Herbarium Specimen Browser:
A Tool for Accessing Botanical Specimen Collections

Erich R. Schneider[1], *John J. Leggett*[1], *Richard K. Furuta*[1], *Hugh D. Wilson*[2], *Stephan L. Hatch*[3]
Texas A&M University
College Station, TX 77843, USA
E-mail: {erich,leggett,furuta}@csdl.tamu.edu, wilson@bio.tamu.edu, s-hatch@tamu.edu

[1]Center for the Study of Digital Libraries, Tel: 1-409-862-3217
[2]Department of Biology, Tel: 1-409-845-3354
[3]S.M. Tracy Herbarium, Department of Rangeland Ecology and Management, Tel: 1-409-845-4328

ABSTRACT

For several years the Texas A&M Bioinformatics Working Group has pursued the construction of a novel digital library resource, an electronic adaptation of the information in the S.M. Tracy Herbarium, a major collection of preserved plants. This paper describes a tool we have developed for panoramically surveying the contents of the collection: the Herbarium Specimen Browser. While some of the Specimen Browser's implementation details (particularly its unconventional use of a full-text retrieval system to store its database, and its specialized mapping software) are of general interest, it also exhibits properties which designers of similar digital library access systems may find worth considering: support for pattern discovery, use of regularity in hypertext link sources and destinations, and employment of Javascript as an interface simplification mechanism.

KEYWORDS: browsing, pattern discovery, mapping, full-text retrieval, WWW, botanical collections

INTRODUCTION

Since mid-1995, the Bioinformatics Working Group at Texas A&M University has been a focus for digital library developments of several kinds. Generally speaking, we have been creating WWW tools for botanists and botanically-interested nonspecialists to explore aspects of botanical data sets, mainly relating to geographic distributions of various plant groups. Also, we have been pursuing the more specific project of transferring the information contained in the contents of the S. M. Tracy Herbarium (a collection of over 200,000 preserved plants with a particular focus on the grasses of Texas) into electronic form. Out of necessity, we initially pursued these developments separately; Web tools were constructed using information gathered by external entities while, simultaneously, a system allowing the rapid input of specimen information from the herbarium was being developed.

After the special input system was completed, input of data from the herbarium began and eventually reached a point where it became feasible (indeed, imperative) to provide Web-based tools allowing group members and the world at large to access the herbarium's resources over the Internet. This paper describes the current state of our ongoing work on the product arising from the combination of our two main activities: the Herbarium Specimen Browser (http://www.csdl.tamu.edu/FLORA/tracy/hsb.html). This paper is an updated version of [8] which described an earlier state of the system.

Following this introduction, we provide some background about our working group and the botanical collections known as *herbaria*. Then, we explain the operation and implementation details of the Specimen Browser, and describe some desirable properties which it possesses and that reflect principles that other digital library designers may wish to consider. We close with speculations about future work.

ABOUT THE WORKING GROUP AND HERBARIUM COLLECTIONS

The Texas A&M Bioinformatics Working Group (http://www.csdl.tamu.edu/FLORA/tamuherb.htm) is an interdisciplinary endeavor with participants drawn from several groups on campus. Fundamentally, participants can be divided into two groups, biologists and computer scientists. Biologically-oriented group members are primarily systematists (specialists in taxonomy) and are affiliated with the Departments of Biology, Rangeland Ecology and Management, and Entomology. The computer science-oriented members are drawn from the Center for the Study of Digital Libraries and specialize in hypermedia and digital library systems.

Digital Libraries 98 Pittsburgh PA USA
Copyright ACM 1998 0-89791-965--3/98/ 6...$5.00

Our working group is fortunate in that even before our current collaboration several of the biologist participants had begun developing Web materials on their own, and were therefore proficient in the Web technologies of the time (HTML markup and the structuring of Web information spaces) as well as in the use of commercial database programs. Consequently, they have been able to maintain information structured according to biological needs and to maintain and develop much of the group's Web infrastructure, leaving the computer scientists to develop the "advanced" Web systems.

One of the group's long-term goals has been the replication of the information in the S. M. Tracy Herbarium in electronic form. The herbarium, one of 2639 in the world, is a collection of plant specimens which have been pressed, dried, and glued to cardstock sheets. Each specimen sheet has a label containing information on the collector, the location of collection, an accession number (a number uniquely identifying the specimen within the collection), and a Latin scientific name of the specimen's species, along with an indication of the taxonomist responsible for associating that name with that species. The process of assigning scientific names to plant species is one of continual revision and fraught with dispute; as a result, many specimen sheets have annotations reflecting re-identification by later investigators. (See [4] for information on "specialist" work practices of this kind.)

