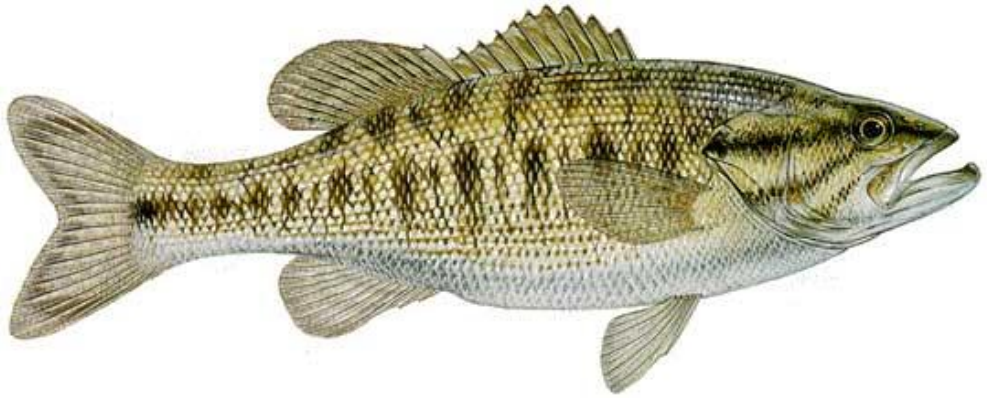
Managing Environmental Flows Information

Eric S. Hersh and David R. Maidment
The University of Texas at Austin
Center for Research in Water Resources

May 19, 2010
EWRI

Executive Summary

- Lots of data, poorly organized and minimally accessible
- A suite of tools and systems is emerging for physical and chemical water data from CUAHSI and others
- This work seeks to extend these systems by adding biological water data to the discussion in a step toward full integration



Outline

- Background
- An Environmental Flows Information System for Texas
- Hydroinformatics for Freshwater and Marine Biota
- Conclusion

Outline

- *Background*
- An Environmental Flows Information System for Texas
- Hydroinformatics for Freshwater and Marine Biota
- Conclusion

Motivation

"The State should be encouraged to mount a comprehensive review and digitization project to recover all of this data and make it available to stream ecologists and other scientists."

-Science Advisory Committee (2004)

"All available data and study reports related to the hydrologic, biologic, geomorphic, water quality, and connectivity of the study area will be assembled."

-Texas Instream Flow Studies: Technical Overview (2008)

Problem Statement

"It is clear from this BBEST's report that the collaborative process envisioned in SB3, designed to achieve a consensus, broke down."

-SAC Review Memo, 3/17/2010

- Trinity/San Jacinto/Galveston BBEST unable to reach consensus, largely attributed to a fine-scale analysis of hydrology with little or no connection to the aquatic biology of the basin and bay system

"***Cyberinfrastructure*** refers to infrastructure based upon distributed computer, information and communication technology. If ***infrastructure*** is required for an ***industrial economy***, then we could say that ***cyberinfrastructure*** is required for a ***knowledge economy***."

-2003 NSF Blue-Ribbon Advisory Panel on Cyber infrastructure

Informatics – the science of information, the practice of information processing, and the engineering of information systems

Hydroinformatics – the study of the flow of information related to the flow of water (and the entire water environment in general)

Instream Flows

Thom Hardy, Texas State University

Water left in a river system, often for managing some aspect of its condition

Instream Flows is the art and science of collecting data in systems we cannot adequately sample using methods developed by committees of technically unqualified participants for organisms we know very little about in order to form concepts about processes we do not fully understand that we represent as mathematical abstractions that we cannot precisely analyze to determine their responses to indeterminate stresses we cannot accurately predict now let alone in the future all in such a way that society at large is given no reason to suspect the extent of our ignorance.

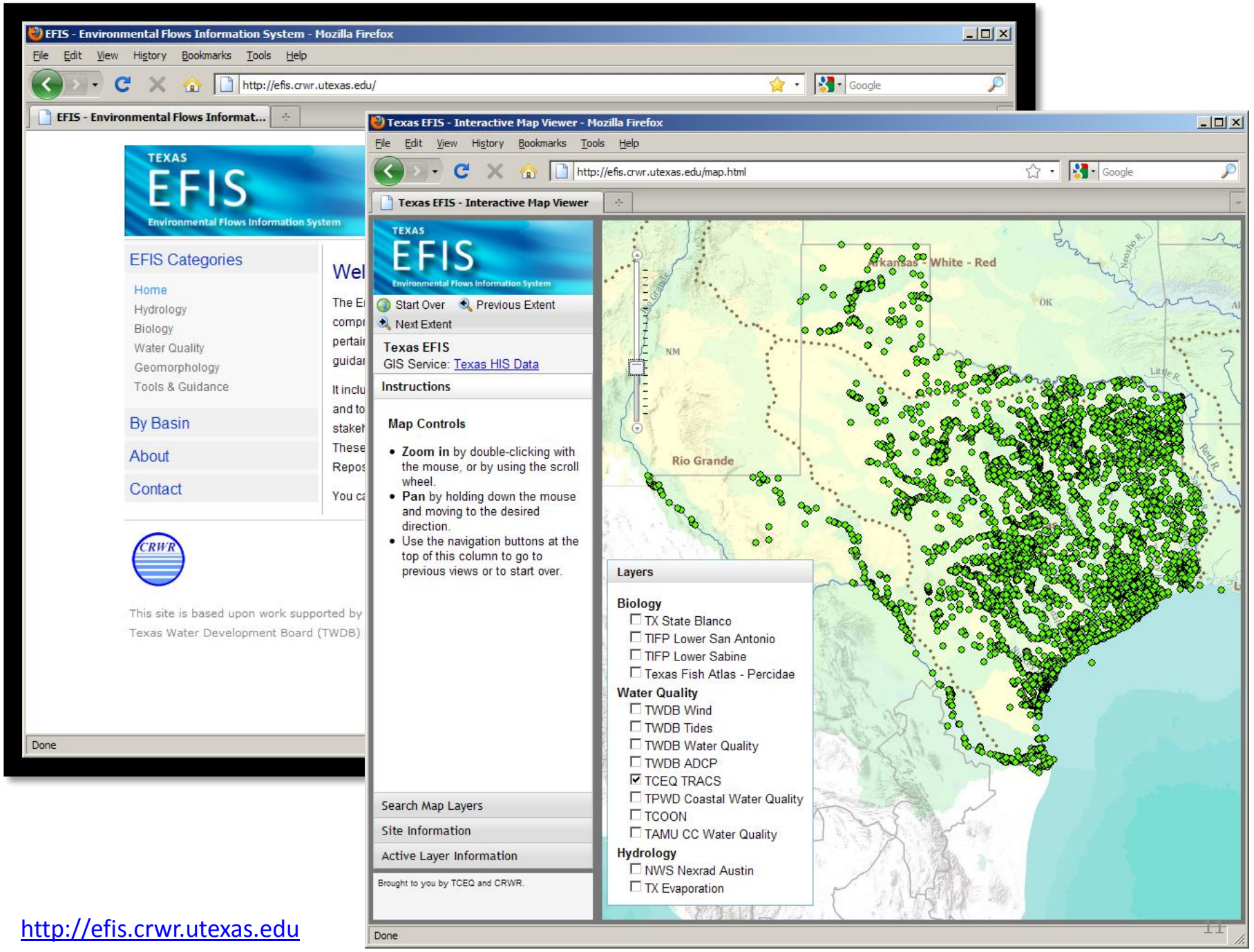
Outline

- Background
- *An Environmental Flows Information System for Texas*
- Hydroinformatics for Freshwater and Marine Biota
- Conclusion

