

UCSD SBRP Information Integration Strategy and Web Site Vision

Draft 1.0 prepared by the SBRP Web Site Development Team:

Robert Tukey, SBRP Program Director
 Keith Pezzoli, Urban Studies and Planning
 Ilya Zaslavsky, San Diego Supercomputer Center
 Richard Marciano, San Diego Supercomputer Center
 Ben Tolo, San Diego Supercomputer Center
 Jason Wiskerschen, San Diego Supercomputer Center
 Michael Bedar, Regional Workbench, SBRP

Introduction	1
Purpose	1
Overview of the key needs	2
A Multi-level and Collaborative Team Approach	2
1. INTRA-NET	2
2. REGIONAL-NET	2
3. NATIONAL-NET	2
The UCSD SBRP portal will provide:	3
Features of the Other 18 SBRP Program Websites	4
Key Users	4
UCSD SBRP researchers	4
Students and school teachers (EOT clients).....	4
UCSD Students.....	4
Industrial partners and community groups (EOT clients).....	4
Outside World, general public, communities and neighborhoods,	4
Mexico.....	4
Site Organization	5
Plan of action	5
Attachment 1: SBRP Program Description.....	6
Attachment 2. Levels of the UCSD SBRP Web Site.....	7
Attachment 3: Graphical Interface for Site Navigation*	8
Attachment 4: Duties And Tasks Organized by Thrust Areas.....	9
Attachment 5: List of functions for a full time Web designer	10

Introduction

The purpose of this document is to collect, analyze and define high-level needs and features of the UCSD SBRP Web Site, and the organization of data and information sharing and integration among UCSD SBRP projects. It focuses on the capabilities needed by the stakeholders and the target users, and why these needs exist. In addition, it provides a high-level architecture for the web site, including data modeling for the site’s content.

Purpose

The purpose of this Vision document is to provide a high-level view of the UCSD SBRP Web Site by answering the following fundamental questions:

- What problems/needs does the system address?
- Who is most likely to be affected by the system?
- What entities are included in the system?

This document will provide the gauge against which all future decisions related to the implementation of this project should be evaluated. The details will be presented in the forthcoming requirements specifications. [Attachment 1](#) contains a summary of the SBRP program aims.

Overview of the key needs

At present, the UCSD SBRP site provides simple content that includes brief descriptions of SBRP projects and core activities, schedules, and links to other related sites. Navigation is organized via left “navigation” frame, which contains links to various SBRP activities (program, news, calendar, publications, outreach, contact information) In its current form, it is not useful for UCSD superfund researchers (beyond providing links and events schedule), and does not reflect the cohesiveness of separate research projects both within UCSD SBRP, and the nation-wide SBRP program. In the future, the web site will become a working test bed for cooperation between individual research projects, a portal for accessing and navigating a project database, and will serve diverse audience: UCSD SBRP researchers, SBRP researchers outside UCSD, students, EOT clients (including teachers and industrial partners), and general public. In other words, while continuing to provide standard information about projects, the site will have elements of a portal/workbench enabling discovery and queries against research data and knowledge. We anticipate that the development of such web site can go in parallel with UCSD SBRP administration’s focus on increasing interaction and interdependence between individual projects, by making them more “information-transparent” to each other, and defining interfaces for data and information exchange and navigation useful for different groups. In other words, we need to build a “working” web site.

A Multi-level and Collaborative Team Approach

To successfully accomplish this task, a team approach is required. The specific areas of expertise required include Environmental Policy and Planning, Geographic Information Systems and spatial mapping, Data Integration and Knowledge Mapping/Navigation, and Web design and graphics. The scope of activity needs to cover internal, regional and national website functions. These three levels, outlined below, need to be integrated:

1. INTRA-NET

Facilitate cross-campus communication, networking and collaboration among the SBRP’s nine basic research projects, three research support cores, administrative core, outreach core, and training core.

2. REGIONAL-NET

Enable the linkage of science to policy, planning, and education in the San Diego-Tijuana global city-region. Meet the need for a regional, binational federated GIS of interactive maps on Superfund toxics, land use, coastal water quality, and other variables.