The specimens in herbaria are vital to the practice of systematic botany, the branch of the field dealing with taxonomy. Herbarium specimens form the foundation of plant nomenclature, in that all scientific names (and the procedures for assigning them) are ultimately linked to *type specimens* in specific herbaria. Also, herbarium collections are important in the construction of floristic manuals or *floras*. A flora is an exhaustive list, for a given region, of a given group of plants (e.g. of all grasses, or of all flowering plants), their distributions within that region, and other information about them. The information in a flora is considered more accurate and reliable when the distributions in it are documented by herbarium specimens in addition to field observations. The over one million specimens housed in herbaria in the state of Texas provide a base of hard data that can be used for these floristic summaries (and any other study dealing with Texas plants).

The Herbarium Specimen Browser's database (containing about 82,000 records at the time of this writing) is primarily drawn from two sources, the collections in the Tracy Herbarium and the Texas A&M Biology Department Herbarium. Information on specimens in the Tracy Herbarium had not been put into electronic form prior to the formation of the working group. This was done under the supervision of group members from Rangeland Ecology and Management using a system specially designed for rapid input of specimen data. (Information on this system, called *Tracy*, is available at http://www.csdl.tamu.edu/FLORA/input/inputsys.html.) Specimen data for the Biology Herbarium did exist, but needed revision to put it into the format of the newly-entered Tracy data. A much smaller amount of data was also taken from a handful of other herbaria in Texas, all entered using the special input system. (We anticipate that, in the long term, data from both privately-operated and university-affiliated herbaria throughout the state will be accessible using our system; the particular data taken reflect initial attempts by herbarium operators throughout the state to develop a common interchange format for this and similar purposes.)

At present only specimens collected within Texas are used by the Specimen Browser. For each of those, the following items have been recorded: accession number and source herbarium, collector's name, a collector-specific number for the specimen, date of collection, county (within Texas) of collection, and scientific name (along with some special codes relating that name to a generally-accepted taxonomy [5]). Future revisions to the Specimen Browser system will allow non-Texas specimens to be used as they are entered; future data-gathering passes are anticipated to input data from annotations and images of the plants themselves.

IMPLEMENTATION AND FUNCTIONALITY OF THE SPECIMEN BROWSER

A hallmark of our working group's Web tool development has been the use of the public domain information retrieval system MG [10]. MG's collection construction programs take sets of arbitrary ASCII documents, compress them, and produce indices needed for querying. MG's querying programs then allow Boolean, ranked, and specific document (i.e. by document number) queries, returning results in a variety of forms, ranging from fully uncompressed documents to lists of document numbers.

It could be said that our tools make use of MG's full-text retrieval facilities to emulate the query functions of a relational database. "Documents" are formed from a table's individual records; each field is prefixed with a unique string to form (in most cases) a "word" which the full-text retrieval system can search for. As a result, one can retrieve the "records" containing desired field values by retrieving documents containing desired "words".

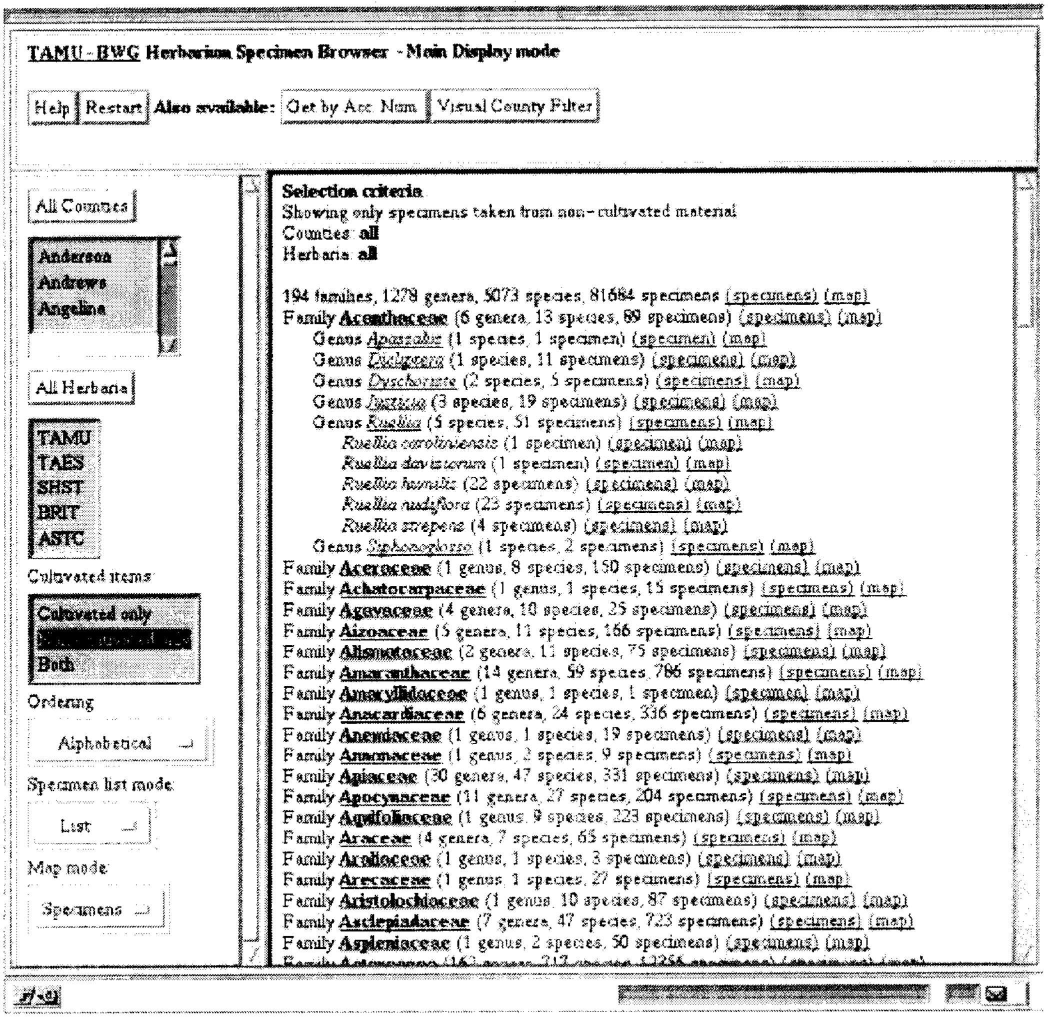
For applications such as ours where database updates are infrequent, the use of MG makes the system more convenient for users than if a standard database system were used. Since the collections are read-only, much of the overhead caused by transaction management and concurrency facilities is eliminated. Also, the retrieval system is optimized heavily with regard to query speed by moving much computation into the collection construction phase.