An Environmental Flows Information System for Texas

- **Six information types:**
  - Point observations data (eg: WaterML/ODM)
  - Geographic (shapefile, feature class, KML, WFS/etc)
  - Documents (DSpace digital archive)
  - Tables (conservation status, guilds)
  - Tools (CALF, TSA, HydroExcel)
  - Links (Fishes of Texas, IHA, SAC)

- **Four access types:**
  - Web Page
  - Interactive Map Viewer
  - Digital Library
  - HydroPortal



http://efis.crwr.utexas.edu

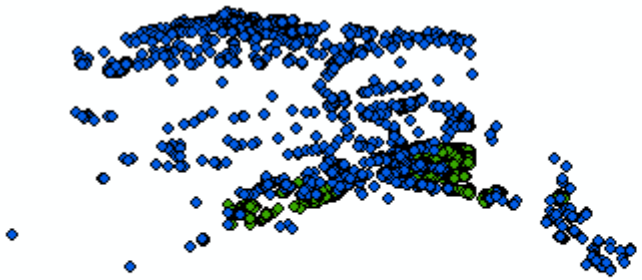
CUAHSI Hydrologic Information System

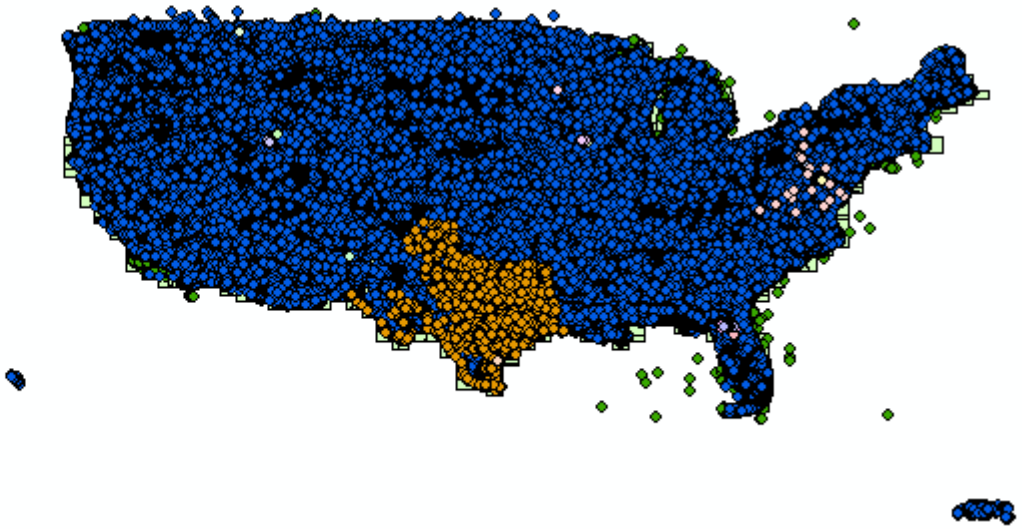
- 43 observation networks
- 15,000 variables
- 1.8 million sites
- 4.3 billion data values

Mission:

To enhance hydrologic science by facilitating user access to more and better data for testing hypotheses and analyzing hydrologic processes

