For: woodruff.hydro.plan.8E9C
Printed on: Wed, Jan 22, 1992 17:49:53

Document: er_final
Last saved on: Wed, Jan 22, 1992 17:46:08

pl2ps 4.0.5 Copyright 1988 Interleaf, Inc.

DWR Browser and Text Retrieval Needs

Draft, January 1992

I. Introduction

I have been interviewing DWR staff to determine their browser–related needs. In my discussions with our engineers, I pursued several topics:

1. Attributes/functions which should be associated with various document classes (I focussed mainly on text).

2. Staff responses to the graphical browser paradigm.

3. Examples of queries they might make to a database system. It worked best if I had them explain the query in terms of the instructions they would give a human assistant performing the search for them.

II. Potential Functions to be Associated with Text Documents

- **Geographical Coordinates**
  There substantial enthusiasm about this. One issue in geo–referenced text is the distinction between the location of the document's generation versus the location of the topic of the document.

- **Time**
  This is obviously useful. Note that geographical coordinates of objects change over time, e.g. river flows move, legal boundaries are redefined, cities expand. Also, in text, the distinction must be made between time of generation of document versus time(s) referred to in document. Another concern with indexing text data from different time periods is temporal context; the meanings of terms change over time because of legislation, social influence, etc. A thesaurus used for text retrieval could take the date a document was generated into account, or alternatively, the user could be provided a facility to determine the lineage of terms.

- **Size/Length**
  Document length or size is a practical consideration which could be displayed.

- **Access History**
  Access patterns for single users or groups of users, i.e. recency of reference, frequency of access, whether a user has seen the data before.

- **Theoretical versus Practical**
  Practical articles focus more on the details of equipment, implementation. This distinction is relevant to DWR; our engineers are often looking for more practical articles because they are designing an actual project. One of the tricks they use is that certain journals are known to be more theoretical.

- **Reading Level**
  This has to do with how well–suited an article is to the reader's background. Reading level of a

document (novice to expert) can be estimated by ratio of text to equations and figures, by vocabulary level, etc.

- **Author's Attitude**
  Two propositions characterizing authors' attitude are "X exists" or "I support X". For example, "find me all articles which say global change warming is happening", or "find me all articles which say good things about DWR". Information about author's attitude and time of generation of a document could be combined to allow users to view the intellectual evolution of a given discipline.

The above functions are all continuous. It's still not clear to me how to do certain types of searches in this paradigm, e.g. specifying a "subject". One of our biologists pointed out that plants and animals are classified hierarchically rather than along a continuum. Another subject–related issue is the classification of the specificity of an article, i.e. whether it's an overview of the subject, or focuses on a specialized aspect of the subject.

III. Browser Issues

- One desirable feature is the ability to select discrete areas in the landscape with a single mouse click, e.g. allow a mode in which the user can click on CA rather than outlining it, or can issue a query like "locate and select Santa Barbara County" (here we get into the issue of how much the browser will overlap with GIS code). Our staff said that there are certain types of searches in which the data in neighboring areas would not be relevant, for example, if they are analyzing CA population, they don't want to view data from anywhere but CA. In this case, they would want to select CA directly, rather than outlining it.

- Our staff said that in many instances they would like to select a geographical area before displaying any icons.

- One of our environmental specialists pointed out that when he is looking at something from a global level, he perceives and examine different features than when he looks at it from a micro level. For example, if he were looking at a map of the world, he would want to see bioregions imposed on the landscape, but if he narrowed in on CA, he would want to see vegetative mapping instead. Is there a way to have the landscape change dynamically as the user zooms in and out?

- Data Integrity
  Data can be self–inconsistent, or self–consistent but conflicting with other data in the repository. Automated detection of this would be very useful to DWR. Are there potential methods of culling out or marking inaccurate or obsolete data? Could there be an "integrity" function based on other information such as the source of the data and measurements against other data?

- Annotation (e.g. to data sets) is useful.
  Note per our last meeting that text will be associated with all types of data.