3. NATIONAL-NET

Link UCSD’s SBRP to similar programs across the U.S. Methods are needed, based on principles of distributed intelligence and federation, to generate synergy across the nation’s diverse sites of SBRP-driven research, outreach, and training.

The SBRP Program is furthest along in developing the Outreach-driven REGIONAL-NET. Along these lines, a “Regional Workbench” program is in place. The Regional Workbench is envisioned as a regional network of knowledge-action collaboratives to integrate research, technological development and efforts to improve quality of life (see [Attachment 2](#)). At present, the Regional Workbench site provides simple content, but the future vision of the site involves much more. The notion of a workbench is of a site that is actively used for discovery within a discipline. A workbench is an interactive site that enables a user to conduct research against existing knowledge, while building upon that knowledge and publishing his/her own work. A good workbench provides the following characteristics:

- Searching of data
- Computation using internal/external resources
- Joining of complex repositories
- Uploading of completed research

Some examples of existing workbenches in other disciplines include the Environmental Workbench (NASA), Biology WB (NCSA → UCSD), Chemical Engineering WB (NCSA), Medline WB, Sociologist WB (Cornell), Archivist WB, Sociology WB, Scientist WB. The utility of the Workbench Approach is not limited to the Regional level of SBRP’s web site. It is envisioned the Workbench Approach will provide utility for all three levels: the intra-net, regional-net and national-net.

The UCSD SBRP portal will provide:

- Searching of project-related data and knowledge
- Ability to query internal research data repositories (from UCSD SBRP projects) jointly with external databases (from EPA, other SBRP sites, etc.)
- Mechanisms for uploading research data and updating project metadata, for indexing and navigation
- Different ways of navigating SBRP resources (via text links, text queries, and spatially)
- Ability to integrate process models discovered by individual projects (the RO1-Core interaction diagrams in the SBRP proposal suggests how to move in this direction).

Building such a portal will be accomplished through consistent description of individual projects, integrating project data in a common repository, and building the web site on top of such a repository. The tools and technologies to be used in building the UCSD SRB data integration infrastructure, will include:

- 1). SDSC Storage Resource Broker (SRB), a data management environment that allows remote ingestion and access to logically-organized large volumes of data and information, and navigating and querying this information with help of Metadata Catalog (MCAT)
- 2) The Topic Maps technology for knowledge-based navigation between components of the Web site and relevant databases (as a result of MOU between Mondecca, a developer of topic map software, on the one side, and SDSC and UCSD USP on the other side).
- 3) XML-based modeling and navigation of project data. Each project can be described with a few metadata tags (Dublin Core, plus a few extensions), to include:

```

<project name>
  <pi>
    <contact information>
    <years of funding>
    <location>
    <keywords>
    <sub_projects>
      <pi>
        <personnel>
        <targeted toxins>
        <targeted genetic response mechanism>
        <methods – linking them to service cores>
        <experimental objects>
        <area_focus>
        <keywords>
        <publications>
        <process models>
        <link to more info>
      ...

```

(Input from SBRP scientists will be key here)

If other NIEHS SBRP projects were to be presented in the same fashion, it would allow querying and navigating among them in a uniform way, and eventually position UCSD SBRP as an SBRP information integration portal for NIEHS.

- 4.) SDSC technologies for integration with related databases on toxics (including spatial databases), online querying of different databases and online mapping.

Features of the Other 18 SBRP Program Websites

Currently, the NIEHS SBRP web page (<http://benson.niehs.nih.gov/sbrp/>) serves as a portal to the 19 SBRP projects, providing program summary, project descriptions, publications and research briefs for each project, and links to project web sites. The available searching mechanism is quite limited. From simple queries “which other SBRP projects are focused on the same toxins”, or “which other projects have explored a similar stress response mechanism, or used the same methods”, to more complex queries such as “whether my research results can be corroborated by results obtained in other SBRP projects”. While in the first case, query answers may be obtained by querying a simple database of SBRP projects (which does not exist yet); the second case requires a more sophisticated information and knowledge modeling effort. We suggest moving in this direction, initially enabling simple searches and eventually offering more complex interaction with the data. Additional search mechanism we can provide is via online mapping of all NIEHS SBRP sites, with linking to their specialization, focus, agenda, etc.