http://his.cuahsi.org

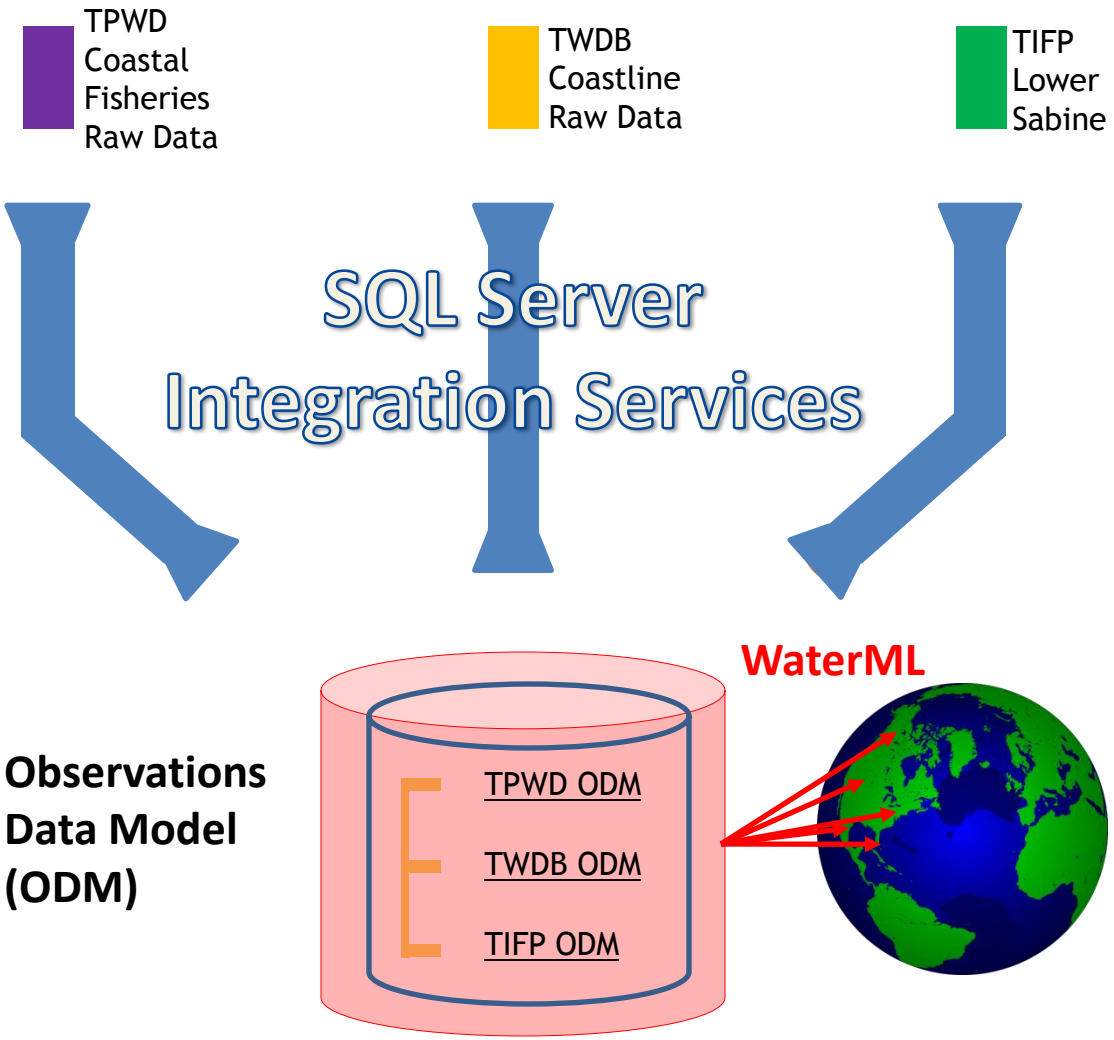




…and Texas HIS

Publishing an ODM Water Data Service

Bryan Enslein, CRWR

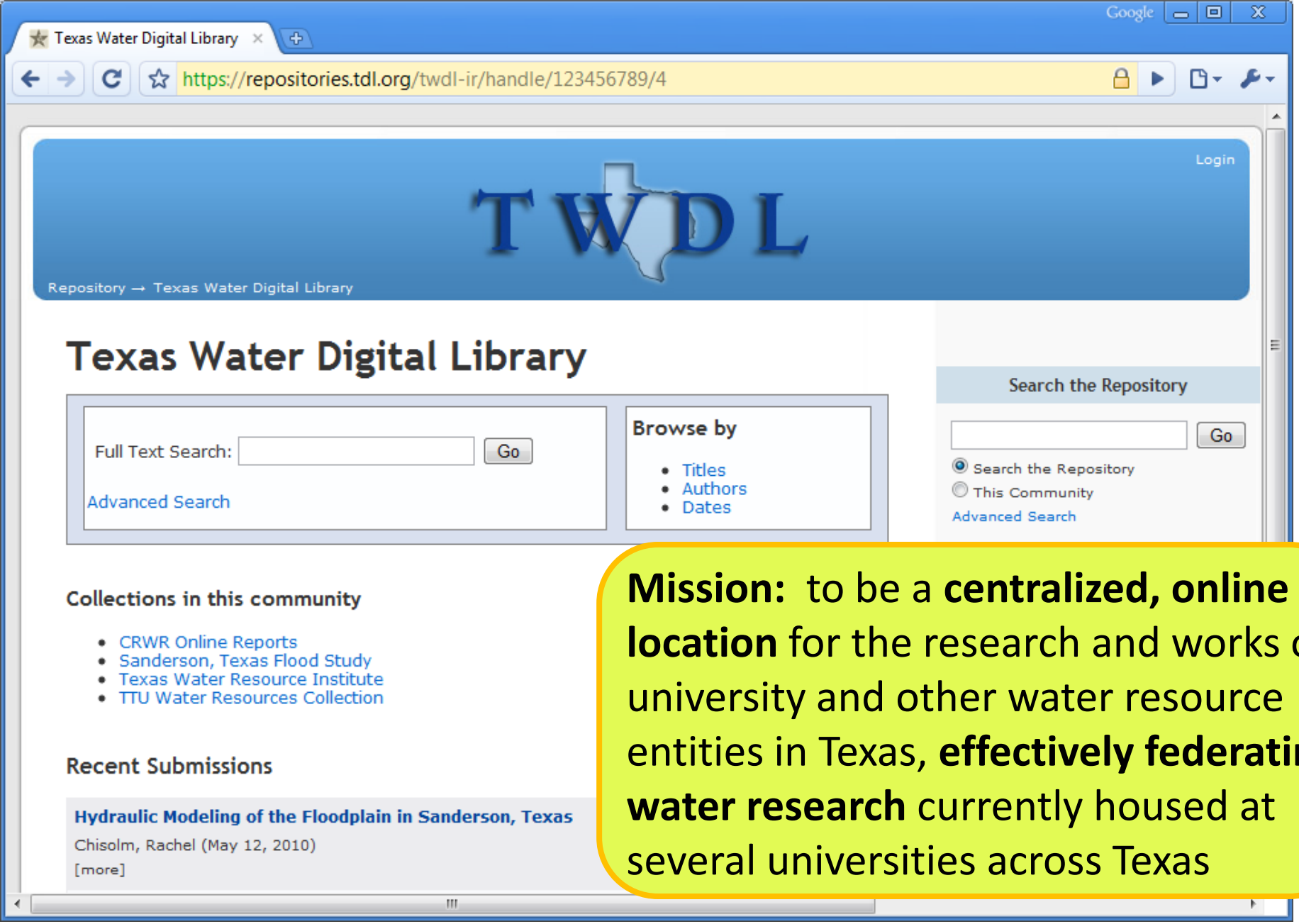


Ingest Data From Different Sources

Transform Data into Uniform Format with SSIS Scripts

Load Newly Formatted Data into ODM Tables in MS SQL/Server

Wrap ODM with WaterML Web Services for Online Publication



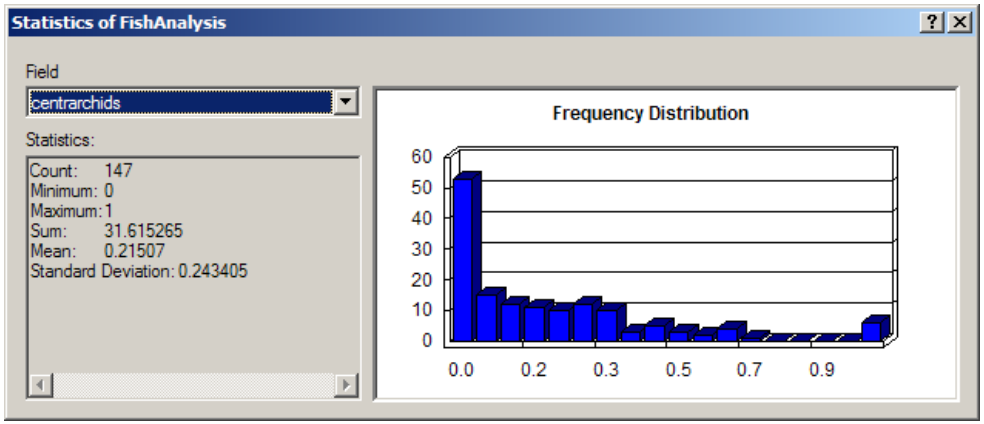
Mission: to be a **centralized, online location** for the research and works of university and other water resource entities in Texas, **effectively federating water research** currently housed at several universities across Texas

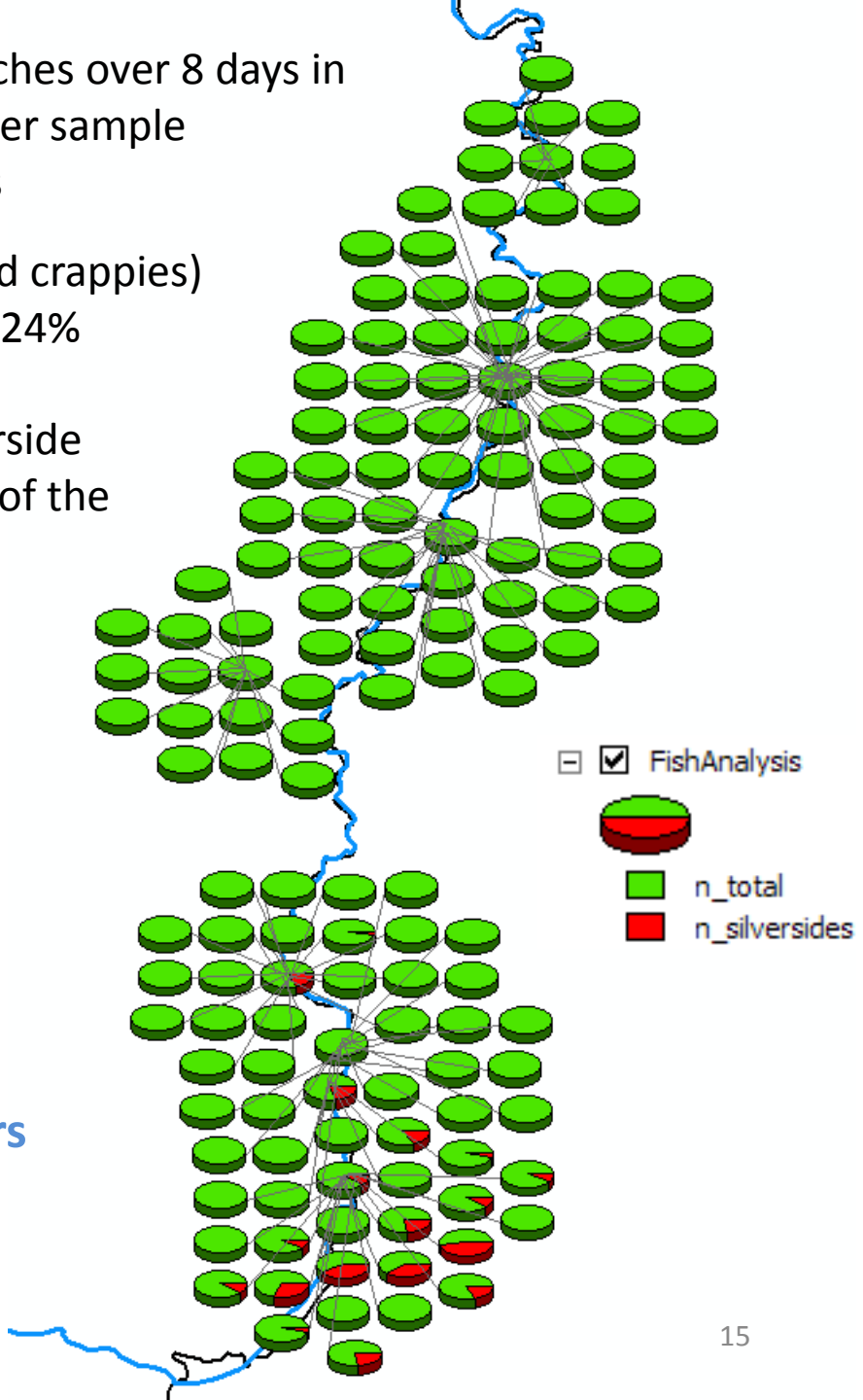
Prototype at: https://repositories.tdl.org/twdl-ir/

