Confronting Standards and Nomenclature in Spatial Data Infrastructures: A Case Study of Urban Los Angeles County Geospatial Water Management Data

Miriam A. Cope¹, Stephanie S. Pincetl²

¹Wesleyan University mcope@weslyan.edu
²University of California, Los Angeles spincetl@ioes.ucla.edu

Abstract

This paper examines the problem of insufficient and often inaccurate water management boundary data in California. Due to fragmented water management in California, no central government agency is responsible for coordinating water data collection, authorship, and dissemination and maintenance. Despite statewide and county scale efforts to build spatial data infrastructures that include water data, the independent and isolated development of geospatial data sustains competing nomenclatures of water management features and poor boundary data in publicly available data sets. This paper examines these nomenclature and spatial inconsistencies and calls for assigning standardized numeric identifiers for California’s public and private water management entities. We contend that resolving the development of a universal ID system not only helps reconcile nomenclature differences propagated across data sets but also serves as a vehicle for forwarding an institutionally integrated water management spatial data infrastructure.

Keywords: urban water management, spatial data infrastructure, Los Angeles

*This work is licensed under the Creative Commons Attribution-Non commercial Works 3.0 License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/ or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

DOI: 10.2902/1725-0463.2014.09.art2
1. INTRODUCTION

This paper examines the status of spatial water management data in urban Los Angeles. Water resources management is decentralized in the state of California, with public, private, and non-profit agencies and organizations responsible for potable water management and storm water, ground water, and waste water management. Water management entities operate in a largely uncoordinated and independent manner that impedes broader understanding of water use and management in urban Los Angeles. Such uncoordinated activity has limited the development of sustainable water management plans under conditions of population growth and climate change.

Another consequence of decentralization is the uneven quality of geospatial water management data produced by organizations and made available to the public across water management entities and districts. Both public and private entities produce maps that differ greatly in sophistication and accuracy, ranging from paper maps with boundaries outlined in pen to geospatial boundary data whose reliability is uncertain. Data collection, editing, and sharing are conducted independently and inconsistently with no coordination among the water management entities themselves. We argue that spatial data for water management can be improved through the introduction of geospatial data and the mandates for a more robust water management spatial data infrastructure (SDI).

When de Montalvo (2003), citing US Spatial Data Transfer Standard, identifies five criteria used to assess spatial data quality: positional accuracy, attribute accuracy, consistency, completeness, and lineage. We find three of those criteria lacking in urban Los Angeles SDI. First, the nomenclature for attribute data is inconsistent across entities’ geospatial data sets. At the most basic level, if the same spatial feature has different names, then organizations and the public fail to share the same understanding about that water management feature. Second, agencies do not apply universal numeric identifiers that persist with data updates. This negatively affects relating data over time and subsequently the lineage of spatial data. Lastly, positional accuracy may differ between water management data sets. Although less of a problem for geo-visualization purposes, positional accuracy is what makes geospatial data unique: that features on the earth’s surface are tied to a particular coordinate or reference point as precisely as possible. Combined, these limitations affect the quality of these spatial data and in applied circumstances as in our case study, limit our understanding of where water is managed and by whom. The significance of SDI standards is summed up in Canada’s Geospatial Data Infrastructure: “Standards are necessary for facilitating robust, open transfer of spatial data packages between platforms... Standards provide key benefits such as encouraging innovation, improving efficiency, reducing transaction costs, increasing transparency and allowing international compatibility for the marketplace” (Canada’s Geospatial Data Infrastructure,
In urban Los Angeles, mandating standards through existing SDI can facilitate data sharing among decentralized actors in water management and help make transparent water management needs that must be understood to better address water use in conditions of climate change and population growth and for the state as a whole.

The fragmented and decentralized water management system in urban Los Angeles is a complex network of competing water management strategies, institutional arrangements and resources—financial, physical and human. Barriers to implementing standards and SDI development in this context may result from an array of organizational practices and beliefs such as the importance of sharing new information; allocation of financial resources necessary for updates versus other goals; and organizations’ understanding of strategic purposes and managing toward specific outcomes (When de Montalvo, 2003).

To examine and uncover ways of developing standards in urban Los Angeles, we build on Hendriks et al (2012) and Grus et al (2010) and conceptualize an SDI as an adaptive regulatory device. For Hendriks et al (2012), an SDI is an infrastructure that regulates behavior, including social, institutional, and technological resources and practices. The term adaptive accounts for the complexity of behaviors, components, and interactions found in social and natural systems. Complexity is further understood as an interactive relationship between institutions, resources, and organizational behaviors that a priori makes standardization and development and outcomes of an SDI difficult to control (Grus et al, 2010). An adaptive regulatory device framework helps capture the variability of geospatial data sets in urban Los Angeles water management and the dynamic relationship among social, technical and institutional factors that influence standardization in SDI development.

We begin our paper with a brief overview of existing SDI literature and examples. We then provide an overview of potable water types in our study area to contextualize the decentralized water management system in California. We then examine the SDIs of two online portals and three offline data sets found in urban Los Angeles. The discussion section examines differences in the context of an adaptive regulatory device framework, accounting for differences and potential for standardization. We contend that the urgency of maintaining reliable and predictable water supplies under conditions of uncertainty, e.g. climate change and population growth, compels a closer examination into the organizational, technological, and social factors that impede SDI development and standardization.

