Electronic Media and Environmental Science: A Discussion from e-Education to e-Data

S. L. Brantley

Director, Earth and Environmental Systems Institute

EarthTalks, January 9, 2012

Cyberinfrastructure and Social Media and Sensors and Databases and Online Teaching and Other Cyberstuff…

~~Electronic Media~~ and Environmental Science: A Discussion from e-Education to e-Data

S. L. Brantley

Director, Earth and Environmental Systems Institute

EarthTalks, January 6, 2012

Jan 16	*Ann Taylor*, Acting Director, Dutton e-Education Institute	"e-Education for Geographers, Geoscientists and Meteorologists"
Jan 23	*Michael Mann*, Professor of Meteorology, Penn State	"The Hockey Stick and the Climate Wars: Dispatches from the Front Lines"
Jan 30	*Richard Alley*, Evan Pugh Professor of Geosciences, Penn State	"Communicating Real Science in a Sound-Bite World"
Feb 6	*Jeffrey Brownson*, Asst Professor of Energy and Mineral Engineering, Penn State	"Emergent e-Education: new strategies for research and engagement linked to the cloud classroom"
Feb 13	*Rick Hooper*, Exec Director, Consortium of Universities for Advancement of Hydrologic Sciences (CUAHSI) *Abstract*	"Sharing Water Data for Research, Education, and Citizen Engagement: The web services revolution"
Feb 20	open	
Feb 27	*Tim Spangler*, Director, COMET, Boulder, CO	"The COMET Program: Resources for Innovative University Instruction and the Evolution of E-learning"
March 12	*Bill Brune, Lee Kump, Karl Zimmerer*, EMS Department Heads	"Panel Discussion: The Present and Future of e-Education in Meteorology, Geosciences and Geography
March 19	open	
March 26	open	
April 2	open	
April 9	*Rodney Erickson*, President, Penn State University	"e-Education and the Research University: A Conversation"
April 16	open	

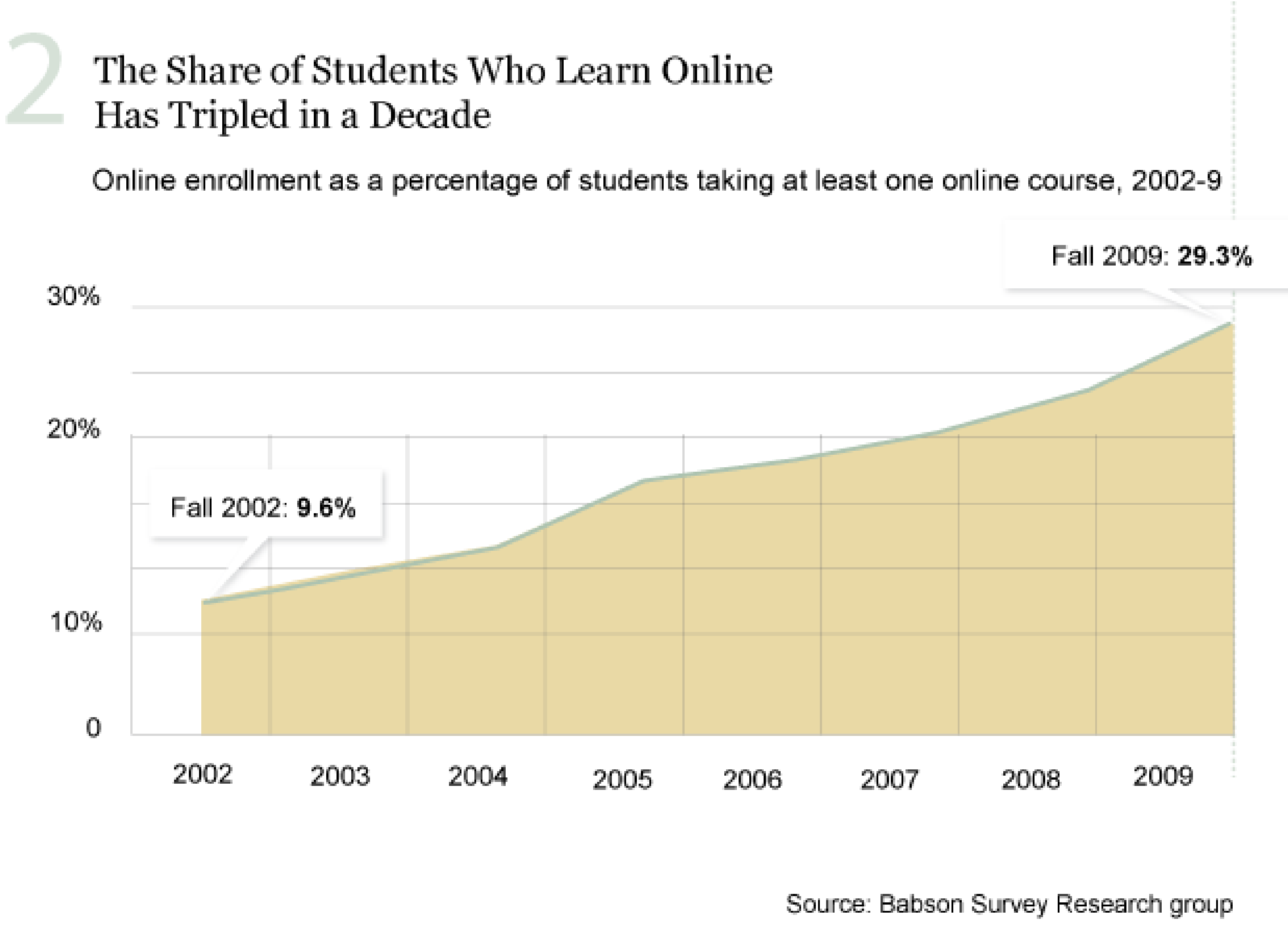
Khanjan Mehta (COE); Eva Zanzerkia (NSF EarthCube); Jim Campbell (USGS, PA Water Science Center)...