1. Most NIEHS SBRP projects have web sites. All websites provide links to individual projects, and many provide links to related EPA/SBRP information sources, published papers, general knowledge on environmental toxicology. However, these additional (and very useful!) resources are not accessible in a way that allows querying them together with research project data.
2. Individual project web sites differ significantly in scope, from a brief description to fairly large web sites, with maps etc.
3. Several web sites have different entry points for researchers, students, and the general public (providing adapted project descriptions for the latter).
4. The web sites serve just an information purpose, and do not provide feedback mechanisms, or ways for the audience to participate and contribute.

Key Users

UCSD SBRP researchers

The Web site and data sharing infrastructure will provide a way for SBRP researchers to store and share their databases, develop consistent metadata descriptions for their projects, and interface with other projects and clients of the SBRP (EOT, industry partners, other NIEHS, etc.)

Students and school teachers (EOT clients)

The site will serve as a portal for students and teachers in developing and studying SBRP-related curricula. Students will be able to access selected research databases, develop their own experiments, and do simple modeling (with Stella – connection with Preuss, Hyam). Additionally, student and teachers will be able to assess and map spatial extents of various toxins, and relate them to population risks.

UCSD Students

Communicating SBRP results and agenda to UCSD students will be accomplished via the Regional Workbench web site, a partner activity of the Superfund web site (see separate info on RWB)

Industrial partners and community groups (EOT clients)

The site will serve as a portal for communication between EMS-aware regional partners, via a forum.

Outside World, general public, communities and neighborhoods.

The site will provide an easily navigable portal to SBRP research in UCSD and beyond.

Mexico

Part of the materials, especially covering Mexico (including mapping) should be in both English and Spanish.

Site Organization and navigation

The site will include at least the following:

- Home
- Topics (aka Knowledge Clusters)
- Search and navigation (various navigation metaphors)
- Learning how to use the web site
- Credits/Partners
- Contact information
- Internal
 - Uploading, web-based data sharing
 - Administration

Attachment 3 is a graphical illustration of one possible site navigation interface

Plan of action

- Define the SBRP information integration strategy and website vision. Identify the immediate, medium term, and long-term goals, objectives and tasks. This calls for a “Vision Statement”
- In addition to ongoing website design upgrades, immediately select two pilot projects (e.g., two labs--such as Tukey and Taylor--and their relationship with the Cores). Prototype integration processes for these pilots thereby setting the stage to bring all the ROIs and Cores into a mutually reinforcing environment for information sharing and collaboration.
- Showcase these pilots and the test bed integration strategies during the December 10-11 National SBRP meeting in Gainesville, Florida.
- For each pilot project:
 1. Link relevant data sources, tools and methods into cross-referenced and integrated systems useful for meeting aims in basic research and training. With input from the scientists build an ontology for the subject domain and evaluate the utility of topic map technology and other semantic navigational tools.
 2. Organize statistical and spatial data collections in several on-line formats, including interactive GIS mapping and 3-D visualization, so as to facilitate research, education, outreach and training. Write-up short tutorials and learning modules for the use of these information systems.
 3. Document the environmental health policy context of the basic research and toxics in question (identify pathways for linking the science to integrated risk assessment, environmental management and/or remediation decisions). This has a regional and national dimension.
 4. In line with university policy, provide state of the art website design and graphical interfaces to meet the above challenges. Analyze and evaluate the usefulness of the website to enable continual improvement.

A more detailed list of task and duties is provided in Attachment four and five.