-Sabine River: 165 samples collected at 8 study reaches over 8 days in 2006; 147 samples yielded fish, averaging 40 fish per sample
-5,811 fish were observed, representing 58 species

-Across all sites, 889 *Centrarchids* (sunfish, bass, and crappies) were observed with a relative abundance of 22% ± 24%

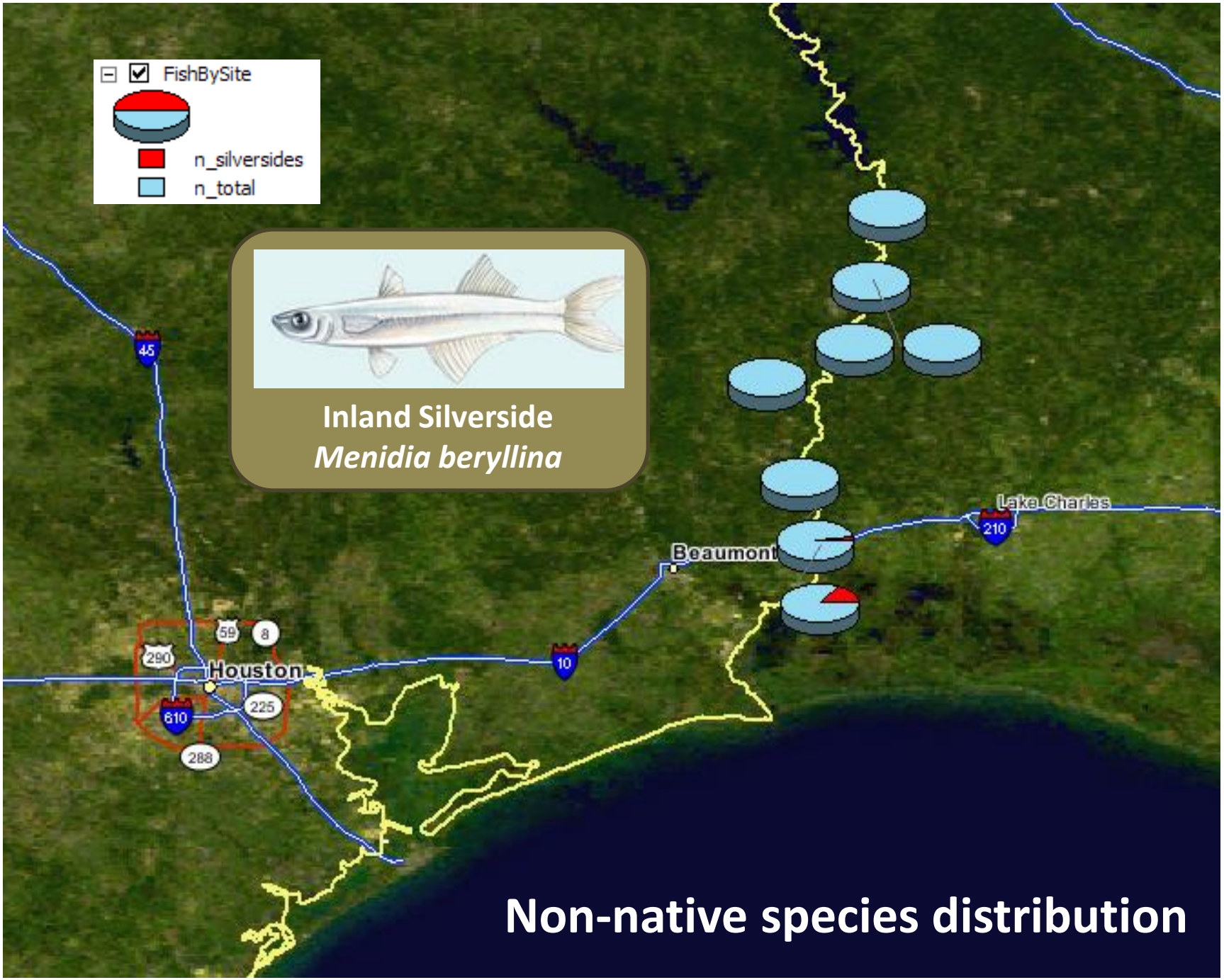
-The only non-native observed was the inland silverside (*Menidia beryllina*). 192 total, ranging from 0-90% of the sample population with a mean of 3% ± 12%





Distribution: Originally found in **coastal waters and upstream in coastal streams** along the Atlantic and Gulf coasts

http://www.bio.txstate.edu/~tbonner/txfishes/menidia%20beryllina.htm


FishBySite
n_silversides
n_total
Inland Silverside
Menidia beryllina
45
Houston
59
8
290
610
225
288
10
Beaumont
Lake Charles
210
Non-native species distribution

Outline

- Background
- An Environmental Flows Information System for Texas
- *Hydroinformatics for Freshwater and Marine Biota*
- Conclusion

The Water Environment

- Distinct types of water data, each with its own character
  - Physical data describe the movement of water and its properties
  - Chemical data describe the constituents moving with, in, and through the water
  - Biological data describe the organisms inhabiting the water environment.

phys/ chem

Surveys

Sensors

Continuous

Longitudinal

vs.