It is important to note that SDI gains have been achieved in the past decade within the state. Indeed, this paper recognizes the history and effort the California GIS
Council undertook to develop the California Geoportal (http://portal.gis.ca.gov) and the resulting administrative, political, and natural resources spatial information made public through the coordination of agencies at federal, state and local levels, as well as the private companies. However, the nature of this SDI remains “ad hoc” because contributions and updates are voluntary; the state as geoportal “host” is not responsible for the accuracy and information of the specific data sets. Thus the problem of inconsistent nomenclature and lack of standards remains.

2. SPATIAL DATA INFRASTRUCTURES

2.1. Philosophy and Examples

The guiding philosophy of a spatial data infrastructure (SDI) is that organizations, institutions and technology promote access, standards, and interoperability of data and platforms necessary for comprehensive and publically accessible geospatial data. Spatial data standards are one component of an SDI that help organize how individuals and entities coordinate and share information across government scales, between public and private sectors, and to ensure quality and consistent spatial data are available to all spatial data users (Mapping Science Committee, 1993; Williamson et al, 2003).

Hendriks et al (2012) argue that SDI definitions largely fit into two broad categories: ones that focus on SDI components (e.g. data, human resources, etc.), or ones that focus an SDI objectives (data quality, efficient, etc.). In both of these types, Hendriks et al (2012) find that technology dominates as the core of an SDI despite the institutional and social contexts in which SDIs are developed. As a result, complexity in developing, implementing, and analyzing SDIs becomes obscured. For example, the Geospatial Information Authority of Japan (GSI) consists of multiple offices with distinct SDI functions including geodetic surveying activities; general management of geospatial information and GIS policy; standardization of geographic information; international cooperation; public relations and hearings; updating the Digital Japan Basic Map, and environmental monitoring (http://www.gsi.go.jp/ENGLISH/page_e30003.html). This SDI emphasizes technical competencies over social to produce highly regulated geospatial information. Japan’s hierarchical system of governance may obscure the degree and complex social arrangements involved in maintaining its SDI infrastructure. In general, the focus on technologies per se (hardware, software data, etc.) is problematic for the critical reason that at their core, technologies are socially and culturally constructed (Pacey, 1983; Pickles, 1995).

It is necessary therefore to a priori define the term infrastructure to explicate the complexity in both defining and analyzing an SDI (Hendriks et al, 2012). An infrastructure is a system that regulates behavior. As a regulatory system, an SDI takes on the role of an intermediary as it becomes a material and discursive
apparatus that links components (technologies, human resources, skills, procedures) to specific organizational objectives (Hendriks et al, 2012). The conceptualization of an SDI as a “regulatory device,” requires that technologies, human resources, and institutions must be weighted similarly upon analysis because any one of those dimensions may affect the efficacy of an SDI.

A number of SDI examples hybridize an emphasis on components and objectives leading to definitions and implementation of SDIs that stress equally the relevance of social and institutional practices and behaviors. For example, U.S. Executive Order 12906 defines the U.S. National Spatial Data Infrastructure (NSDI) as “the technology, policies, standards, and human resources necessary to acquire, process, store, distribute, and improve utilization of geospatial data” (Executive Order 12096, 1993). In this example, an SDI is the combination of human resources, procedures, education and regulatory frameworks necessary to ensure data quality, data standards, sharing, and integration remain efficient and possible (Rajabifard et al, 2003; FGDC, 2005). Technical goals such as interoperability, access, and standards are translated as social outcomes in which national data are expected to support the needs of the federal government and interests and applications of entities throughout the public sector (Mapping Science Committee, 1993).

Canada’s Spatial Data Infrastructure is defined also as, “the data, standards, policies, technologies and partnerships that are in place to allow the sharing and visualization of information on the Internet” (http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/10783). Although goal oriented, its definition balances an emphasis on technology with the reality of institutional, standards and social linkages that inform its production. The United Nations Spatial Data Infrastructure emphasizes institutional and technical “coherence” for efficient geospatial data sharing and support of SDI development in Member Countries (UNSDI, http://www.unsdi.nl/what-we-do/unsdi-mission-statement/index.html). Its components consist of capacity building mechanisms through training and education, institution building through support of policy development and agenda building, and generating geospatial information. The technical infrastructure is defined as the components that support metadata, geospatial data, and internet connectivity. A qualitative analysis can uncover the degree to which the technical, capacity building, and governance components of the UNSDI intersect. Here, this example points to an integrated approach to SDI development.

An SDI in which social, technical and institutional components and outcomes interface may also be conceptualized as a “complex adaptive system” (Grus et al, 2010). Following the work of Barnes et al (2003), Grus et al (2010) define a complex adaptive system (CAS) as one in which “elements interact dynamically to exchange information, self-organize and create many different feedback
loops...and the systems as a whole have emergent properties that cannot be understood by reference to the component parts.” A distinguishing feature of complexity is that “the whole of the system is different from the sum of its parts” (Grus et al, 2010). In the sense that an SDI relies on data, standards, stakeholder involvement, institutional frameworks and human resources, an SDI is indeed a complex structure in service of and construction by the entities that produce it. The interaction between components and objects is seen in terms of understanding the entirety of an SDI, which is the aggregation and totality of human resources, technological capacities, regulatory frameworks, and individual skills that facilitate the development and exchange of accurate geospatial information (Williamson et al, 2003; Hendriks et al, 2012).