The Specimen Browser itself is implemented as a collection of CGI (Common Gateway Interface) programs running within a frame-based and Javascript-coordinated framework. Figure 1 shows the Specimen Browser in use. All three frames are dynamically generated. The top frame contains a title and an indication of the browser's operating mode (which is also indicated implicitly by the contents of the other frames, but the top frame provides a constant point of reference). The top frame also contains buttons which allow movement to other relevant operating modes, as well as the always-accessible "Help" and "Restart" options.

The frame on the left is the *control frame* which contains a number of controls used to change the display of information in the *main display frame* on the right. The controls fall into two classes: controls which filter the set of specimens the display is based on, and controls which generally change the display's form (for example, changing the order in which a list is sorted). The precise controls appearing in this frame depend on the Specimen Browser's operating mode; in some modes, the frame is empty.

The frame on the right is the previously-mentioned *main display frame*. It is the location of the actual displayed information derived from the specimen database. What exactly it displays is dependent, again, on the Specimen Browser's operating mode. Figure 1 shows the Specimen Browser in its initial mode, *Main Display* mode, which shows:

- The specimen selection criteria currently in force (specified by the filtering controls in the control frame)
- A line indicating the number of specimens which meet those criteria, and the numbers of species, genera, and families those specimens are members of.
- A list of all families represented by specimens meeting the current criteria, the number of genera, species, and specimens they contain, and several links.



TAMU - BWG Herbarium Specimen Browser - Main Display mode

Cultivated items:

Ordering: Alphabetical

Specimen list mode: List

Map mode: Specimens

Selection criteria
Showing only specimens taken from non-cultivated material
Counties: **all**
Herbaria: **all**

194 families, 1278 genera, 5073 species, 81684 specimens (specimens) (map)
Family **Acanthaceae** (6 genera, 13 species, 89 specimens) (specimens) (map)
Genus *Apassalus* (1 species, 1 specimen) (specimen) (map)
Genus *Dicliptera* (1 species, 11 specimens) (specimens) (map)
Genus *Dyschoriste* (2 species, 5 specimens) (specimens) (map)
Genus *Justicia* (3 species, 19 specimens) (specimens) (map)
Genus *Ruellia* (5 species, 51 specimens) (specimens) (map)
Ruellia caroliniensis (1 specimen) (specimen) (map)
Ruellia davisiorum (1 specimen) (specimen) (map)
Ruellia humilis (22 specimens) (specimens) (map)
Ruellia nudiflora (23 specimens) (specimens) (map)
Ruellia strepens (4 specimens) (specimens) (map)
Genus *Siphonoglossa* (1 species, 2 specimens) (specimens) (map)
Family **Aceraceae** (1 genus, 8 species, 150 specimens) (specimens) (map)
Family **Achatocarpaceae** (1 genus, 1 species, 15 specimens) (specimens) (map)
Family **Agavaceae** (4 genera, 10 species, 25 specimens) (specimens) (map)
Family **Aizoaceae** (5 genera, 11 species, 166 specimens) (specimens) (map)
Family **Alismataceae** (2 genera, 11 species, 75 specimens) (specimens) (map)
Family **Amaranthaceae** (14 genera, 59 species, 786 specimens) (specimens) (map)
Family **Amaryllidaceae** (1 genus, 1 species, 1 specimen) (specimen) (map)
Family **Anacardiaceae** (6 genera, 24 species, 336 specimens) (specimens) (map)
Family **Anemiaceae** (1 genus, 1 species, 19 specimens) (specimens) (map)
Family **Annonaceae** (1 genus, 2 species, 9 specimens) (specimens) (map)
Family **Apiaceae** (30 genera, 47 species, 331 specimens) (specimens) (map)
Family **Apocynaceae** (11 genera, 27 species, 204 specimens) (specimens) (map)
Family **Aquifoliaceae** (1 genus, 9 species, 223 specimens) (specimens) (map)
Family **Araceae** (4 genera, 7 species, 65 specimens) (specimens) (map)
Family **Araliaceae** (1 genus, 1 species, 3 specimens) (specimens) (map)
Family **Arecaceae** (1 genus, 1 species, 27 specimens) (specimens) (map)
Family **Aristolochiaceae** (1 genus, 10 species, 87 specimens) (specimens) (map)
Family **Asclepiadaceae** (7 genera, 47 species, 723 specimens) (specimens) (map)
Family **Aspleniaceae** (1 genus, 2 species, 50 specimens) (specimens) (map)