Studies

Samples

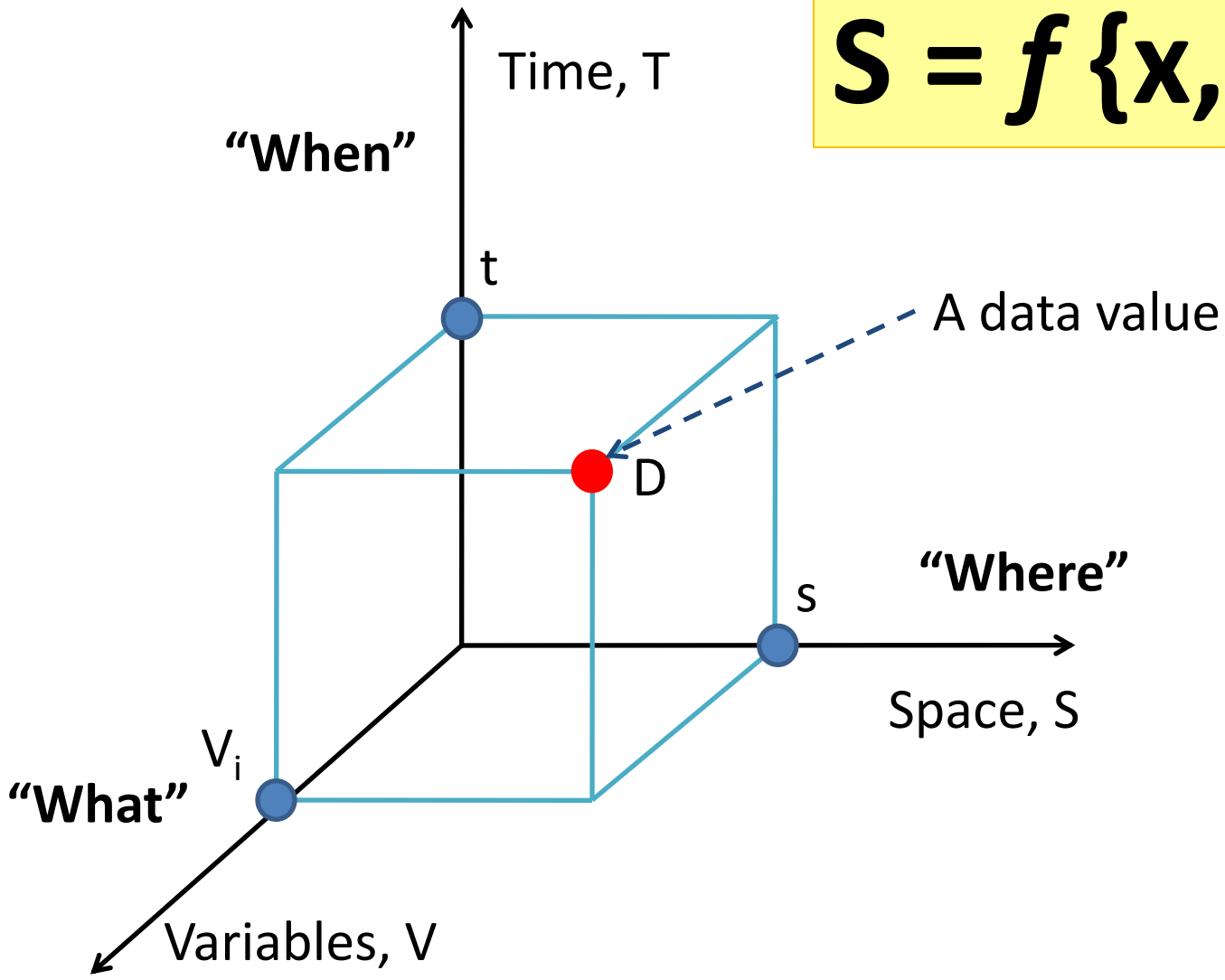
Discrete

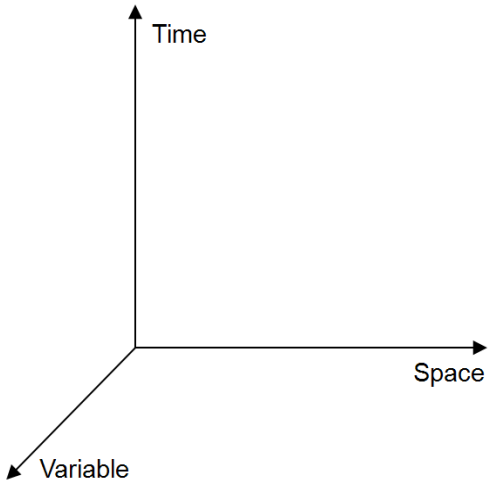
Lateral

biol

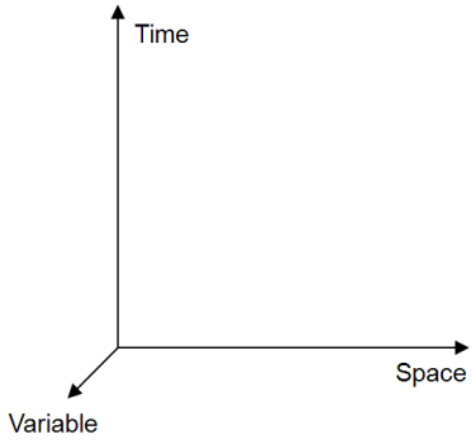
The Data Cube

$$S = f\{x,y,z,t,v\}$$

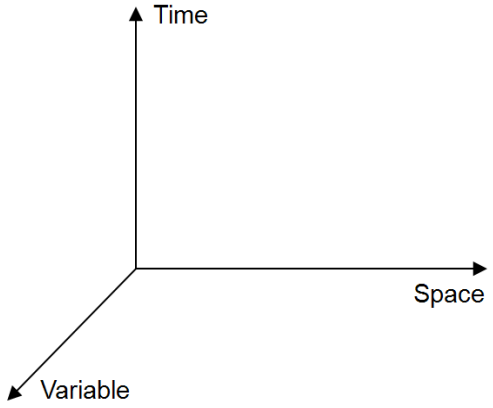




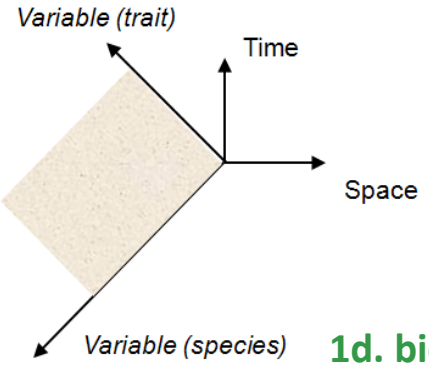
1a. generic data cube



1b. physical



1c. chemical



1d. biological

$$S = f\{x,y,z,t,v\}$$

taxonomy

phenotype

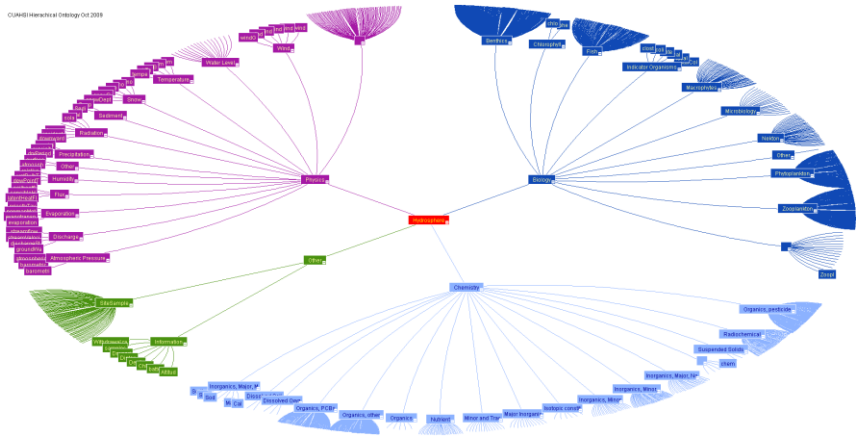
Measurements, traits, and characteristics, such as length, mass, sex, or count.