IV. Sample queries

The goals of DWR are essentially the following: (1) predict how much water of what quality is needed where [*census data, land use analysis*] and the priorities for supplying this water [*health and safety, economics, legislation*]; (2) find the means to supply as much of this water as possible [*hydrology, water*

quality studies, meteorology, construction], taking into account impact on environmental and cultural resources [*plant and animal species, archaeological data*]; and (3) provide the public with information [*publications, movies, press releases*].

To determine the types of queries DWR makes on these topics, I did the following: noted sample queries stated by staff during interviews; requested information on literature searches done by Resources Agency staff at the California State Library; and located DWR contracts relating to studies and literature searches.

A. Interviews

- Show land ownership and land use change through animation.
- Find all articles on obstruction of stream flow.
- If we put a barrier at location X, how will it affect salinity and water elevation at location Y? [*The engineers have to make sure that enough water is released into the Delta to prevent the sea water from intruding.*]
- If we enlarge forebay X, or pump a different amount of water out of forebay X, what will happen to flows, elevations, and salinity at location Y?
- What barriers have been introduced into the Delta during the last 20 years?
- What dredging [*deepening*] of the Delta has occurred during the last 20 years?
- What are the types of pesticides used on a certain crop?
- Find all research on Twitchell Island that says Lilaeopsis masonii [*an endangered plant in the Delta*] exists there.
- Is food limiting to striped bass larvae in estuaries where striped bass are found?
- Show the distribution of kit foxes in CA.
- Are there fewer kit foxes now than last year?
- Given a map of all endangered species, show all regions where blunt–nosed leopard lizards live, then show all studies on the lizards in this area.
- What is the incidence of ponds in relationship to species X?
- Find all tech reports on archaeological findings in area X.
- Find all legislation related to archeology in this area [*archaeological findings and endangered species sightings are not good public information because sightseers could disturb the resources*].
- Find all newspaper articles that are against the State Water Project.

B. CSL

The California State Library in Sacramento performs literature searches for various agencies. At DWR's request, they are going to do an informal analysis of some of the searches relating to natural resources management which have been done in the past. As part of this analysis, they will determine the types of parameters people have used to modify subject searches.

C. Contracts

Consultants often do our bibliographic searches, as well as unique studies. These contracts are not a comprehensive representation of DWR activities because DWR tends to contract in areas in which it has less expertise. If better research tools were commercially available, it might be cost efficient to do more of these in–house.

The following are excerpts from some of the Division of Planning contracts (in which DWR contracted with private institutions or other State Agencies, or was the contractor itself). My selections are based on a quick perusal of the material from the past decade or so. For the the sake of brevity, I've only shown short passages, some of which I have paraphrased, but I'll be happy to provide more information on any subjects of interest.

i. Agriculture, Economics

- Assess the likely short–run and long–run financial impacts of a range of alternative future water supply levels and quality standards on agriculture and related industries. This includes assessing the short–run responses of water districts and agriculturalists to recent water scarcities and identifying the likely long–run responses to continued constraints. The opportunities and constraints for adjustments to alternative supply and quality levels will be analyzed and the broader economic/financial impacts of adjustments on the agricultural sector and the State economy determined.

- (1) Evaluate the effects of newly developing technology on the State's agriculture, e.g. wind and solar energy and their economic feasibility in agriculture.
  (2) Evaluate fossil energy sources and future availability and future price.
  (3) Design a model that assesses the effect of varying supply and price of one farm commodity on related commodities.

- New [*contract date: 1979*] energy technologies are evolving, some of which have large water requirements for cooling or resource processing. It is imperative for water and energy planning at both the State and national level to assess the future water demands by energy technologies and possible impediments to energy development posed by limited water supplies.

- Perform a cross section survey of CA growers that will provide a statistically valid sample of regional crop production costs. The sample frame will be stratified by scale of farm production, crop, and county. [DWR makes direct use of farm production data in various analysis, including water allocation studies, project evaluations, and drainage considerations.]

ii. Water Quality

- Estimate the hydrologic conditions of the Sacramento–San Joaquin Delta that would occur in the absence of the Central Valley Project and the State Water Project.