Institutions constitute elements and outcomes of an SDI and simultaneously inhibit effective SDI development data sharing and standards maintenance between key spatial data developers and the public. Multiple governmental agencies and departments still inherit responsibility for managing different mapping activities such as cadastral surveying, planning, geodetic control and national mapping (Williamson et al, 2003). Further, departments and organizations that have the need for similar data nevertheless face competing goals, which evince little perceived need or cooperation in data development or exchange (Pinto and Onsrud, 1995). Organizational beliefs about the importance of sharing new information, the allocation of financial resources necessary for updates, and beliefs about whether sharing data serve an organization’s strategic purposes contribute to a climate of “willingness to share” (When de Montalvo, 2003). The social pressure an organization perceives it experiences as well as beliefs about control over data also determine the final outcome in spatial data sharing arrangements and opportunities such as ones an SDI present.

We synthesize Hendriks (2012) and Grus et al (2010) and conceptualize an SDI as an adaptive regulatory device. An “adaptive regulatory device” means SDIs proceed through iterations and that SDIs evolves as technologies, organizational needs, and requirements change. We add the term “adaptive” for two reasons. First, the term helps account for the challenges posed in the social processes required to reengineer institutional arrangements necessary for SDI development (Williamson et al, 2003). Second, “adaptive” recognizes the dynamic changes and interactions that influence technical and institutional SDI development within and across agencies.

3. STUDY AREA

Our study area is urban Los Angeles County. Urban Los Angeles County is the area south of and excluding Santa Clarita, and is home to just over 10 million people (U.S. Census Bureau, 2013). The county contains 88 incorporated cities (LACDPW, 2006) and covers 4,058 square miles (U.S Census Bureau, 2013).
Precipitation in urban Los Angeles averages 15.5" annually and typically occurs in winter months, December through March (LCDPW, 2006). There are three major watersheds in our study area, the Los Angeles River Watershed, San Gabriel River Watershed and Coastal. Land use in these watersheds is very diverse. Upper portions of both the Los Angeles River and San Gabriel River watersheds are covered by Angeles National Forest and are largely undeveloped. However, the remaining portions contain dense areas of industrial, commercial, and residential development (LCDPW, 2006).

Water suppliers depend to varying levels on water from one of three imported water sources: the Los Angeles Aqueduct importing water from the Owens Valley, Colorado River, and the State Water Project (Green, 2007). In 1910, the City of Los Angeles completed the Los Angeles Aqueduct to deliver water from the Owens Valley to the population of the City. In 1928, 11 cities including Los Angeles founded the Metropolitan Water District of Southern California (MET), a contractor to the Bureau of Reclamation to receive water from the Colorado River. The MET stores and manages its annual allocation to ensure water supply through periods of drought. The State Water Project, completed in 1973, delivers water from the Sacramento-San Joaquin Delta, to entities over 444 miles south in Southern California. The MET is the largest State Water Project contractor by Acre-Feet, contracting on paper over 2 million acre feet of water from both the Bureau of Reclamation and the state Department of Water Resources, which manages the State Water Project (CDWR, http://www.water.ca.gov/state_water_project_home.cfm).

Although beyond the scope of this paper, groundwater rights play a significant role in the water supply system in urban Los Angeles. Due to depleting groundwater levels, and over pumping, public and private entities brought lawsuits that were eventually settled by the California Supreme Court, and resulted in the creation of court appointed authorities called "watermasters," who are responsible for ensuring water rights owners pump only their allocated amounts and to replace overdrafts. Urban Los Angeles has 7 adjudicated groundwater basins.

4. MANAGEMENT ENTITIES IN LOS ANGELES COUNTY

Since our SDI research examined primarily spatial information related to and often authored by potable water suppliers, we note here the institutional architecture of potable water management in our study area and provide additional detail below. In urban Los Angeles, there are four types of potable water suppliers: Special Districts, cities, public utilities, and mutual water companies. Special districts and cities are public entities; public utilities are private companies that are regulated by the Public Utilities Commission to theoretically help ensure fair pricing of a public good (water) that is delivered for profit. Mutual water companies are companies that sell water to members at cost. Mutual water companies are private non-profit
companies with membership restricted to those with water rights appurtenant to their land.

4.1. Public Entities

4.1.1. Special Districts

Following the passage of the Wright Act (1887), irrigators and domestic water users could form units of local government called “special districts” that structured the access, management, delivery, and payment of water within that district (Pincetl, 1999). Other types of special districts were created for fire, hospitals, libraries, and mosquito abatement. California state law defines a special district as “any agency of the state for the local performance of governmental or proprietary functions within limited boundaries” (Government Code §16271 [d]; quoted from Mizany and Manat, 2002).

Water districts are types of special districts that may be created under one of 38 general acts and 97 special acts authorized by the State of California. Los Angeles County has 24 total water districts (Table 1), formed through one of 6 types of general acts. As units of government, water districts have the power to charge fees for services or collect taxes to cover the costs of providing services (California Department of Water Resources, 1973; Mizany and Manatt, 2002).