Ontologies & Semantic Mediation

Category	Records		Code	
	Count	% of TRACS	Count	% of TRACS
TRACS SWQM	7,591,675	100.00%	4,412	100.0%
Site and Sample	184,935	2.44%	82	1.9%
Benthic Macroinvertebrates	49,402	0.65%	1,323	30.0%
Fish	32,710	0.43%	311	7.0%
Phytoplankton	10,099	0.13%	371	8.4%
Zooplankton	10,344	0.14%	266	6.0%
Nekton (non-fish)	2,942	0.04%	31	0.7%
Macrophytes	449	0.01%	54	1.2%
Total, Biologic Data in TRACS	**290,881**	**3.83%**	**2,438**	**55.3%**

eg: 'reservoir inflow' vs. 'streamflow'

Category	Description
Site and Sample	Including: sampling effort, methods, and equipment; substrate, channel geometry, geomorphology, streamflow, cover, vegetation, watershed, aesthetics, and weather
Benthic Macroinvertebrates	Animals without backbones which live all or part of their lifecycle in or near the bottom of freshwater systems. Including: Platyhelminthes (flatworms), Annelids (worms, leeches), Arthropods (mites, insects, crustaceans), and Mollusks (clams, mussels, snails)
Fish	Vertebrate cold-blooded animals that live their entire lives in water, breathe by means of gills, and move by means of fins (with some exceptions)
Phytoplankton	Microscopic, free-floating or suspended plants and algae which have movement depending on currents and are primary producers
Zooplankton	Microscopic animals capable of movement and are secondary producers. Including: crustaceans and rotifers, diatoms, dinoflagellates, and copepods.
Nekton (non-fish)	Free swimming organisms, exclusive of fish as defined above. Including: Decapods (shrimp, prawns, crayfish, crabs), jellyfish, squid, turtles, frogs, alligators
Macrophytes	Large vascular aquatic plants, growing in or near water that are either emergent, submergent, or floating. Including: cattails, rushes, arrowhead, waterlily



http://test.hydroseek.net/ontology/CUAHSIOntologyMay2009.html

Monitoring Site Locations

SpatialReferences
SpatialReferenceID {PK}
SRSID
SRSName
IsGeographic
Notes

ReferencLatLong (1..1 — 0..*)
ReferenceLocalCoords (1..1 — 0..*)

Sites
SiteID {PK}
SiteCode
SiteName
Latitude
Longitude
LatLongDatumID {FK}
Elevation_m
VerticalDatum
LocalX
LocalY
LocalProjectionID {FK}
PosAccuracy_m
State
County
Comments

OffsetTypes
OffsetTypeID {PK}
OffsetUnitsID {FK}
OffsetDescription

OffsetLocationOf (OffsetTypes 0..1 — DataValues 0..*)
Characterize (DataValues 0..* — OffsetTypes 0..*)

Observation Values

Locate (Sites 1..1 — DataValues 0..*)
Generate (DataValues 0..* — Sources 1..1)

DataValues
ValueID {PK}
DataValue
ValueAccuracy
LocalDateTime
UTCOffset
DateTimeUTC
SiteID {FK}
VariableID {FK}
OffsetValue
OffsetTypeID {FK}
CensorCode
QualifierID {FK}
MethodID {FK}
SourceID {FK}
SampleID {FK}
DerivedFromID
QualityControlLevelID {FK}

Create (DataValues 0..* — Methods 1..1)
ResultIn (DataValues 1..1 — Samples 0..1)
AreDerivedFrom (DataValues 1..* — DerivedFrom 0..*)

DerivedFrom
DerivedFromID {FK}
ValueID {FK}

Data Sources

Sources
SourceID {PK}
Organization
SourceDescription
SourceLink
ContactName
Phone
Email
Address
City
State
ZipCode
MetadataID {FK}

DescribeSource (Sources 1..1 — ISOMetadata 0..*)

ISOMetadata
MetadataID {PK}
TopicCategory
Title
Abstract
ProfileVersion
MetadataLink

e.g. "USGS"
"USEPA"

Data Collection Methods

Methods
MethodID {PK}
MethodDescription
MethodLink

Samples
SampleID {PK}
SampleType
LabSampleCode
LabMethodID {FK}

Analyze (Samples 0..* — LabMethods 1..1)

LabMethods
LabMethodID {PK}
LabName
LabOrganization
LabMethodName
LabMethodDescription
LabMethodLink

Variables

DescribeOffsetUnits (Variables 1..1 — OffsetTypes 0..*)

Units
UnitsID {PK}
UnitsName
UnitsType
UnitsAbbreviation

DescribeVariableUnits (Units 1..1 — Variables 0..*)
DescribeTimeUnits (Units 1..1 — Variables 0..*)

Variables
VariableID {PK}
VariableCode
VariableName
VariableUnitsID {FK}
SampleMedium
ValueType
IsRegular
TimeSupport
TimeUnitsID {FK}
DataType
GeneralCategory
NoDataValue

e.g. "Water temperature"
"Dissolved Oxygen Concentration"

Value Grouping

Group (DataValues 1..* — Groups 0..*)
Represent (Variables 1..1 — Groups 0..*)

Groups
GroupID {FK}
ValueID {FK}

Describe (Groups 0..* — GroupDescriptions 1..1)

GroupDescriptions
GroupID {PK}
GroupDescription

Categorical Data

CharacterizeCategorical (Variables 1..1 — Categories 0..1)

Categories
VariableID {FK}
DataValue
CategoryDescription

Data Qualifiers

QualityControlLevels
QualityControlLevelID {PK}
Definition
Explanation

SpecifyQualityOf (QualityControlLevels 1..1 — Variables 1..1)

Qualifiers
QualifierID {PK}
QualifierCode
QualifierDescription

Qualify (Qualifiers 0..1 — DataValues 0..*)

e.g. "Sample holding time exceeded"

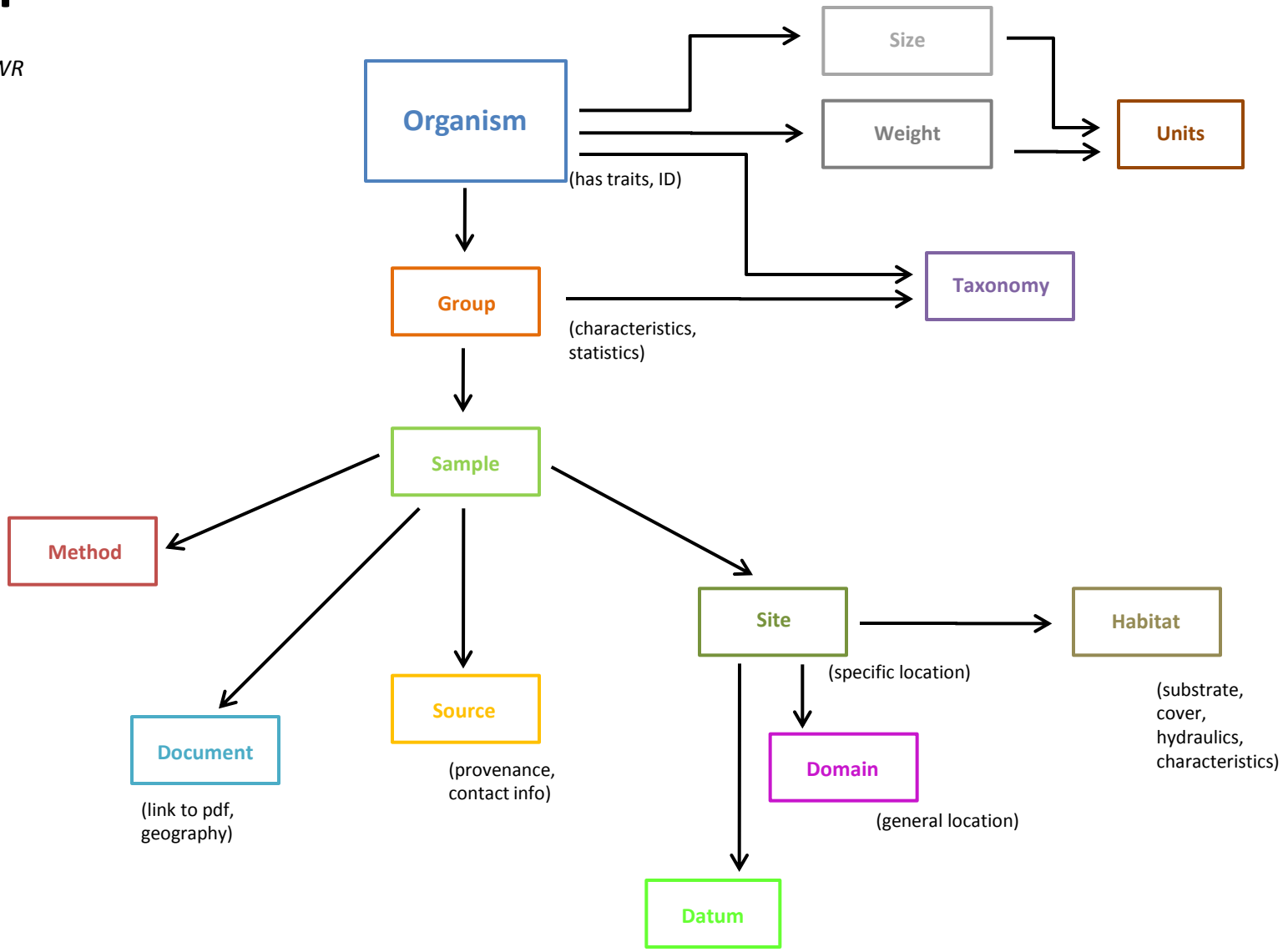
SeriesCatalog
SeriesID {PK}
SiteID {FK}
SiteCode
SiteName
VariableID {FK}
VariableCode
VariableName
VariableUnitsID {FK}
VariableUnitsName
SampleMedium
ValueType
TimeSupport
TimeUnitsID {FK}
TimeUnitsName
DataType
GeneralCategory
MethodID {FK}
MethodDescription
SourceID {FK}
Organization
SourceDescription
QualityControlLevelID {FK}
BeginDateTime
EndDateTime
BeginDateTimeUTC
EndDateTimeUTC
ValueCount