- Geotechnical investigation and analysis of possible seepage from the proposed Los Banos Grandes Reservoir and possible impact of such seepage on existing shallow ground water lands in nearby farm areas. [Legislation authorizing Los Banos Grandes Reservoir directed DWR to consider possible impact of this project on high ground water and drainage problem areas on nearby irrigated farm lands.]

- Chemical analysis of soil samples for pesticides, pesticide degradation products, and petroleum hydrocarbons. [It is necessary to evaluate the possibility of soil contamination by toxic organic chemicals of land acquired for the development of the Kern Water Bank, a ground water storage project.]

- Review operation history at oil fields and related activities to identify possible sources of toxic organic chemical, heavy metal, and other toxic element contamination. The possible sources shall include, but not necessarily be limited to: (1) spills; (2) waste disposal practices; (3) improperly abandoned wells; (4) mud pits associated with drilling activities; (5) leaking storage facilities and pipelines; and (6) failure of production and injection wells allowing contamination of fresh water. [Above information to be obtained by interviews and review of the pertinent records, literature, and aerial photography.]

- Identify specific sites where significant toxic organic, heavy metal, and other toxic element contamination is likely to occur, and (1) prepare map showing location of all waste disposal areas; (2) prepare map showing location of all abandoned and/or destroyed wells; and (3) prepare map showing location of all present or former mud pits associated with drilling operations.

- Review agricultural land use history and historical pesticide usage in the study area.

iii. Hydrology

- Provide computer operations to study the impact of deepening the Sacramento River Deep Water Ship Channel on State Water Project operations.

- Implement a computer model which simulates the Central Valley Project and State Water Project system operation of reservoirs, river flows, and conveyance facilities.

- Develop an artistic representation of flow patterns in the Sacramento–San Joaquin Delta. Investigate and develop methods to visually describe: (1) various hydrologic conditions in the Delta including reverse flow conditions in the lower San Joaquin River due to State Water Project operations and (2) salinity intrusion and salinity gradients in the Delta which occur with varying flow patterns including reverse flow conditions in the lower San Joaquin River.

- Determine the hydrogeologic characteristics and properties of the Indian Wells Valley ground water system.

- Estimate the annual unimpaired flows during the past 400 years of the Santa Ynez River at Bradbury Dam, and make hydrological analyses of this time series. Such an estimate of unimpaired flows is to be made by reconstruction of flows prior to 1918 from tree ring data, and unimpaired flows of the Santa Ynez River after 1918 to be provided by the State.

- Investigate the use of weather modification techniques to augment snowpack and streamflow in the Sierra Nevada. The plan will be sufficiently specific to guide an actual seeding operation in the

Feather River Basin without further elaboration. The plan will specify types of clouds to be seeded, seeding agents to be used and their rates of application, locations for ground–based generators, aircraft seeding tracks (if applicable), suspension criteria, and a proposed method of evaluation.

iv. Environmental

- Review the literature on the environmental effects of shrubland management. [There are differences of opinion within the scientific community as to the effects of vegetation management on water and sediment yield, etc. A review of the literature is needed to evaluate the extent of existing knowledge and design a research program to fill in the gaps found to be essential to valid definition of the effects of shrubland management.]

- Provide an environmental evaluation of the effects of weather modification and vegetation management in the Santa Ynez Watershed above the Cachuma Reservoir. [These contracts relied on previous studies and the readily available literature, as well as on new studies, e.g. an analysis of existing vegetation composition in the watershed, and the development of vegetation composition scenarios with and without enhanced controlled burning for vegetation management.]

- Estimate the areal extent of each of seven management categories within three administrative areas in three watersheds. The seven management areas are: geologic hazard, riparian zone, reservoir buffer, chamise climax, north–facing slope ecotype, oak savanna, and other wildlands except barren.

- Prepare a report analyzing the degree of success of an experimental transplanting program for a state protected plant.