4.1.2. Regulation of Water Districts

While general legislative acts allow for a vote at the local level to form a special district, final boundaries are approved of or denied by a Local Area Formation Commission (LAFCO). (Mizany and Manatt, 2002; California Department of Water Resources, 1973). The State of California passed the Knox-Nisbet Act of 1963 which authorized the establishment Local Area Formation Commissions to approve or deny proposed boundary changes for cities and special districts as well as initiate proposals to dissolve or consolidate special districts (Assembly Committee on Local Government, 2012). LAFCOs’ objectives include preventing urban sprawl, ensuring the protection of open space and agricultural lands, and ensuring that both high density and rural areas received crucial community services (CALAFCO, http://www.calafco.org/about.htm). There are 58 Local Area Formation Commissions in California, one in each county.

4.1.3. City owned public utilities (“Cities”)

City owned public utilities (Cities) manage the purchase, distribution systems, billing, and upkeep necessary to deliver water supplies to residential and commercial customers. These city-owned public water utilities may import water directly, such as the City of Los Angeles (which owns and operates the Los Angeles Aqueduct) or purchase water from a water wholesaler as do, for example, the cities of Downey and Santa Fe Springs which purchase water from the Central
Basin Municipal Water District. Some cities manage groundwater rights. For example, the City of Azusa owns water rights along the San Gabriel River, and purchases imported water from a wholesaler. However, it also pumps the majority of its water supplies from one of 11 active wells in the Main San Gabriel Basin (Azusa Light and Water, 2010). In urban Los Angeles County, 41 cities (Table 1) supply water to residential, commercial and/or industrial customers.

Table 1. Type and Number of Public Potable Water Suppliers in Urban Los Angeles County

<table>
<thead>
<tr>
<th>Public Water Supplier Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Districts</td>
<td>24</td>
</tr>
<tr>
<td>Cities</td>
<td>41</td>
</tr>
</tbody>
</table>

4.2. Private Entities

4.2.1. Utilities

Privately owned (or investor-owned) public utilities are another type of water supplier (Table 2). Private companies can be either for profit or not-for-profit. If such a company is for-profit, it is regulated by the State through the Public Utilities Commission (PUC). The PUC also regulates privately owned natural gas and electricity suppliers. There are 114 privately owned public water utilities in the state, regulated by the California Public Utilities Commission (CPUC, http://www.cpuc.ca.gov/puc/water/). The CPUC reviews the pricing structure of these for-profit water utilities every three years. This recognizes that water is a public good and charges must remain (theoretically) affordable for consumers. There are 11 water utilities in Los Angeles County, 9 in the urban Los Angeles County study area.

4.2.2. Mutual Water Companies

There are 25 mutual water companies in the County of Los Angeles. These are non-profit companies that sell water to members at cost. As not-for-profit private companies, they are not regulated by the PUC. However, like all water suppliers, they are monitored by the Department of Public Health and must report to and satisfy federal/EPA water quality requirements. Each mutual water company varies in scale and size.
Table 2. Type and number of private potable water suppliers in urban Los Angeles County.

<table>
<thead>
<tr>
<th>Legal Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilities (profit)</td>
<td>8</td>
</tr>
<tr>
<td>Mutual Water Companies (non-profit)</td>
<td>25</td>
</tr>
</tbody>
</table>

All of these legal types of water suppliers collect and manage spatial data independent of one another, with few exceptions. A standardized data collection process however would formalize a geospatial nomenclature and promote efficiency within organizations as they exchange data and manage their districts and service areas.

5. ACCURACY AND STANDARDS OF GEOSPATIAL WATER MANAGEMENT DATA

5.1. Background

The state of California initiated a spatial data infrastructure (SDI) in 2000. The development process brought together stakeholders from federal, state, local agencies, tribal groups, academia, and private sector actors. This resulted in the current California Geoportal, a geo-web that enables extensive data sharing and data management options across scales of government and between public and private entities. However, gains in geospatial data access and interoperability still have not translated to institutional requirements to ensure continued updating, standardization, and oversight of geospatial information.

In the context of water resources management, no single agency is responsible for identifying or updating spatial water use and management data. Agency staff members such as GIS managers and engineers independently post data to the California Geoportal; however there is no procedure for guaranteeing spatial or attribute data accuracy, or timely updates. Instead, state, county and local agencies primarily ensure internal processes for data access, editing and develop metadata to meet their own needs and directives independent of SDI development.

Further, water management agencies and companies we researched interrelate through complex, networked pathways based on water delivery and sales. For example, the Metropolitan Water District of Southern California is a special district with its own board of governors. However, it is constituted by city and special district members. Both cities and special districts make decisions independent of the MET, yet simultaneously depend on the MET for water supplies since it is the contractor for State Water Project and for Colorado River Water. These relationships have implications for data collection.
We faced two interrelated challenges in GIS data collection for Los Angeles County: identifying all potable water suppliers and locating geospatial boundary data. We needed to first empirically identify and verify the names and number of potable water suppliers to know who is involved in the management of water resources in our study area. Based on internet research, phone calls, and literature searches we identified approximately 100 public and private potable water suppliers in our study area. We confirmed existing suppliers through the primary sources listed in Table 3.