Outline of the Talk

- Thoughts about how our lives are changing as university teachers due to the cyber-revolution
- Thoughts about how our lives are changing as university environmental researchers due to the cyber-revolution
- Hypotheses about the future
- Discussion

How our lives are changing as environmental teachers

How many students are learning online?



Also:

Sloan foundation reports that online enrollments grew by 10% in 2010 against 2% for the sector as a whole

Chronicle of Higher Education, B20, Special Section , Online Learning; Nov 11, 2011

But online learning is not restricted to universities: Mobile phone training programs (for example) are growing too

40.7% of companies are considering using mobile devices for training

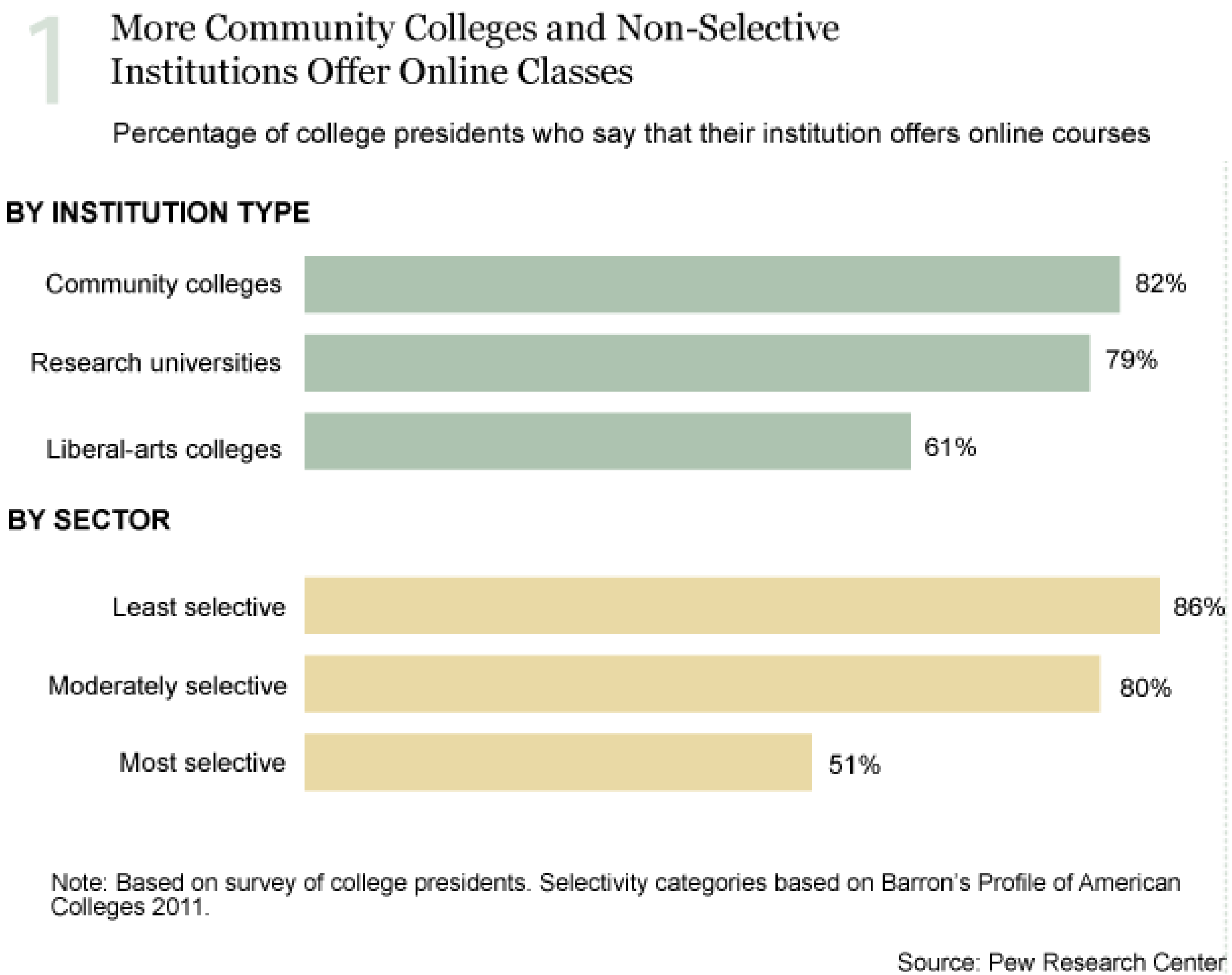
10.1% of companies are developing mobile device applications for training