- Document and characterize long–term trends in the abundance and composition of key biological components of SF bay during the next 10–20 years. The study will use long term monitoring of selected biological components (plankton, benthic infauna, and macroalgea) of the Bay ecosystem to identify and evaluate environmental stress.

- Conduct a literature review of past relevant fish and wildlife studies in the Cottonwood Creek study area. Perform field surveys to determine bird and mammal populations. [DWR will use the results of the fish and wildlife studies to ascertain whether a proposed project is environmentally feasible.]

- Examine ways to mitigate the impacts of the Los Banos Grandes project on (1) the San Joaquin kit fox, Swainson's hawk, bald eagle, peregrine falcon, and greater sandhill crane, and (2) State/Federal candidate mammal and bird species. Also, address the relationship of the population of kit fox at Los Banos Grandes with the kit fox population as a whole.

v. Cultural

- Survey the literature and other information sources with respect to potential reservoir sites and submit a report setting forth the presently known aspects of cultural heritage which are significant in the reservoir areas. [Study includes researching available published and unpublished archaeological, historical, and ethnographic references dealing with the proposed reservoirs outlined above and the areas immediately surrounding them. Reports dealing with similar reservoir proj-

ects will also be used for comparative data and to outline the next phase of cultural resource investigation necessary in connection with advanced project planning.]

- Conduct an archaeological survey of the Los Banos Grandes project areas. In addition to providing a complete inventory of the project areas' cultural resources and evaluating their significance, the contractor will also provide a recommendation of measures to mitigate any impacts by the project on these resources.

vi. Project

- Conduct a feasibility investigation on enlarging Shasta Lake for the purpose of increasing water and energy supplies.

- Apply value engineering in a broadly defined sense to review and evaluate project design proposals and cost estimates to determine whether it will be possible to develop a facility less expensive than the proposed Auburn Dam which will accomplish essentially the same purposes for water supply, power, flood control, and corollary recreation and instream benefits.

- The study will be a feasibility investigation of the engineering, economic, institutional, legal, and environmental considerations of using the Oroville watershed for precipitation enhancement programs. A reconnaissance–level study was judged unnecessary due to the information available from other studies along the Sierra Nevada range. Work includes review of precipitation enhancement projects conducted by PG&E, the Bureau of Reclamation, and others.

- Determine, based on geological engineering, economic, institutional, legal, and environment considerations, the feasibility of using the Chino Ground Water Basin for the underground storage of water.

- Provide a findings report which investigates the physical, economic, and environmental effects of a controllable, submerged, mechanical barrier in the vicinity of Carquinez Strait and develop operational criteria for a model of the barrier to be used in conjunction with the Corps of Engineers' Bay–Delta model. [A number of interests within the water community felt it was appropriate at this time to review past studies and identify environmental and technical concerns on this issue.]

- Provide consulting services for the development of conceptual designs and a model testing program for large water control structures to be located in the Sacramento–San Joaquin Delta. [Study includes detailed review of engineering literature to locate texts and journal articles describing similar types of structures. Computerized searches will be made at University of California libraries, through the Archives of the Water Resources Center at Berkeley, and other libraries where appropriate.]

- Determine implementation procedures, necessary improvements, project costs, revenues, funding, and a benefit/cost ratio for the wildlife management plan which is being considered as an alternative to an overland water facility for Sherman Island.
  (1) implementation and project management procedures for a wildlife management plan (e.g. Define who shall possess the legal ownership of land titles.)
  (2) description of needed and existing improvements (e.g. island levees, roads, pumps)
  (3) consumptive water use differences between marsh and riparian habitat, and historic crop pat-

terns
(4) water quality and quantity needed to maintain wildlife habitat
(5) description of water rights
(6) water and marsh management strategies
(7) description of existing wildlife populations and Department of Fish and Game future objectives
(8) waterfowl disease potential and probable control measures
(9) mosquito abatement considerations
(10) inventory of existing vegetation
(11) soils map
(12) history of past land use
(13) topography (estimate of expected survey costs)
(14) potential archaeological considerations
(15) estimated capital costs
(16) predicted funding needs
(17) benefit/cost ratio