Table 3. Potable Water Supplier Data Sources

<table>
<thead>
<tr>
<th>Source name</th>
<th>Information Collected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metropolitan Water District of Southern California</td>
<td>Qualitative, District and City water suppliers (names and contact information)</td>
</tr>
<tr>
<td>Water Resources Collection and Archives</td>
<td>LA County Public and Private Water Agencies/Companies Database</td>
</tr>
<tr>
<td>LA County GIS Data Portal</td>
<td>LA County Water Companies (2008)</td>
</tr>
<tr>
<td>Special District and City Government Websites</td>
<td>Current customers, both retailers and direct customers.</td>
</tr>
</tbody>
</table>

Using the databases, websites, and interview data, we corrected spelling, abbreviations, private or public status and updated contact information.

Once we had an accurate list of current water suppliers, we could collect GIS boundary data to which we would join our data. We discovered that the adoption and extent of GIS varies by agency, city, and district. For example, the West Basin Municipal Water District has an extensive GIS intended for internal use, while Foothill Municipal Water does not use GIS technologies for internal use. Some of the mutual water companies we met with have no GIS boundary files, but rather rely on paper maps. The permutations of how data are constructed, used and shared within agencies, across agencies, and to the public are discussed next.

The second major step involved building a base map of potable water suppliers. We located publically accessible potable water supplier shapefiles through online portals and individual agencies. Our initial GIS development revolved around two data sets, “Private Water Districts” and “Water Purveyor Service Areas,” both available through online portals discussed below. Typical data management issues emerged: establishing correct projections and coordinate systems and resolving attribute data differences between multiple data sets. We relabeled records and merged data in instances where service areas had been dissolved or taken over by another water supplier. As we learned more about the spatial and attribute information, we found that 1) shapefiles of each type of did not consistently contain the same total number of records and 2) neither of the data sets consistently contained the same universe of types of potable water suppliers (e.g. special districts, city suppliers, etc.). Lastly, neither “Private Water Districts” nor “Water
Purveyor Service Areas" data sets had a formalized ID system. For research purposes, we could assign a numeric identifier to non-spatial data or geocode the non-spatial data then conduct a spatial join. We could also join data using entity names, however, text based identifiers (e.g. names) are prone to typographical errors and/or abbreviations as discussed above. Consider the problem of similarly named special water districts in urban Los Angeles County. These include: the San Gabriel Valley County Water District, San Gabriel Valley Municipal Water District and the Upper San Gabriel Municipal Water District. Any combination or missing letter can change the result of the query. The result is poor matching between records, incomplete records, or null returns if the text in a query does not exactly match the text in the database or record.

The most robust immediate solution is a geodatabase developed for cloud and enterprise access. We asked Los Angeles Department of Public Works (LADPW), the contact department for the water purveyor data set, about fields we thought could be universal IDs, “WTR_EOC” and “WTR_EOC_ID.” However, we learned that these were automatically generated IDs and were relics of data conversion processes from ArcInfo to ArcGIS.

The lack of a universal ID for potable water suppliers prompted our analysis of differences between water management geospatial data and information standards and call for a universal ID in our discussion. Below, we provide an overview of the accuracy and differences of water management information, spatial and non-spatial, and then situate our discussion in the context of current GIS management practices.

5.2. Online Portals

5.2.1. California Geoportal (Formerly Cal-Atlas)

The first portal we accessed for water supplier data was Cal-Atlas. Cal-Atlas evolved into the California Geoportal, as of 1 January, 2013. Our study required use of the Private Water Districts data set in which cities, special districts, and private companies (utilities and mutual water companies) were identified at sub-County scales. The private entities included in this data set have water rights that predate state or federal water contracts, hence the name of the folder and inclusion in the data set (Interview with the Chief of Land Resources, Bureau of Reclamation, July 2013).

The boundary data are cooperatively shared by Bureau of Reclamation, the Mid-Pacific GIS Service Center and DWR. In this data set (http://www.atlas.ca.gov/download.html#/casil/boundaries), metadata detail the coordinate system and positional accuracy standards. The original data were reviewed for accuracy by the Bureau of Reclamation. There is no procedure for reporting boundary updates or name changes. In conversation with the Bureau of
Reclamation GIS contact, there haven’t been any modifications to the “private water district” set since 2004. There is no mandate for agencies to update or report their data in state specified procedures. This goes hand in hand with the voluntary culture of the California Geoportal, in that agencies are encouraged to contribute data, but not necessarily update it, except at their convenience.

For the purposes of transparency and accuracy, the “Private Water Districts” folder nomenclature is problematic for the very reason no legal entity or political boundary is called a “Private Water District.” The folder name reflects terminology based on an internal agency need and not an institutional reality. However, the name is misleading for the general public who are simply looking for potable water supplier data. The data are still spatial features of entities that are interpreted and spatially examined beyond the mandate of managing water contracts. Additionally, the term “water districts” has a very specific meaning in the context of water management. Water districts are a class of locally formed governments with legislation structuring their role in water management. Grouping them with private companies makes it more challenging for the public to sort through data and discern geospatially some basic information about water management. The polygon records however, consist of cities, private companies and special districts. None of these entities have a unique ID that could help link spatial and non-spatial information over time to records that get updated.