Attachment 1: SBRP Program Description

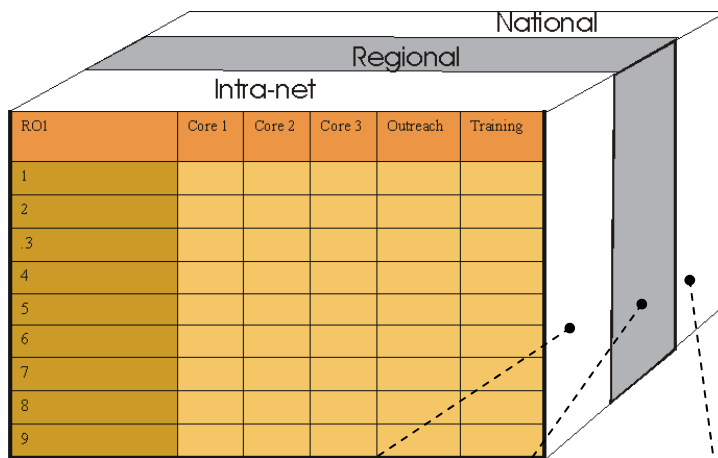
<http://superfund.ucsd.edu>

Molecular Mechanisms and Models of Exposure

The central goal of this program is to implement modern scientific approaches to identify and characterize genomic stress responses elicited by waterborne pollutants found at Superfund sites. The program is comprised of nine research projects (seven biomedical, two non-biomedical), and three research support cores (DNA micro array technology, macromolecular analysis, and mouse genetics and phenotyping). Administrative, outreach and training cores are also included.

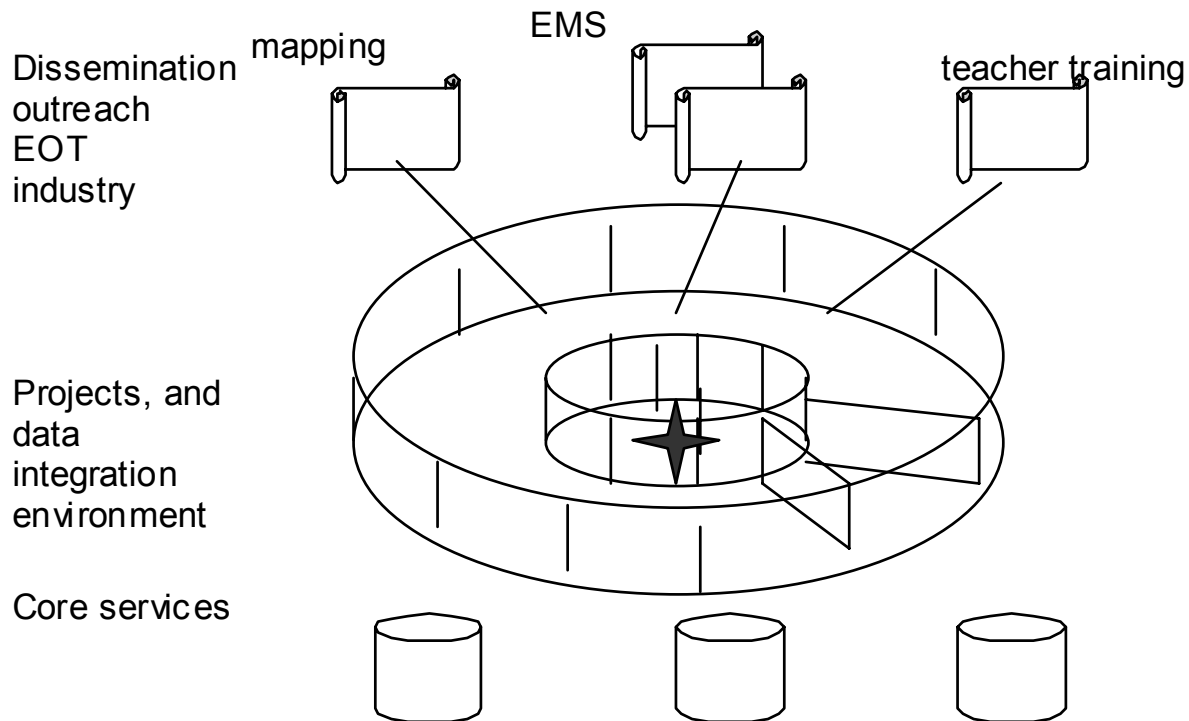
The central hypothesis of this program is that analysis of patterns of gene expression *in vivo* and *in vitro* will identify additional biomarkers of environmental injury and lead to more accurate mechanistic endpoints that can be used for risk assessment and remediation decisions. By developing innovative methodologies to identify biological responses that are caused by exposure to environmental contaminants and by defining the exposure pathway, this program will have the tools to assess mechanisms of toxicity mediated through cell signaling and gene expression. The basic research component will focus on the role of Superfund chemicals on signal transduction, the identification of target genes regulated by these chemicals and the development of models for bioremediation and detection. Three projects investigate mechanisms of oxidative stress, signal transduction, and transcriptional regulation. Molecular approaches to study expression and regulation of Phase I and II drug metabolism are the focus of two projects, and one project studies the mechanistic and molecular actions attributed to cholinesterase inhibition by organophosphates. The final biomedical project is investigating endocrine disruptors in the environment and their influence on gene expression. The two non-biomedical projects are integrated along the same themes that were developed in the biomedical projects. The first project examines how plants are protected from metal toxicity, and the signal transduction pathways activating these responses, as well as creating transgenic plants and lower organisms that can be used for highly cost-effective bioremediation of metals. The second non-biomedical project examines the transformation of metals by bacteria to understand the mechanisms involved in the natural attenuation of metal pollution and develop new bioremediation strategies.