15% of companies are already using mobile device applications for training

34.7% of companies have no plans to use mobile devices for training

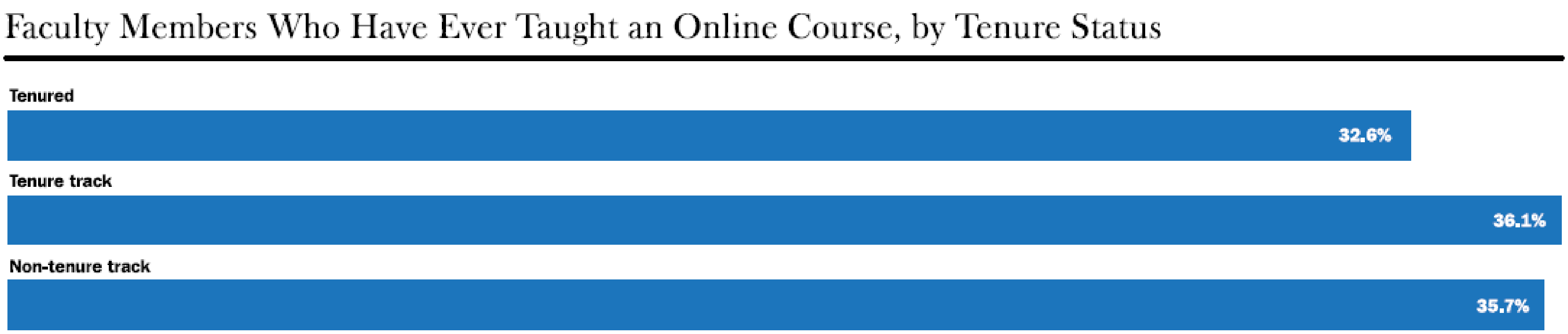
(“Mobile Learning in the Palm of your Hand”, by the American Society for Training and Development; quoted in Sky Magazine, October 2011)

Who offers online courses?



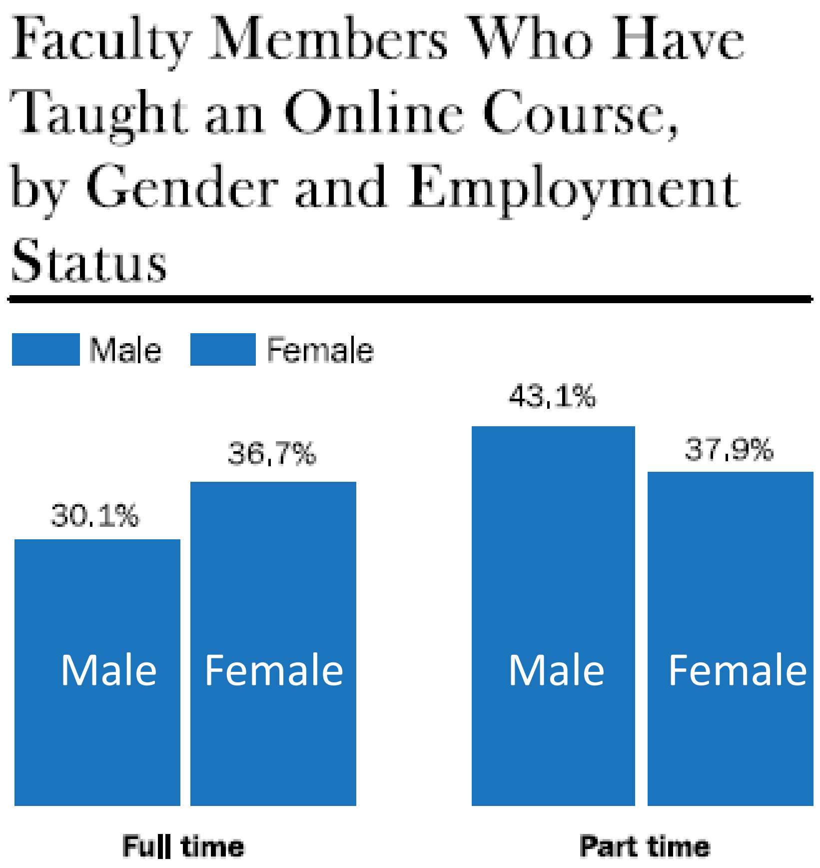
Chronicle of Higher Education, B20, Special Section , Online Learning; Nov 11, 2011

Who teaches online courses?



http://chronicle.com/article/Faculty-Views-About-Online/125200/; accessed 1-5-11

Who teaches online courses?



http://chronicle.com/article/Faculty-Views-About-Online/125200/; accessed 1-5-11

EMS: Online education by the numbers

- EMS entered the distance education arena in 1998: Geography launched online certificate program in GIS, delivered through World Campus
- College now has 3 online post-baccalaureate certificates; 2 online masters degrees, 1 online bachelor degree program
- 80 EMS courses are available on the web: reaching more than 5100 undergrads and >1200 adult professionals worldwide
- Geography: 71 residential graduate students; 153 online graduate students in MGIS

What is causing the growth of online education: New online capabilities or trends

- Skype-based seminars
- Live question and answer sessions
- Chat rooms
- Email
- Twitter
- I-phone apps
- Web-inars
- Video conferencing
- Web cameras are everywhere
- Social media (Facebook, Yammer, LinkedIn)
- Gaming technologies
- Simulation engines
- Mobile phones for training and interaction with students
- I-pads for on-demand training
- Digital whiteboards
- Blogging
- Adaptive learning modules

"Most colleges…have yet to find ways to use technology to really transform education…Technology has fundamentally changed the productivity of every industry in America except education. In nearly all of higher education, it is an add-on cost."