5.2.2. Los Angeles County GIS Data Portal
The LA County GIS Data Portal contains an array of spatial data available for downloading and interactive mapping. Categories and sources are clearly marked. We downloaded the “Water Purveyor Service Area” shapefile located in the “Administrative and Political Boundary” category. The metadata state the water purveyor data were last updated in 2009, although there is no record of which polygons were edited. This data set is maintained by LA County Department of Public Works (LACDPW). A separate list (non-spatial) of water companies was uploaded in an excel file to the portal in May 2011 (http://egis3.lacounty.gov/dataportal/2011/01/27/water-purveyor-service-areas/). The county database contained addresses and contact information of both public and private potable water suppliers. We used it together with the WRCA inventory to ensure a complete list of potable water suppliers.

Unlike the Private Water Districts file from the California Geoportal, metadata were not supplied with this shapefile. This limited our understanding of field names contained in the attribute files. For example, it was unclear whether name in the attribute data referred to a public water system name or a water system owner, the latter of which could be a public or a private entity. We contacted LACDPW for clarification.
Over the course of several interviews, we learned that thirty years ago, LACDPW understood the need to have electronically available data given its engineering and planning needs. LACDPW did not have a mandate to create a GIS, but did so anyway because they continually found uses for it. At the time, the County Engineer’s office located in the Department of Public Works was small (in staff) but pursued the development of a GIS. Their first GIS file was a county parcel file created by converting CAD files into GIS.

For our data, we learned that the Land Development Division (http://dpw.lacounty.gov/ldd/) in LA County Department of Public Works was responsible for approving subdivision and development plans. Part of the Tract & Parcel map filing process for subdivision approval required the developer to identify the water system to which the development would connect. The Land Division used Thomas Guide maps to locate the proposed development area, draw boundaries around the proposed subdivision, the label the area with the name of a water purveyor. We did learn that some water suppliers may have, over time, provided their own boundary files or maps for digitization.

Over ten years ago in 2003, LADPW digitized the water purveyor boundaries in the Thomas Guides as initially sketched by the Land Development division. Lost in this institutional process was whether the name field in the “water purveyor service area” data set referred to the name of a public water system or the company that owned the system, or if even those were one and the same. Boundaries in this data set have not been updated for several years. Although Public Works continues to build, research, and share data they also don’t have the human resources necessary to regularly maintain and update all publically available data.

5.3. Offline Data (Agency Housed Data)

5.3.1. LAFCO

Certain data sets were only available upon request and not through an agency GIS data portal. We proceeded to collect potable water supplier boundary data from the Local Area Formation Commission of Los Angeles County (LAFCO), the agency that regulates city and special district boundaries. We received both water and sanitation special district boundary data from LAFCO. The types of special water districts in Los Angeles County and under direct LAFCO Los Angeles purview are: County Water Districts, Irrigation Districts, Municipal Water Districts, and Water Districts. LAFCO reported that it did not have a universal numeric ID system assigned to special districts under its purview. Instead, LAFCO relies on names of districts for database queries, or uses auto-generated values to identify records that are then edited or updated within the agency (LAFCO Interview May 30, 2013). This trend persisted at the state level, although it affects non-spatial water use data as well.
5.3.2. DWR
The California Department of Water Resources (CDWR) is the agency responsible for collecting Urban Water Management Reports from potable water suppliers every five years. We learned that DWR uses two databases to collect and manage water supplier data. The first database is DWR’s online water submission tool called the DWR Online Submittal Tool (DOST) (http://www.water.ca.gov/urbanwatermanagement/dost/). DOST hosts Urban Water Management Plan (UWMP) data that urban water suppliers are required to complete every 5 years. An “urban water supplier” refers to a public or private water supplier that provides water for municipal purposes either directly or indirectly to more than 3,000 customers, or that supplies more than 3,000 acre-feet (AF) of water annually (CDWR, http://www.water.ca.gov/urbanwatermanagement/docs/water_code-10610-10656.pdf)

Some of the mandatory reporting water data include water sources, water supplies (current and projected), water shortage contingency plans—including expenditure plans, lower income water demand (current and projected), wastewater and recycled water use and projects. A second database is hosted in Access. This is the Water Management Plan Tracking System. Through email correspondence, we learned that Access database is: “primarily used to house contact information, receipt dates, regional information, etc.” Currently, there is no way to correlate the information between these two databases (DOST and UWMP) except by manually matching record IDs in DOST to the correct water supplier record in Access. This process is repeated every 5 years when agencies submit UWMPs and as needed because the automatic numbers generated by DOST change with every update. Therefore, a single supplier or entity can have multiple numbers assigned to it.

We emailed DWR staff and conducted phone interviews to follow-up on the potential for developing standardized supplier names and unique IDs. The DWR’s mandates and mission suggest that 1) they would systematize the nomenclature and ID of water suppliers and 2) the volume of data they collect would benefit from a standardized ID system. The biggest challenge facing the DWR, however, is the question of who to invite to the table to systematize nomenclature and develop IDs. Precisely because no single agency oversees all potable water suppliers, and because organizations will continue to use the same information differently, it would be difficult to determine the agency or department best suited for unique ID development. Other institutional barriers emerged, such as who would update unique IDs and how updates would be propagated through individual agencies. On the technical side, the DWR also mentioned the problem we want to resolve, which is to link agencies and utilities to the water systems they manage. A numeric identifier for water suppliers would facilitate creating the one-to-one and one-to-many relationships that would show which water suppliers own and/or manage water distribution systems.
5.3.3. **Metropolitan Water District of Southern California (MET)**