Attachment 2. Levels of the UCSD SBRP Web Site



Intra-net	Regional-net	National-net
<p>Site designed to facilitate communication and collaboration among the SBRP’s nine projects and Cores, including Outreach and Training. Site provides user-friendly research tools, data sources, and topic maps for navigating digital libraries and links. Also serves as the SBRP clearinghouse for current news, events, seminars, and other programmatic issues.</p>	<p>Site designed to link SBRP basic science to partners in education, industry and government. The San Diego-Tijuana city-region serves as the test bed for applications. The Outreach Core website will include a binational federated GIS of interactive maps on toxics, land use, coastal water quality, and other variables for purposes of environmental management, industrial ecology and regional planning.</p>	<p>Site designed to link UCSD’s SBRP in value adding ways with other SBRPs across the nation. Site designed as a knowledge network organized around certain themes and topics judged to be priorities by the participating partners. The intent is to leverage resources in ways that generate synergy across the nation’s diverse SBRP research, education, outreach and training platforms.</p>

Attachment 3: Graphical Interface for Site Navigation*



*One possibility is to have this set up as a 3D navigation, a merge of Mondeca hyperbolic tree and a 3D solid visualization a'la Vis lab.

Attachment 4: Duties And Tasks Organized by Thrust Areas (Collaborative Team Approach)

Project Management, Policy and Planning Context (Pezzoli)
<ul style="list-style-type: none"> • Manage the overall organization of the collaborative team effort. Tasks here include documenting the process; setting up meetings, interviews, and focus groups; facilitating the work of Zaslavsky, Marciano and Tolo (e.g., identifying and collecting data sources, providing feedback/suggestions on mapping, knowledge management, and web design challenges) • Continue to serve as Outreach Core Leader, w/ co-Leader Leffert (see aims in the original proposal) • Document the environmental health policy context of the basic research and toxics in question (i.e., identify pathways for linking the science to integrated risk assessment, environmental management and/or remediation decisions, beginning with the strategically selected pilot projects) • Network with Mexico to enhance the binational dimension of our efforts. This includes (i) building, with Marciano and Zaslavsky, the first ever transborder 3-D model of the San Diego-Tijuana city-region including topography and bathymetry data (i.e., this is an essential base map for toxicology, medical geography, integrated risk assessment, etc.), and (ii) an Outreach project in Tijuana.
Research Methods and Tools 1: Geographic Information Systems and On-line mapping (Zaslavsky)
<ul style="list-style-type: none"> • Develop “spatial navigation” mechanisms for SBRP web site • Team up with Marciano and Pezzoli; develop pilot ontologies for data integration between two labs. • Adapt the online interactive mapping system for use by students and other education/outreach partners, and work with them towards its inclusion in the curricula. • Write-up short tutorials and learning modules for the use of these information systems. • Integrate environmental data from various online sources/databases (EPA, San Diego County permits, SANDAG, etc.) into online maps reflecting chemicals/pollutants of concern to UCSD SBRP. • Create interactive maps for analysis of spatial distribution of toxics, and health risks.
Research Methods and Tools 2: The Semantic Web, Knowledge-based integration (Marciano)
<p>Contribute to the development of Superfund-related environmental science and toxic studies by targeting educational and industrial concerns by:</p> <ul style="list-style-type: none"> • Build with Pezzoli, Zaslavsky, SIO and Mexican partners a transborder model of the San Diego-Tijuana region using elevation and ocean depth data. Explore the resulting outreach and educational opportunities. • Team up with Pezzoli and Zaslavsky to develop pilot ontologies for data integration between two UCSD labs. XML technologies and topic map / semantic navigational tools will be explored. Demonstrate and evaluate these tools for the Web site developed by Tolo. • Weave in the best candidate technologies developed in parallel through other SDSC projects, including NARA (National Archives & Records Administration) where knowledge-based persistent archives are researched and tested, NHPRC (National Historical Publications & Records Commission) where knowledge / topic maps are being prototyped and federal ontologies are being mapped.
Web Design, Graphics and Programming (Tolo)
<ul style="list-style-type: none"> • Provide state-of-the-art Web design, graphics, and programming necessary to meet the above challenges. • Analyze and evaluate the usefulness of the Web site to enable continual improvement. • Develop and implement interactive tutorials online to enhance the efforts and effectiveness of the Education Outreach and Training Cores.