Horsburgh et al (2008)

ODM 1.0

BioODM

v.1.2
Eric S. Hersh, UT-CRWR
9/25/2009



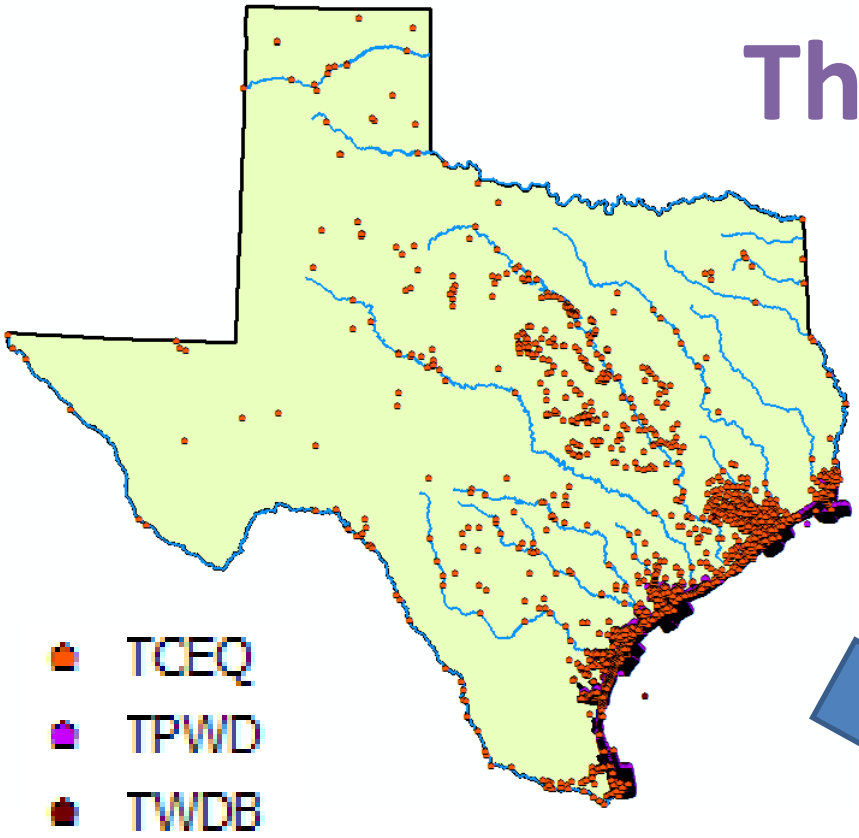
Organism
OrganismID {PK}
SizeID {FK}
WeightID {FK}
Sex
Lifestage
Anomalies
Age
GroupID {FK}
SampleID {FK}
TaxonomyID {FK}
OrganismComments
Group
GroupID {PK}
MinLength
MaxLength
Count
TaxonomyID {FK}
GroupComments
Taxonomy
TaxonomyID {PK}
Kingdom
Phylum
Subphylum
Superclass
Class
Subclass
Infraclass
Superorder
Order
Suborder
Family
Genus
Species
Subspecies
TSN
TaxonomyComments
Sample
SampleID {PK}
MethodID {FK}
SampleDistance
SampleDuration
SampleDrift
SampleDateTimeLocal
SampleDateTimeUTC
OffsetUTC
OffsetVertical
OffsetLateral
SampleMedium
QualityControlLevel
SiteID {FK}
SourceID {FK}
DocumentID {FK}
HabitatID {FK}
SampleComments
Method
MethodID {PK}
MethodDescription
MethodReference
MethodComments
Document
DocumentID {PK}
DocumentTitle
DocumentLink
DocumentAuthors
DocumentYear
DocumentOrganization
DocumentDiscipline
DocumentKeywords
DocumentCitation
DocumentAbstract
DocumentGeography
DocumentComments
Source
SourceID {PK}
SourceName
SourceDescription
SourceLink
ContactName
ContactEmail
SourceComments
Site
SiteID {PK}
SiteName
Latitude
Longitude
Elevation
DatumID {FK}
Accuracy
RiverLocation
DomainID {FK}
SiteComments
Domain
DomainID {PK}
Waterbody
NHDPlusCatchment
NHDPlusSubbasin
NHDPlusBasin
County
State
DomainComments
Habitat
HabitatID {PK}
HabitatType
HabitatDescription
CoverType1
CoverPercent1
CoverType2
CoverPercent2
CoverComments
MeanVelocity
MaximumVelocity
MeanDepth
MaximumDepth
Width
PercentDetritus
PercentVegetation
HabitatComments
PercentClay
PercentSilt
PercentSand
PercentGravel
PercentCobble
PercentBoulder
PercentBedrock
PercentOther
SubstrateComments
Datum
DatumID {PK}
HorizontalDatum
VerticalDatum
DatumComments
store these elements in a GIS
Size
SizeID {PK}
SizeName
UnitsID {FK}
SizeComments
Weight
WeightID {PK}
WeightName
UnitsID {FK}
WeightComments
Units
UnitsID {PK}
UnitsName
UnitsType
UnitsAbbreviation
UnitsComments

Field Name	Data Type	Description	Example	Constraint
TaxaID	Integer, Identity	Unique integer identifier for each taxonomic classification	42	Mandatory; Unique Primary Key
Family	Text (50 char)	Scientific family name	*Centrarchidae*	Optional
Genus	Text (50)	Scientific genus name	*Micropterus*	Optional
Species	Text (50)	Scientific species name	*salmoides*	Optional
TaxaComment	Text (256)	Comments related to the taxonomic identification and/or classification	-	Optional