Robert W. Mendenhall, President of Western Governor's University

What is causing the growth of online education: An example of a new online capability even for advising

- “Think of the problem in terms of a supermarket cereal aisle, says Tristan Denley, provost of Austin Peay State University, in Clarksville, Tenn. You find every choice known to man. But unless you've opened the box, you have very little information to judge what's inside. How do you pick one?

- Part of the answer, he says, is technology that can look at people like you who have made such decisions in the past, and see whether those decisions worked out. In April, Austin Peay debuted software that recommends courses based on a student's major, academic record, and how similar students fared in that class.

- Some professors fretted about students misinterpreting the Netflix-like tips as commands, but the Gates Foundation quickly ponied up $1-million to refine the software so other colleges can adopt it.

- Now Austin Peay plans to expand on its work with a new tool that offers tips for making a more important decision: picking a major.”

Quoted from the Chronicle of Higher Education; http://chronicle.com/article/A-Moneyball-Approach-to/130062/; accessed January 2012

What is causing the growth of online education: Drive toward sharing course materials

"By using the technology to teach – to deliver the content of a course – we are able to free students to study what they need to learn, and to do so at their own pace. **Learning becomes the constant and time becomes the variable, rather than holding time constant and letting the learning vary.** In an online environment that truly takes advantage of technology, the faculty role may change from delivering content to mentoring students."

Robert W. Mendenhall, President, Western Governors University

Lots of examples of how or where course materials are shared or sold:

- The Univ of Southern CA has shared resources to create or improve online courses
- Open Education Resource University – a consortium of universities worldwide grants credit to students who can pass assessments
- MIT's lectures and materials are available free online..students can earn a certificate if they pass assessments
- The New York Times Knowledge Network
- University of the People: nonprofit, tuition-free, 1200 students, 120 countries, processing fees of $10-$100; volunteer professors; www.uopeople.org

What is causing the growth of online education: Cost of education

- Ex U.S. students have debts of $1 trillion (The Economist)

- College tuition has risen four times faster than inflation (B. Smith, The Chronicle of Higher Education, Nov 11, 2011)

- "The ranks of the most expensive colleges have grown again:

  100 institutions …charg[e] $50,000 or more for tuition, fees, room, board in 2010-11…

  58 universities and colleges ..charged that much in 2009-10, …

  5 colleges were priced over $50,000 [in 2008-9].

- [2010] marks a milestone as the first public institution has joined that elite club: the University of California at Berkeley is charging out-of-state residents $50,649 for tuition, fees, room, and board. (The price for in-state residents is only $27,770.)

- All of the other 99 colleges charging $50,000 or more are private. They made up 9 percent of the 1,058 private institutions reporting any amount for tuition, fees, room, and board."

(from Jeffrey Brainard, October 31, 2010, Chronicle of Higher Education)

Penn State as of 2011/2012 room + board + tuition in-state is $26742; out-of-state is $36626

Which ten university presidents were invited to the White House on Dec 5?

- Dr. King Alexander, President, California State University – Long Beach
- Dr. Francisco Cigarroa, Chancellor, University of Texas System
- Dr. Jared Cohon, President, Carnegie Mellon University
- Dr. Freeman Hrabowski, President, University of Maryland – Baltimore County
- Dr. William “Brit” Kirwan, Chancellor, University System of Maryland
- Dr. Larry Shinn, President, Berea College
- Mr. Thomas Snyder, President, Ivy Tech Community College
- Dr. Holden Thorp, Chancellor, University of North Carolina – Chapel Hill
- Dr. Nancy Zimpher, Chancellor, State University System of New York
- Dr. Robert Mendenhall, President, Western Governor’s University

Western Governor's University

- Public, all online university that was founded in 1996 by 19 state governors (The Economist)
- Began offering courses in 1999; now offers 50 degrees
- Now enrolls more than 25000 U.S. students
- Professors decide what students must know
- University buys teaching materials from publishers and pays student mentors for each student
- Professors design assessment tools: proctored tests, projects, papers, etc.
- Teams of graders are employed to do the grading
- Third party assessments used where possible
- Accredited by Northwest Commission on Colleges and Universities (one of the 6 regional accrediting agencies)

"Faculty remain critically important, but their focus moves from preparing lectures to monitoring data about student participation and performance, engaging in rich dialogue with individuals…embellishing curricula where appropriate, introducing supplemental resources, and developing new content modules [when needed]."

Diane Auer Jones, former U.S.Assistant Secretary of Education, Chron. Higher Ed, Nov 2011

Financial models for online degree programs: Non-Profits (NP) and For-Profits (FP)

-Penn State is a NP but uses a FP model for tuition for online degree and certificate programs generating a revenue stream for academic units