Figure 1: the Specimen Browser in use

The family names in this display are HTML anchor sites. Selecting one of them causes an "expansion" of the display to show a listing of the genera (represented by specimens) contained in that family; one can then select one of the genera to see a list of species in the genus. Selecting an already expanded item causes its "contraction". Figure 1 shows the results of expanding the family Acanthaceae, and within that family, the genus *Ruellia.* Selecting "Acanthaceae" again would cause all sub-items under it to disappear.

As already mentioned, the filter controls in the control frame specify a set of selection criteria which determine which specimens are used in constructing the displayed list. Currently there are three classes of criteria: county of collection, herbarium in which a specimen is located, and whether or not a specimen was taken from material under cultivation. Use of the controls to alter any of the criteria causes an immediate update of the display frame's contents. Any expanded items are still shown as expanded, but only families, genera, and species represented by specimens meeting the new criteria are displayed; also, the totals indicating numbers of genera, species, and specimens are updated to reflect the new criteria. Actions on the filter controls and updates to the display are coordinated by Javascript functions.

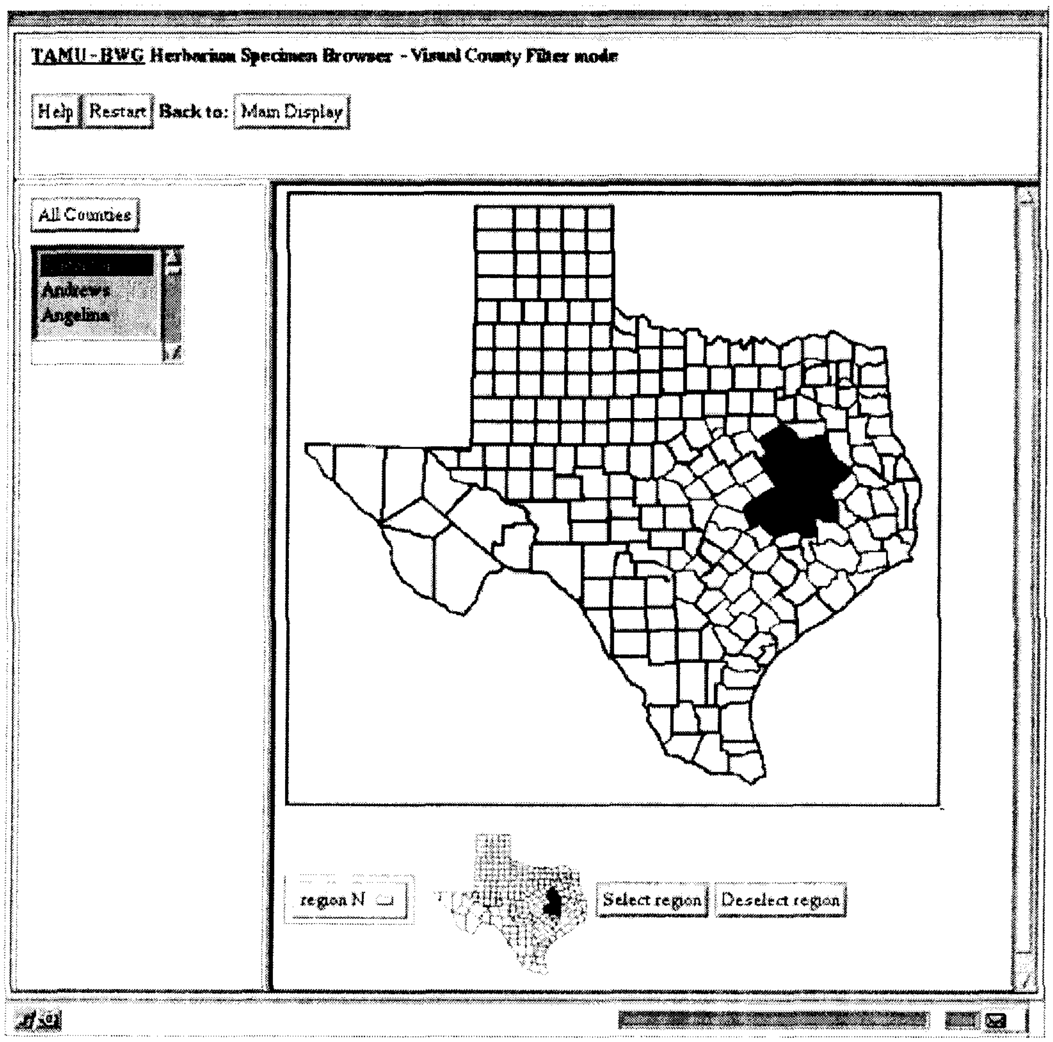
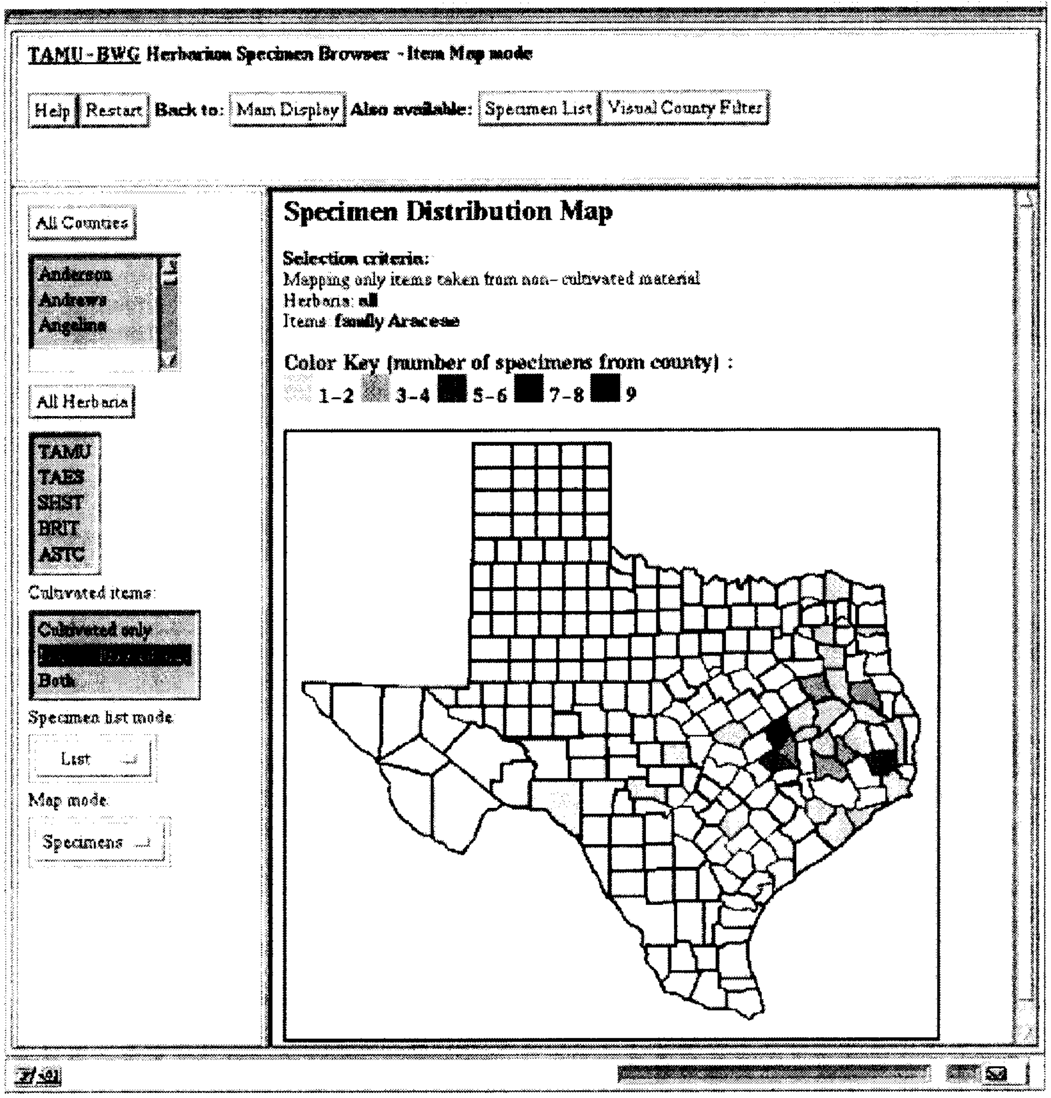