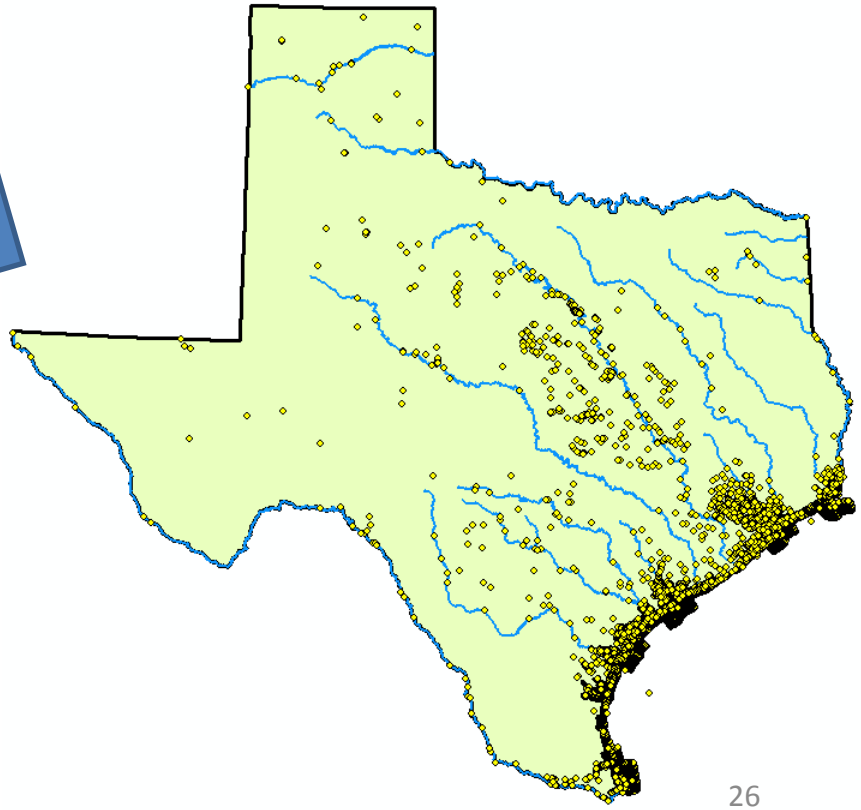
- **2-D variable space**
- **Collections-based focus**
- **Sampling effort**

ITIS Integrated Taxonomic Information System

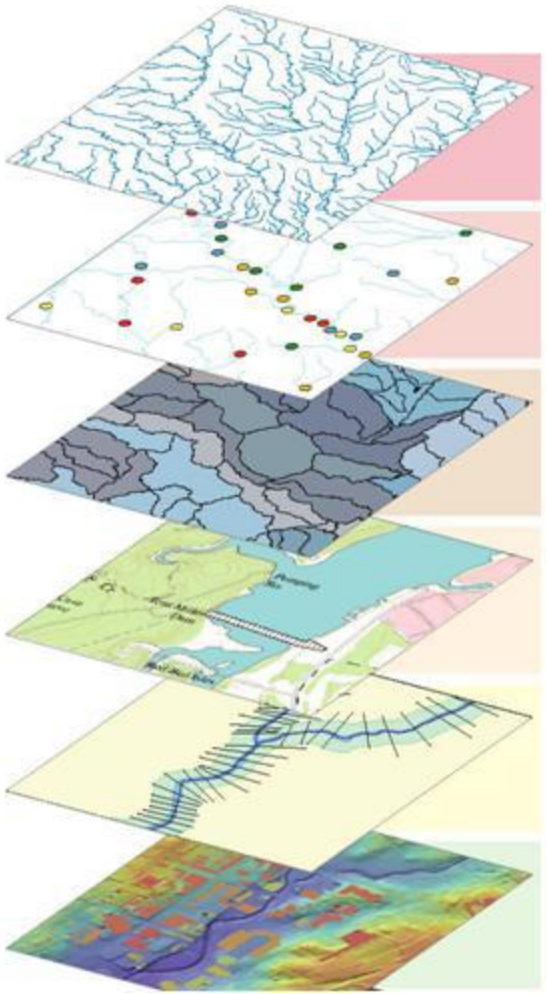
Themes



"It's the salinity, stupid."
- Paul Montagna, SAC, 2/09



Environmental Flows Themes



Geospatial Data

Water Quality

Climatology

Hydrology & Hydraulics

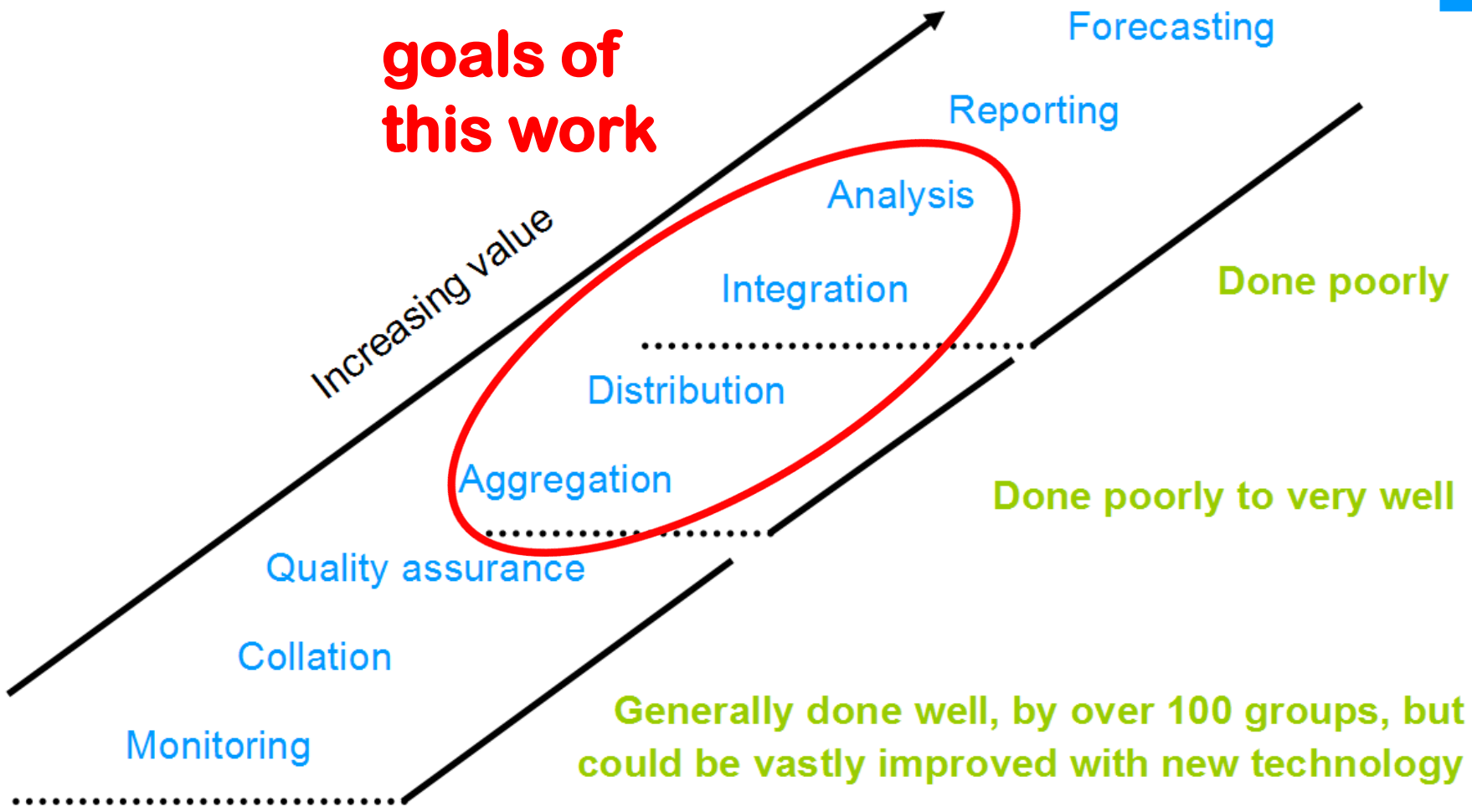
Geomorphology & Physical Processes

Aquatic Biology

Outline

- Background
- An Environmental Flows Information System for Texas
- Hydroinformatics for Freshwater and Marine Biota
- *Conclusion*

The water information value ladder.



From Robert Vertessy, CSIRO, Australia

thanks.