	NP	FP	FP	FP	FP	FP	FP	NP
Program	**Western Governors University**	**Capella**	**Walden**	**University of Phoenix**	**Grand Canyon University**	**Kaplan**	**Devry**	PSU
B.S. Business Management	$2,890 per six month term. If it takes four years= **$23,120**	180 credits x $310* per credit = **$55,800**	181 credits x $280 per credit = **$50,680**	120 credits x $570 per credit = **$68,400**	120 credits x $450 per credit = **$54,000**	180 credits x $371 per credit = **$66,780**	Total tuition** = **$65,496**	8 semesters of 12+ credits = $49,000
B.S. Information Technology (or Equivalent)	$2,890 per six month term. If it takes four years= **$23,120**	180 credits x $310* per credit = **$55,800**	181 credits x $280 per credit = **$50,680**	120 credits x $570 per credit = **$68,400**		181 credits x $371 per credit = **$66,780**	Total tuition** = **$65,496**	
B.S. Nursing (RN to BSN)	$3,250 per six month term. If it takes two years= **$13,000**	90 credits x $310* per credit = **$27,900**		60 Credits x $500 per credit = **$30,000**	36 credits x $450 per credit = **$16,200**	90 credits x $315 per credit = **$28,350**		
Master of Business Administration	$3,250 per six month term. If it takes two years= **$13,000**	48 credits x $678 per credit = **$32,544**	36 credits x $810 per credit = **$29,160**	36-54 credits x $715 per credit = **$25,740 - $38,610**	54 credits x $530 per credit = **$28,620**	60 credits x $441 per credit = **$26,460**	48 credits x $751.66 per credit = **$36,080**	
M.S. in Educational Leadership (or equivalent)	$2,890 per six month term. If it takes two years= **$11,560**	48 credits x $415 per credit = **$19,920**	36 credits x $465 per credit = **$29,160**	40 credits x $575 per credit = **$23,000**	38 credits x $495 per credit = **$18,810**			

*$310/credit cost based on an average of $275 per credit for lower-division courses and $345 for upper-division courses.

** Total tuition cost provided by DeVry based on tuition costs of $597/credit for credit hours 1-11 and $360/credit for credit hours 12 and above.

Chart from http://www.wgu.edu/tuition_financial_aid/overview

Basic economic model of the university from 1500s to 1900s

- Subject matter experts are scarce so it makes sense to build big campuses and attract professors and provide teaching on campus
- Large fixed costs: adding few more profs is relatively cheap…substantial fixed costs w/ low marginal costs to offer one more class
- Profs attract students
- Strongest signals of value in this model is physical aspects of campus and faculty credentials
- Accreditation measures these variables

From article by Burck Smith, member of the American Enterprise Institute's Higher Education Working Group, in The Chronicle of Higher Education, Nov 11, 2011

Economic model of online university

- Uses real-time and asynchronous communication
- Low fixed costs w/ low marginal costs to offer one more class
- Location of profs and students is irrelevant
- Content can be cheap or free
- Online software for learning is becoming cheaper
- Strongest signals of value in this model are outcomes from courses such as Gen Ed courses and skills-based courses
- Accreditation (today) measures whether the college is set up to deliver online courses; whether college is using best practices; whether faculty have been trained to use online services; whether student-support services are sufficient

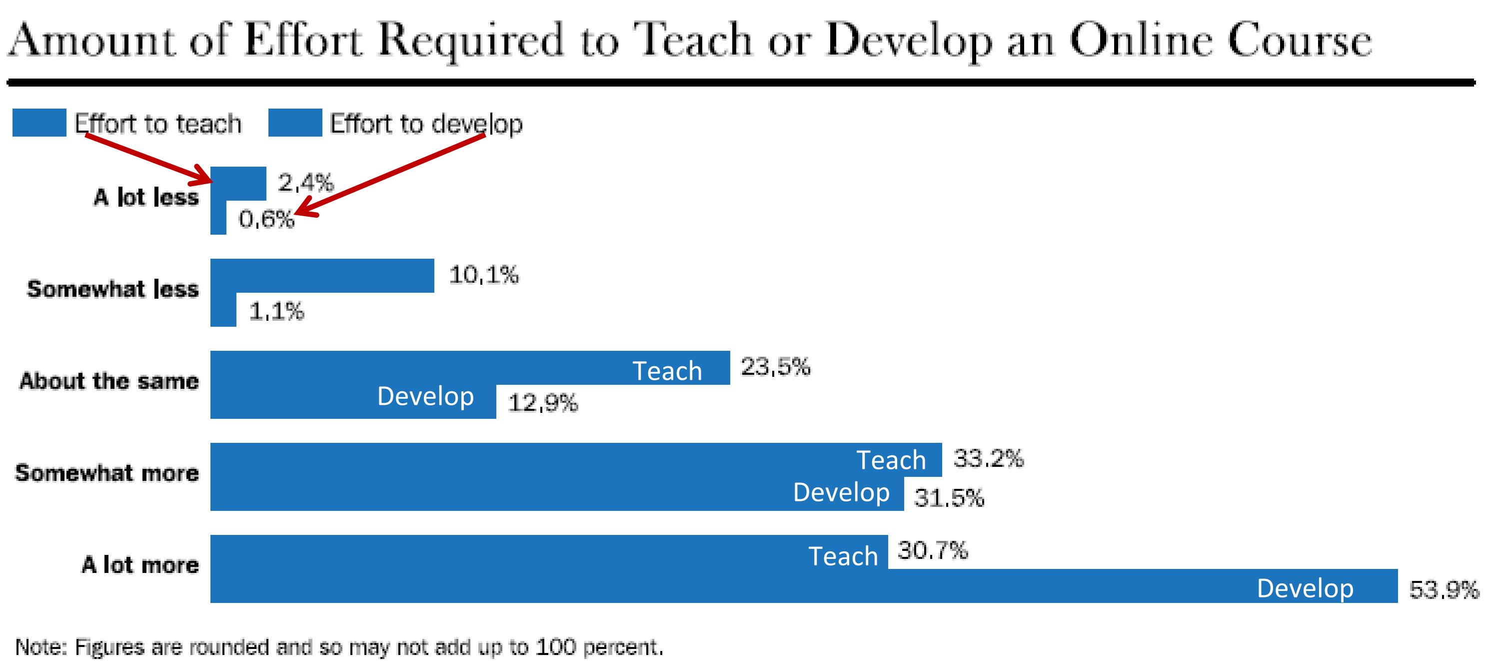
From article by Burck Smith, member of the American Enterprise Institute's Higher Education Working Group, in The Chronicle of Higher Education, Nov 11, 2011