Figure 2: Visually constructing a region of inquiry

The controls in the control bar are all textual. Filtering on county of collection can also be done graphically, by switching to the *Visual County Filter* operating mode via the button in the top frame. This causes a map of Texas to appear in the display frame, with the currently-selected (via the control frame) counties colored in. (See Figure 2 for an illustration.) Clicking a county on the map will cause the corresponding entry in the list in the control frame to be selected or deselected appropriately, as well as updating the map; clicking a name on the list will cause the map to be updated in an analogous way. In this manner, one can build up a region of inquiry; when finished, returning to the operating mode one came from will present the display one left, updated appropriately with respect to the new list of selected counties. (Visual County Filter mode also provides a facility for selecting or deselecting preset blocks of about 16 counties at a time, which is useful for quickly building up a large region of inquiry.)

Simply displaying which families, genera, or species meet a given set of criteria would be very straightforward using a Boolean search. Generating running totals of specimens and other taxonomic categories is not so easy. Consider the notion of precomputing them for each possible set of counties in Texas; since there are 254 counties in the state, this would require precomputing 2^{254} totals for every item, not even considering other filtering criteria. What we do instead is sort the "documents" in the MG collection by a depth-first traversal order relative to the taxonomic tree and precompute lists indicating what categories cover what document ranges. This allows us to perform something like the SQL "select-group by" statement using the full-text retrieval system.



TAMU-BWG Herbarium Specimen Browser - Item Map mode

Back to: Main Display Also available: Specimen List Visual County Filter

Cultivated items: [illegible]

Specimen list mode: List

Map mode: Specimens

Specimen Distribution Map

Selection criteria:
Mapping only items taken from non- cultivated material
Herbaria: **all**
Items: **family Araceae**

Color Key (number of specimens from county) :
░ 1-2 ▒ 3-4 ▓ 5-6 █ 7-8 █ 9

Figure 3: Mapping specimen density

Each item in the list of families, genera, and species has a "specimens" link next to it. This is used to access detailed information on the specimens representing that item. The control bar contains a specimen list mode selector, which can be set to either "list" (the default) or "full data". Selecting a "specimens" link on the display frame switches the system to the *Specimen List* operating mode. If "list" output has been chosen, a bulleted list of specimens, listing, for each specimen, its source herbarium and accession number, scientific name, collector, and county of collection is displayed. Each item in the list is a link which puts the system into *Single Item Display Mode* and shows all available information about the specimen. If "full data" output is chosen, a list of all available data for each specimen is shown, bypassing the intermediate list. Note that the filters which were available in Main Display mode are also available in Specimen List mode (including the visual county filtering option), and the specimen list will dynamically update as filtering criteria are altered. Also, one can switch between "list" and "full data" specimen list output at will. A button in the top frame allows one to return to the main display.

Each item on the main display list also contains a "map" link. Following such a link switches the system into *Item Map* mode. A control frame selector determines whether a map of the density of specimens throughout Texas or of the density of species will be displayed (they may differ, as more than one specimen for a species may exist for a county). Figure 3 shows the result of requesting a map of the density of specimens of the family Araceae. The individual colored counties on the map may be clicked, switching the system into *Single County Specimen List* mode, listing all of the specimens found in that county as in Specimen List mode, but omitting the "by county" filter option (since only one county is being viewed). As in Specimen List mode, all filter options are available in Item Map mode; in fact, the two modes are analogous to such a degree that one can switch between them using a top frame button.

Most of the Web tools the working group has developed have included a clickable map feature. The maps are generated from a file representing the connected regions of the map in a run-length encoding (i.e. a list of *(region, number of pixels)* pairs representing the map as a left-right top-down raster scan) which are also used to easily map *(x, y)* coordinates to regions without bounding polygons and winding rules. We believe this technique has a great deal of applicability to "irregular" image maps of all kinds, and appears to be much faster (less than a second, as opposed to tens of seconds) than using a full-fledged GIS system to generate the maps. (It is necessary to state, however, that our application does not require very fine resolution or sophisticated spatial queries, both of which are handled well by GIS systems.)

To make constructing the maps more efficient, another MG collection is generated from the specimen database during the update process, this time with the records sorted by county. Document numbers are retrieved via the query mechanism in the same way as described above for the main list, but the groups formed are county clusters rather than taxonomic categories. Certain specimen records are specially tagged as representatives of their species to make species-density mapping easier, by insuring only one "representative" exists per county.

PHILOSOPHICAL POINTS BEHIND THE SPECIMEN BROWSER

The following are some general points of philosophy we feel this tool exemplifies and which other digital library designers may find useful.