Lastly, we entered into an agreement for the use of the MET’s member agency boundary data. Two significant issues emerged in this process relevant for future SDI participation and development. One, we received two different boundary data sets from the Engineering Department. These data sets differed in terms of attribute data and spatial accuracy, with the second set sent to us described as the more precise boundary data available. The second issue is that the MET authors its own boundary data of municipal water districts. Almost every municipal water district in urban Los Angeles is a member of the MET. However, these same municipal water districts are also Special Districts overseen by LAFCO and LAFCO also maintains spatial boundary data for these districts as well as cities, including the MET city members. Despite the overlap in data and data resources, and despite LAFCO’s regulatory authority of MET member agencies, each entity individually produces their own data sets for the same purpose. On one hand, the existence of compartmentalized geospatial data production makes sense because these entities have very different missions. On the other hand, the respective missions need not be compromised due to data sharing and an exchange of spatial information about the same geographic features. Human resources, budgets and expertise could be corralled into a joint mission of providing authoritative and accurate boundary data for official and public use.

Not only did we receive two separate data sets from the MET’s Engineering Department which differed, as we were told in terms of spatial accuracy (minor edits) and attribute data that came with each data set. In further conversation, we also discovered that the MET’s GIS Department relies on yet another spatial boundary set, reflecting differences in beliefs about accuracy between data sets and differences in data management styles within the MET. Each department might have its reasons for sharing data internally or with external users depending on perceived loss or gains of autonomy and perceived strategic outcomes (When de Montalvo, 2003).

6. **DISCUSSION**

6.1. **New Directions in SDI: CDPH Water Boundary Tool**

The California Department of Health’s (CDPH) Environmental Health and Tracking Program recently launched an interactive water boundary tool (http://cehtp.org/p/page.jsp?page_key=762) after four years in development. Through the water boundary tool, CDPH seeks: “to facilitate the creation, collection, and vetting of digital maps for every water system in California” (CDPH, http://www.ehib.org/projects/ehss01/water/ddwem_letter.pdf). CDPH communicates what it identifies as mutually beneficial goals through water system manager/engineer participation. First, updated boundaries are intended to support CDPH’s mandate, to prepare responses to emergencies, facilitate research,
increase collaboration between water systems and utilities (CDPH, http://www.ehib.org/projects/ehss01/water/ddwem_letter.pdf). CDPH partnered with UC Davis which led the inventory and collection of geospatial water management information, and linking it to the DPH Public Water System database. CDPH took the lead on developing online interface which is the interactive water boundary tool.

The Water Boundary Tool has a public interface that allows non-registered users to download and utilize the shapefiles. We downloaded the shapefiles for our study area which were the same ones we had accessed from the California Geoportal and from LA County GIS Data Portal. However, CDPH had 37 updated water service boundaries in our study area that were not reflected in the other databases. Some of these updates may cover ten years of potential changes to water system boundary data.

The Water Boundary Tool data add enormous value to existing water supplier service areas because they link the public water system ID to the shapefile. This raised the possibility that perhaps there is a unique identifier for water suppliers and the fragmented arrangement of water management contributes to this information remaining in organizational silos or hidden from other agencies, departments, or companies. However, in our conversation with UC-Davis, we confirmed that water system names (pwsname) in the attribute data are not the same as the owners or managers of the water system itself. This distinction is critical because it means that instead, the numeric IDs (see Figure 1, pwsid) relate to the infrastructure (the public water system) and not to the manager or owner of the system.

We ascertained this distinction by phone because there was an owner field in the attribute data. However owner refers to the owner of the data set and not the water system. When managers or engineers register their information, they are assigned a unique owner ID in order to trace their updates. Even if the owner and pwsid were the same, there is still the problem that the field name refers to the water system (i.e. infrastructure) without there being a separate field for system owner and/or manager. The attribute data nomenclature therefore reflects the purpose of the boundary tool, which is to update water management data.

This tool launched early in 2013. It is a significant step in SDI development, that is, adaptive regulatory behaviour. This SDI responds to an internal DPH need to assist entities and communities to better prepare for emergencies. It also seeks to adapt existing technologies and resources to meet those needs and make available geospatial information and technologies to all users. As of 2014, roughly 5-10% of suppliers had submitted data. This tool builds on existing SDIs and data sets, encourages participation of local experts through outreach and training has a dedicated support staff to further collaborate with local water managers and
engineers. It also raises the possibility that DPH could be the initiator and host for SDI-linked data.

**Figure 1. Attribute Table of DPH Water Boundary Tool Showing pwsid Field.**

![Attribute Table of DPH Water Boundary Tool Showing pwsid Field.](image)

6.2. **Addressing Current SDI Models**

Although a statewide spatial data infrastructure exists for water data, the water management entities we researched continue to maintain their own data collection processes, standards, and data sharing protocols. Agencies coordinate data sharing only through voluntary participation rather than required procedures. This leads to redundancy in spatial data production and limits the accuracy of information upon which government staff and the public rely. Even the naming conventions used to save and manage of data sets differ between agencies.