Issues around online programs: Accreditation

- Accreditation: the 1965 Higher Education act made student eligibility for federal student aid contingent upon college accreditation
- Six regional accreditation agencies were set up; but they neither have staff nor legal authority to conduct investigations; reviewers are volunteers from peer institutions of higher education
- In 2001, the 6 agencies adopted a common set of broad standards for accrediting online programs; standards updated in 2006
- Eric Kelderman: "The hundreds of for-profit colleges that rely heavily on online education receive nearly 90% of their revenue from federal student aid."
- In 2009, Dept of Education's Office of the Inspector General recommended limiting the accrediting authority of North Central's Higher Learning Commission (which oversees colleges in 19 states mid-country) because they had accreditted American InterContinental University, a for profit college; Govt Accountability Office also alleged abuses in recruiting/enrollments
- Led to a congressional hearing in 2010 and stricter rules for credit hours
- Enrollments at the 10 largest for-profit colleges were down on average 14% in 2011

From: E. Kelderman, "Online programs face new damands from accreditors", The Chronicle of Higher Education, Nov 11, 2011

Issues around online education: Effort required to teach an online course



Ray Schroeder (Director of Online Learning, Research and Service at Univ Illinois Springfield: "*On the average, I have about 100 exchanges per student over the course of the semester, which is more than I would have in a traditional classroom*." (Sky Magazine, Oct 2011)

Issues around online programs: Questions about quality

- Not everything can be taught online
- But … a 2010 meta-analysis and review of online-learning studies (U.S. Dept of Education), concluded that online learning was as good as or slightly more effective then traditional face-to-face institutions. (Chronicle of Higher Ed.)
- Some studies have shown lower graduation rates for online programs – but of course different populations of students are served.
- PSU offers online courses and grades for these courses are not distinguished on a transcript from on campus courses – doesn't this mean we think online vs. campus delivery is equivalent?

Positives and Negatives of New Cyber-stuff for Teaching

Positives

- Online classes can reach more people than on-campus classes
- Cheaper for a university to produce
- Can share online materials among professors and universities
- Some materials can be taught online very effectively
- More students will be working while getting a degree…they can do this while online
- Devices can learn preferences/needs of learners and can tailor education to the student