Overviews and filtering

The S. M. Tracy Herbarium can, at first glance, be thought of as a library, consisting as it does of tens of thousands of artifacts, each embodying information and being of individual intrinsic interest, organized both for easy access by interested parties and for curation by collection maintainers. Unlike more conventional libraries, though, properties of the collection as a whole are as interesting if not more interesting to many patrons than properties of individual items. These collection-wide properties are not recorded outside their embodiment in the collection itself.

Much of our past and current work (especially that in mapping geographic distributions) is motivated by the desire to give biologists access to such "emergent metadata" through meaningful overviews of data sets. In this sense our work has an affinity with other digital library projects such as the Visible Human project [7]. The idea is to provide a general overview which allows the discernment of global patterns, coupled with the ability to quickly investigate details if desired.

The idea of the expanding, contracting, and filterable list (similar to Nelson's notion of *stretchtext* [6]) came about as an attempt to realize this. The initial family-level overview allows one to see how specimens are distributed through the collection by family. Interesting families can then be expanded if desired and the resultant subtotals displayed. If one wishes to restrict one's attention to a specific geographic area, one can do so while still maintaining the context of one's attention to particular taxonomic items.

Viewers looking for generalities should not be forced to rely on their own memories. This motivated the implementation of the "list" versus "full data" options for viewing specimen sets. The list option allows one to look for certain general patterns, such as preponderances of collectors, without having to page through large amounts of other data. The full data option, however, allows one to see everything that is recorded about small sets, rather than forcing one to visit each specimen in turn via the list and remember the details.

It is not only important to make patterns visible, but also to avoid the impression of false patterns. Initial experiments with our maps used red and green for the high and low ends of ranges, with a blending to indicate the middle. Unfortunately, this created a midpoint color which had greater visual salience than either endpoint, creating false impressions. The effect disappeared when we switched to a single-color scheme. (Bertin's work [1] [2] contains many useful guidelines for map designers regarding what can and cannot be signified by color, value, shape, etc., and how those variables should relate to the actual data to avoid false patterns. Similar insights can be gained from the work of Tufte [9].)

Regularity

The displays in our system are very rich in links. This gives the impression of an extensive information field which viewers can explore in an unrestrained manner.

However, we avoid disorientation by having trigger actions in a uniform matter. In addition, link sources are uniform - simple rules indicate if a link should be present and are never violated. (They are thus instantiations of what DeRose calls *annotation*, as opposed to *associative*, links [3].) This is not to say that all designers should attempt to impose uniformity on their information spaces, but it demonstrates the possibilities inherent in a simple hypertext model for constructing tools to explore detailed information spaces with regular contours.

Javascript as a simplifying mechanism

The Specimen Browser blends together static items, CGI calls, and Javascript in nontrivial ways. However, the overall effect is one of simplification. Consider that the four options of specimen list, full specimen data, species density map, and specimen density map are implemented using two controls on the control bar and two links per item. Without Javascript this would require four links per item, unnecessarily increasing screen clutter. (A general compromise might be phrased as: "use separate links for fundamentally different operations, and controls to provide generally-applicable customizations.") Javascript also allows user actions on the controls in the control frame to trigger immediate updates of the display frame; without it, some additional, superfluous user action would be required to trigger this.

Javascript is often used today to create flashy substitutes for standard lists of links or scrolling marquees in browser status bars. However, with careful use it can expand on HTML's limited link model and create effects using small sets of composable components that, before, would require large, cluttered lists of links.

Open versus closed systems

Before developing the Specimen Browser, our working group had developed several systems for viewing species distributions based on published floras for particular states of the US. These all had the general form of a shaded distribution map, subregions of which could be selected to give a species list, which list had links yielding further distribution maps. Initially, we considered constructing such a system for viewing the Tracy specimens, but issues of data imprecision and irregularity and the sheer number of species involved prompted the stretchtext-style display. Later, we realized the opportunity for providing more integrated, interrelated displays by moving to the present frame-based framework.

While the new system was much better liked by biologist members of the group than previous efforts, it became apparent that we had developed a closed system which was not easily linked into by outside frameworks. Addressing this became a priority when a system dealing with the plants of California was being developed and we wished to provide specimen-based distribution maps of the California species in a Texan context. As a result, we ended up going back and constructing a separate open system not detailed here. In many ways, the open system is less easy for users to use, but does admit outside linkage.

Our experience with the Specimen Browser thus provides a clear demonstration of a common design choice in digital library systems, between closed systems that provide tailored operations for user tasks and more open systems that admit easier "federation" with outside efforts.

FUTURE WORK

Further enhancements to the Herbarium Specimen Browser's pattern-finding facilities can be envisioned. For example, sorting schemes could be applied to specimen lists to allow examination, in that operating mode, of clusters of specimens from the same county, or collected by the same person. Filtering on the identity of the collector is also possible, but one difficulty that must be overcome is that the same collector is often indicated in many different ways on different specimen sheets. Functionality to support identification of variant names by specialists with knowledge of botanical history would be of use here.