The lack of will to require normalized geospatial data standardization and sharing processes hinders a comprehensive understanding of water use and management information by both the public at large and by entities responsible for decision making in government and private organizations. In particular, through our contacts and interviews, we discovered that water management agencies do not require a standard numeric ID system to identify potable water suppliers independent of spatial and non-spatial data sets and databases. Additionally, data sharing within and between organizations also remains limited. Lastly, the nomenclature for attribute data in unique data sets was inconsistent. While not a panacea to resolving institutional problems of isolated water management, unique identifiers for water suppliers is a necessary requirement to trace water use and boundary information over time and between regulatory agencies.
We suggest requiring a FIPS code equivalent for every water management entity in the state. FIPS stands for the Federal Information Processing Standard, the standard created by the United States government in which a unique spatial identifier is created and tied to U.S. census designated geographies such as a county or place. All U.S. designated geographic areas have a unique FIPS code. The FIPS code field can then be the common field through which disparate data, residing in different platforms or formats, can be linked. Once assigned to water management entities, a FIPS code equivalent can be used to “join” or “relate” spatial and non-spatial water use data information and ideally be carried over into other databases. The fixed ID would have the property of being traceable across platforms regardless of data formats and internal database systems.

The institutional challenge would be to ensure an authoritative index is propagated throughout water management entities’ databases, data collection forms and processes. Mandating such action would be one way to move the project forward as it would override entities’ internal beliefs and strategies about data sharing and transparency and establish a common culture and expectation of data access and standards. Meanwhile, establishing consistent nomenclature could occur prior to assigning universal identifiers and assist with indexing processes. However, the numeric ID would help ensure that the quality of standards are met in the normalizing process, since again there would be an index through which instances of attribute data could be checked.

Assigning a FIPS code would contribute to the call by the Little Hoover Commission (2000), which recommended that the state “enact legislation that would make special districts more visible and accountable” (Little Hoover Commission, 2000). With regard to making financial information and activities transparent, the Little Hoover Commission further stated, “To be useful, financial information should be provided in standard, uniform and easily understood formats” (Little Hoover Commission, 2000, p. 26). Numeric identifiers can help managers and agencies trace information about water districts, from changes in boundaries, to mandated reporting information, and other data about ownership and management tied to that spatial feature. It provides, in short, a critical mechanism for transparency and accountability in all aspects of water management.

7. CONCLUSION

The emergent theme from our research suggests that the lack of structure is not due to a technological constraint. Rather, the current institutional arrangement of water management perpetuates a culture resistant to change that could otherwise increase transparency of water management data and increase efficiency in data sharing between agencies and the public. GIS managers and engineers from federal, state, and local agencies expressed a need and interest in standardized data however the scale of this task from both institutional and technical angles
seemed to overwhelm the possibility of centralizing the design and maintenance of a numeric ID system and corresponding nomenclature for water.

We suggest that resolving the nomenclature differences and establishing a universal ID system goes beyond addressing unique technical problems and serves also as a vehicle for forwarding an institutionally integrated water management spatial data infrastructure. In the most populous county in the nation, Los Angeles County, there continues to be little understanding of who controls water service areas, and within which jurisdictions. Analyzing this issue through an SDI perspective has the potential to improve interoperability, efficiency and access to water management information within and between organizations, but more importantly, reveals how decentralized water management and use of geospatial data that is necessary for addressing and analyzing water management issues within agencies and between agencies and the public. Assigning universal numeric identifier and standardization of nomenclature would serve as first steps towards unifying the isolated knowledge of water resources in California, and making more transparent the entities and issues involved in water management.

Structural changes to data management do not occur in a vacuum but hinge on the human contributions necessary to assess the value of linking information, improve knowledge transfer, and build a system that better traces changes over time. A core group of representatives, GIS managers and/or district managers, dedicated to re-engaging SDI development is critical. Forming a core group necessitates organizing the appropriate agencies and individuals as discussed by the DWR. It also means transforming the current ad hoc SDI culture from a voluntary to regulatory performance. The lack of IDs is not a technological problem but rather indicative of a lack of systematic coordination between individuals and agencies responsible for water management that has been historic in California (Pincetl, 1999, Hundley 2001, Green, 2007). The lack of a robust universal ID system shows embedded institutional beliefs that resist data sharing, the relinquishment of control over information, or the dedication of human capital necessary to provide advantages embedded in 21st century geospatial technologies and systems in the cause of transparency. The case studies of spatial data access described above demonstrate the dismal state of current geospatial boundary data for water management entities, including attribute information. The need to standardize embedded numeric IDs is a cause and effect of the state of incomplete spatial information on water suppliers. It seriously impedes accountability as it requires Herculean efforts to assemble accurate data, and undermines the ability to understand what is going on in water management locally, regionally and across the state. It makes it almost impossible to join other data to water data, such as water quality, census data, further obscuring analysis of management and socio-demographic information, or political decision making. While the human resources to accomplish this task will be important, the public benefits will be felt for decades to come if this initiative is undertaken.
ACKNOWLEDGEMENTS

Research was funded through the National Science Foundation award number 1204235. The authors wish to thank Kelly Torres for her enormously helpful conversation about the CDPH Water Boundary Tool; Madelyn Glickfeld for providing background information and thoughtful comments; and Debbie Cheng, PhD, for very helpful feedback. The authors especially thank the many city, county, LAFCO, and State of California public employees who gave their time and help in interviews throughout the GIS data collection process.

REFERENCES