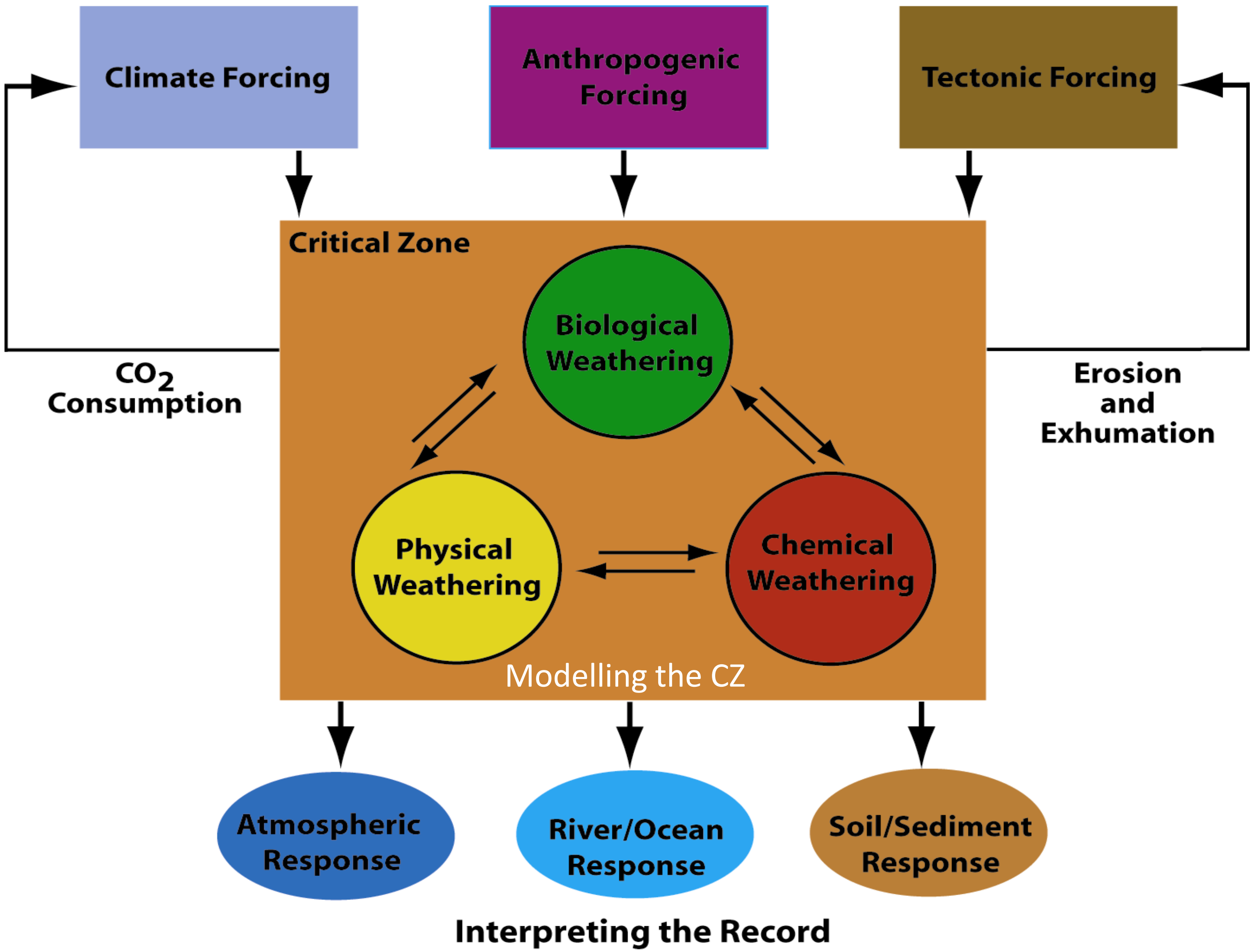
Negatives

- Students may drop out at higher rates (but different populations are involved)
- Human-human engagement is lost, and so are some aspects of learning
- How do we know students are taking the tests?
- Students do not learn to interact in teams face to face
- Laboratories and field trips are a challenge or impossible
- Some students do not learn as much using some online technologies
- Issues around accreditation

How our lives are changing as environmental researchers

Environmental scientists observe the natural system, develop sensor networks, share data sets, and model data to produce knowledge

For example, in Critical Zone science we study earth's surface today in order to project how it will look and act tomorrow



Sensors in the Environment

- Sensors now available to measure aerosols, scintillation in ionosphere, temperature, wind speed, wind direction, solar insolation, soil moisture, rainfall, streamflow velocity, plate motions, etc. etc. (Still hard to make many continuous chemical or biological measurements but that is coming too)
- Mobile phones can collect data
- Enviro/social scientists can monitor Twitter and other social media to learn about environmental change
- Communication technologies can gather the data and log it and stream it in real time to the web
- Other scientists are putting together cyberinfrastructure to house all kinds of environmental data and to find it and use it and model it

NSF EAR Data Sharing Policy

- Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing.

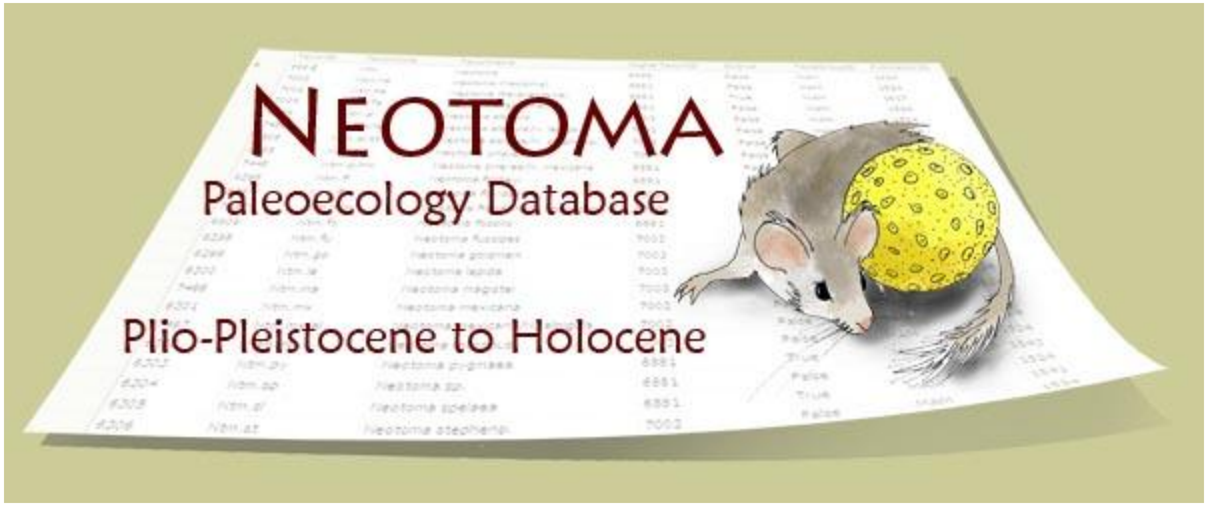
More NSF Data Policy

- Plans for data management and sharing of the products of research. Proposals must include a supplementary document of no more than two pages labeled "Data Management Plan". This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see AAG Chapter VI.D.4), and may include:

  the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;
- the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);
- policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;
- policies and provisions for re-use, re-distribution, and the production of derivatives; and
- plans for archiving data, samples, and other research products, and for preservation of access to them.

Neotoma Paleoecology Database

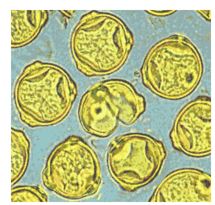
www.neotomadb.org/ *(Slide from Russ Graham, Penn State)*



Neotoma is a multiproxy paleoecology database that includes fossil data for the past 5 million years, the time during which modern species, including humans, and modern ecosystems appeared.