Attachment 5: List of functions and percent time (One Full-time Web designer model)

% of time	Function/ Task No.	List the functions and tasks in descending order of importance starting with the essential functions. Number each function and write ESSENTIAL after each essential function.
60%	1.	<u>Website design and implementation (Essential)</u>
D	a.	<ul style="list-style-type: none"> Manage the development, integration and operation of the Superfund Basic Research Program's (SBRP) internal, regional and national website functions. Continually improve the website's overall structure, usefulness and integrity.
D	b.	<ul style="list-style-type: none"> Build the SBRP intra-net to facilitate cross-campus communication, networking and collaboration among the SBRP's nine basic research projects, three research support cores, administrative core, outreach core, and training core.
D	c.	<ul style="list-style-type: none"> Build the SBRP regional-net to enable the linkage of science to policy and planning in the San Diego-Tijuana global city-region. Specifically, provide web-based assistance the Outreach Core in its efforts to develop a regional, binational federated GIS of interactive maps on Superfund toxics, land use, coastal water quality, and other variables.
D	d.	<ul style="list-style-type: none"> Build the SBRP national-net to link UCSD's SBRP to similar programs across the U.S. Establish methods (based on principles of distributed intelligence and federation) to generate synergy across the nation's diverse sites of SBRP-driven research, outreach, and training.
20%	2.	<u>Information collection, publication, and visualization (Essential)</u>
D	a.	<ul style="list-style-type: none"> Collect and federate information sources into cross-referenced and integrated systems useful for meeting SBRP aims in research, outreach and training. Write-up short tutorials and learning modules for the use of these information systems.
W	b.	<ul style="list-style-type: none"> Present statistical and spatial data collections in a variety of on-line formats, including interactive mapping and 3-D visualization, to effectively communicate information to lay and technical audiences.
10%	3.	<u>Information systems research, analysis and evaluation (Essential)</u>
W	a.	<ul style="list-style-type: none"> Investigate and test topic map technology and other semantic navigational tools (e.g., inxight, mondeco, webmap) for possible use in the SBRP website.
D	b.	<ul style="list-style-type: none"> Keep abreast of advances in HTML, JavaScript, XML, web graphics, accessibility standards, and browser capabilities, and work together with SBRP stakeholders in adopting and adapting to changes in these technologies on the SBRP web site.
Q	c.	<ul style="list-style-type: none"> Analyze and evaluate the usefulness of the SBRP website's functions, content, tools, data sources and links for research, outreach and training.
10%	4.	<u>Synergistic activities and supervision of student research interns (Essential)</u>
W	a.	<ul style="list-style-type: none"> Catalyze synergy with existing web-based projects underway at UCSD including Environmental Informatics, Cal (IT)2, GEON, KINDRED, Regional Workbench, CSRC, GDC, GIS data archiving and access, and Internet2. Participate in meetings and outreach events as deemed necessary.
W	b.	<ul style="list-style-type: none"> Take advantage of the San Diego Supercomputer Center's capacity for integrating distributed persistent digital archives, hierarchical storage systems, databases, data-handling systems, and digital libraries into integrated scientific information repositories.
Q	c.	<ul style="list-style-type: none"> Create opportunities (lab rotations) for graduate and undergraduate student researchers to help build the SBRP's regional-net (Regional Workbench Program) to enable the linkage of science to policy and planning in the San Diego-Tijuana global city-region.