS feeds etc. with links to current news, research papers, genbank sequences, etc. for a particular taxon
- Web-based links

Which types of data are you interested in for checklist information about invasive species?

- Distribution rate
- Potential animal and human health impacts
- Specific details of control attempts and results of such attempts. Also by known distribution.
- Common names
- Ensure that standardised definitions are used so that the information is consistent and transferable. Also, give a confidence rating on the quality/accuracy of the data
- Regulatory status
- Whether they were "observed" to be invasive, or "judged" to be invasive (e.g. after Pest Risk Assessment).
- Whether there is a legal or other relatively official status given to them such as "quarantine pest", "unwanted", "environmental weed" etc.
- Biological control agents evident.
- For plants, growth form is an important character that is not addressed in taxonomy
- Associated taxa (pollinators, parasites, pathogens, predators, etc.)

What would you use to filter spatial/temporal information?

- By known distribution, by invasiveness, by impacts, by reproductive, dispersal, establishment and prevention/control potential.
- Eco-functional group (e.g. phytoplankton, macrozoobenthos, fish, etc.)
- By impact; both the mechanics of impact (e.g. predation, "strangling" by vines, loss of production, outcompetes native plants, human health impacts, clogs waterways etc.) and types of native biodiversity impacted. Also, by impact type--whether biodiversity is impacted (i.e. are they a CONSERVATION concern as opposed to a health or economic concern)

What would you use for filtering checklist information about invasive species?

- Control and prevention potential

What existing protocol would you recommend?

- Really doesn't matter as long as it's reliable (both the information and the interface).
- Global Invasive Species Database.
- The easier the better.
- I have no notion of informatics but could ask our informatician to fill this questionnaire.

Which schemas/ grammars would be acceptable for us to copy or extend, to implement this protocol?

- To extend Plinian Core for the species page information.
- Interactive online maps are not very useful as a reference, downloadable datasets are needed. Citation instructions should be given on the website as well, with complete metadata.
- One customised for invasive species would probably be the best option.
- Any schema is acceptable provided it is logical, user-friendly/intuitive design, and unclunky
- See above -- just add a header line to each table and provide metadata linking any relational tables. Which DarwinCore!!!

Results of the GISIN needs assessment survey

- As many as possible, provided it is clear how terminology / concepts are used, so "apples" and "apples" are put together as much as possible.

Other than Windows, Linux, Unix, and Mac, which operating systems are acceptable for a toolkit to support?

- MS Access, MS Excel, shapefile imports.
- Write it in Java so the O/S is more or less irrelevant.
- Not sure. Our site is ColdFusion and mySQL database.
- None. You should support browsers that work on all the above.
- Should consider Windows Mobile for portable devices.

Which Internet frameworks do you feel are acceptable for a toolkit?

- Requirements should be met by any scripting language.
- asp.net
- Don't know, but make it user-friendly.
- It doesn't matter; it's up to the provider.
- Use HTTP requests. The user should be unaware of what you're using on the back-end.
- As long as users everywhere will be able to use it with their basic internet systems, it doesn't matter.

Which programming languages do you feel are acceptable for a toolkit?

- Javascript
- Don't know, but make it user-friendly.
- None. Again an interface should expect users to use HTML web forms, not program in a language. Make it simple!!!
- ASP.NET, VB.NET
- The toolkit should run on the widest variety of platforms possible, and not require the purchase of proprietary software for design or implementation.

Which toolkit options would you find desirable?

- Don't know, but make it user-friendly and easy to use.
- Whichever is most likely to bridge the digital divide.

Do you have any additional comments, suggestions, or requests?

- The protocols should be sufficiently transparent so that any development group could participate regardless of their particular software and hardware.
- The system needs to figure out ways to gather the information from scores of databases into one system, and to make it easy for non-computer folks.
- A user-friendly, easy to use interface is just as important as the content of an Internet-based database.
- Keep minimum data requirements SIMPLE and keep the user (=collaborator) part SIMPLE or I fear it won't be widely used.
- Based on these questions, it seems you are headed toward a high-tech, complicated, standardized system. Instead, please keep it simple.
- Please persevere although it is not an easy task that you have set yourself.