The complement to filtered overviews is directed searching, which is currently not heavily supported; at present, one can only directly request specific specimens by herbarium and accession number (via *Get By Accession Number* mode, which was not detailed above). More support for direct searching needs to be added. This would be particularly useful for nonspecialist users should support for searches on common, rather than scientific, names be added.

One particular aspect of the information space that is amenable to processing but is currently not utilized is the time dimension. Time-series display of, say, the activities of a given collector represented in an herbarium might be interesting, but it is not clear how to do this in a straightforward yet effective way.

With the entry of all Tracy Herbarium specimens from Texas complete, the working group must now expand the system to take in other areas. (As the herbarium has many specimens from Mexico and Central America, expansion into this area is a priority.) Doing this in a way that will remain as easy for users to grasp will require enhancement to the mapping subsystem. We envision the construction of a Java applet which will operate on encoded files similar to those we use presently, but which will allow continuous, dynamic zooming. This must be done so as not to compromise rapid response, a design consideration we have considered paramount.

The entire process of assigning names to species is quite contentious. Presently the system's database contains special codes linking the names of items it contains to another taxonomy [5] which is widely accepted, though not without question. One of our planned enhancements will allow easy switching between views of the specimen set using the "native" names (those on the specimen sheets) and views using synonymous names accepted by this more generally accepted taxonomy. This is a small attempt at dealing with the much greater problem of allowing individual investigators to impose their own taxonomic "view" on the specimens in the system's database. (Systematists tend to agree in general on the scientific names of plants, but also tend to have strongly held individual views on proper naming within their area of specialty.)

Related to both the naming and mapping issues is the issue of efficient "clipping" at various taxonomic levels for maps. Presently, maps can be drawn which show the densities of specimens and species across the state. Extending this to allow mapping of genus and family densities should be easily done. It remains to be seen how easily this can be achieved if alternate taxonomies can be imposed as described above. However, in such a case this may in fact be necessary, as the reassignment of names can cause a loss of information (such as when one species is split into two - how should specimens of the "split" species then be treated?).

We also intend to integrate, hopefully in a seamless way, email links to database maintainers (possibly with responsibilities divided on taxonomic lines), for ease in making corrections or asking questions. (One benefit of our emphasis on overviews is that certain kinds of data entry errors are easily noticeable to specialist viewers.)

It will also be interesting to see what other kinds of searches we can perform using MG. Searching for records on multi-word fields (like collector name) is easily done. However, complex searches on date ranges, for example (such as finding all specimens collected between a pair of dates), are difficult to perform efficiently using our current date representations; in the future we will be investigating alternate representations more suited to the searches a full-text retrieval system can perform.

REFERENCES

1. Bertin, J. Graphics and Graphic Information Processing. Walter de Gruyter, Berlin, 1981.

2. Bertin, J. Semiology of Graphics. University of Wisconsin Press, Madison, 1983.

3. DeRose, S. J. Expanding on the Notion of Links, in Proc. Hypertext '89 Second ACM Conference on Hypertext (Pittsburgh, November 5-8, 1989), ACM Press, pp. 249-257.

4. Furuta, R. K., Marshall, C. C., Shipman, F., and Leggett, J. J. Physical Objects in the Digital Library, in Proc. DL '96 First ACM International Conference on Digital Libraries (Bethesda, March 20-23, 1996), ACM Press, pp. 109-115.

5. Kartesz, J. T. A Synonymized Checklist of the Vascular Flora of the United States, Canada, and Greenland, 2nd. Edition. Timber Press, Portland, OR, 1994.

6. Nelson, T.H. Literary Machines. Mindful Press, Sausalito, CA, 1993.

7. North, C., Shneiderman, B., and Pleasant, C. User Controlled Overviews of an Image Library: A Case Study of the Visible Human, in Proc. DL '96 First ACM International Conference on Digital Libraries (Bethesda, March 20-23, 1996), ACM Press, pp. 74-82.

8. Schneider, E. R., Leggett, J. J., Furuta, R. K., Wilson, H. D., and Hatch, S. L. WWW tools for accessing botanical collections, in Proc. WebNet 97, World Conference of the WWW, Internet, and Intranet (Toronto, November 1-5, 1997), AACE Press, pp. 449-454.

9. Tufte, E. R. The Visual Display of Quantitative Information. Graphics Press, Cheshire, CT, 1983.

10. Witten, I. H., Moffat, A., and Bell, T. Managing Gigabytes: Compressing and Indexing Documents and Images. Van Nostrand Reinhold, New York, 1994.

ACKNOWLEDGEMENTS

This research was supported in part by the Texas Advanced Research Program (Grant Number 999903-230).