Eon	Era	Period		Epoch	Start Date (mya)
Phanerozoic	Cenozoic	Quaternary		Holocene	0.01
				Pleistocene	1.64
		Tertiary	Neogene	Pliocene	5.2
				Miocene	23.3
			Paleogene	Oligocene	35.4
				Eocene	56.5
				Paleocene	65

Time Interval Covered



Pollen

Vertebrates

Plant Macrofossils

Beetles

Current Data Components
More to come +++

Outreach
&
Educational Activities
(K-12 & University)

Philosophy

Data Sharing

Open and Easy Access

High Quality Data: contributed and maintained with quality control by disciplinary communities

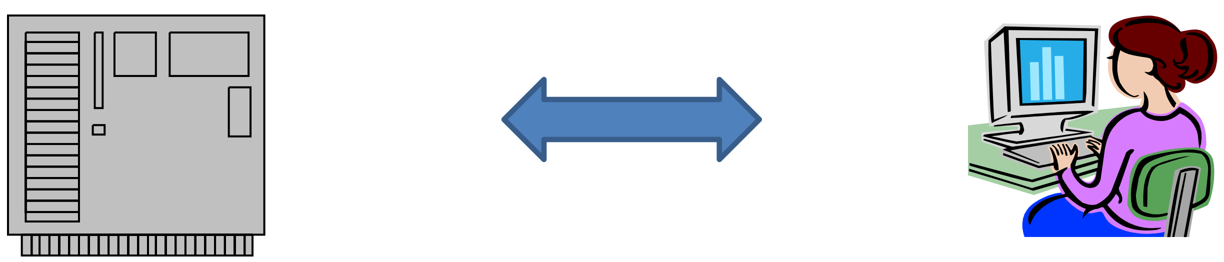
Easy Interfaces with Other Databases

Stimulate New & Innovative Research

Cost Effective Data Management

Neotoma is a single merged database

Neotoma is NOT a distributed database



Database Standardization

Easily Searched

One Format for Data Entry for All Data Types

Facilitates Comparison of Different Data Sets

Cost Effective

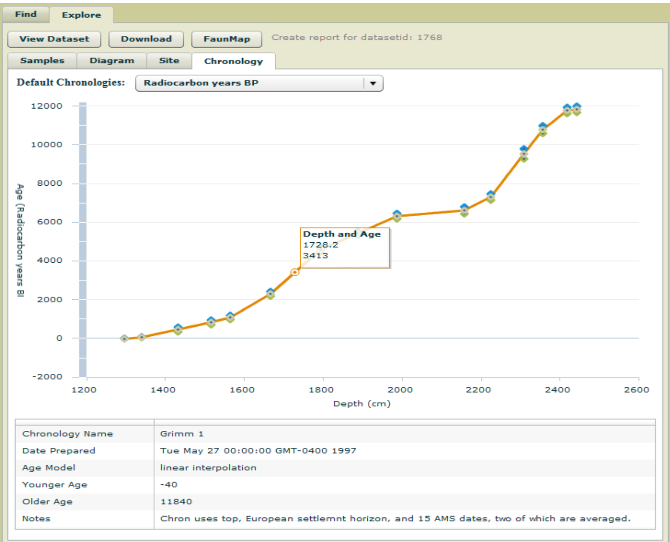
Reduced personnel to administer

Common tools for data entry and analysis

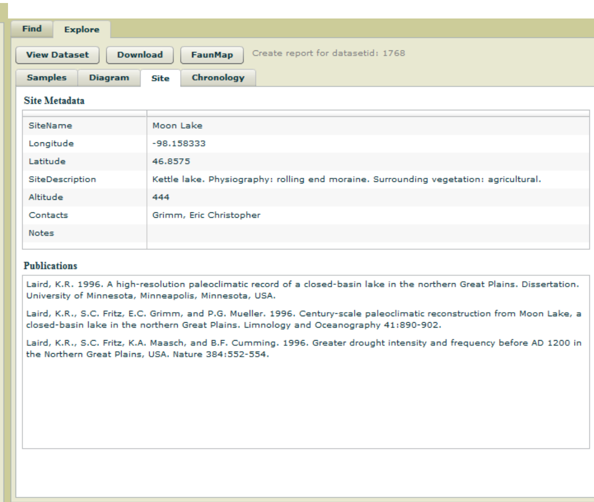
Neotoma Explorer Output Types:

Search sites/datasets by spatial, temporal, and metadata criteria.

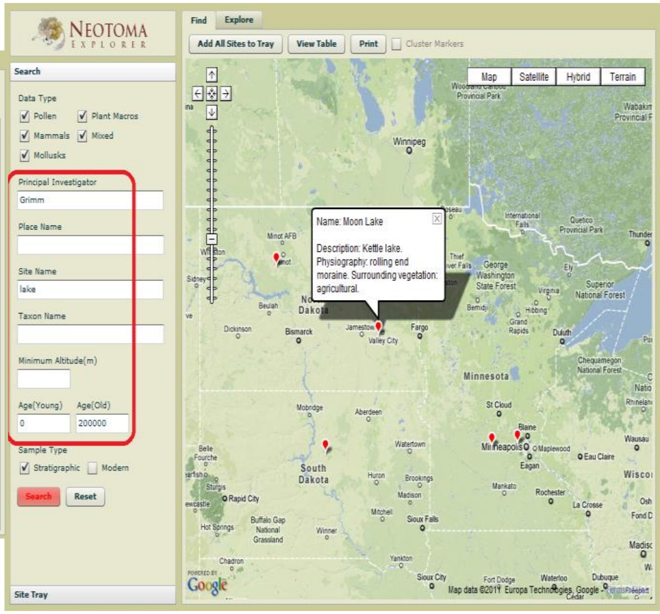
Data Summary (Chronology)

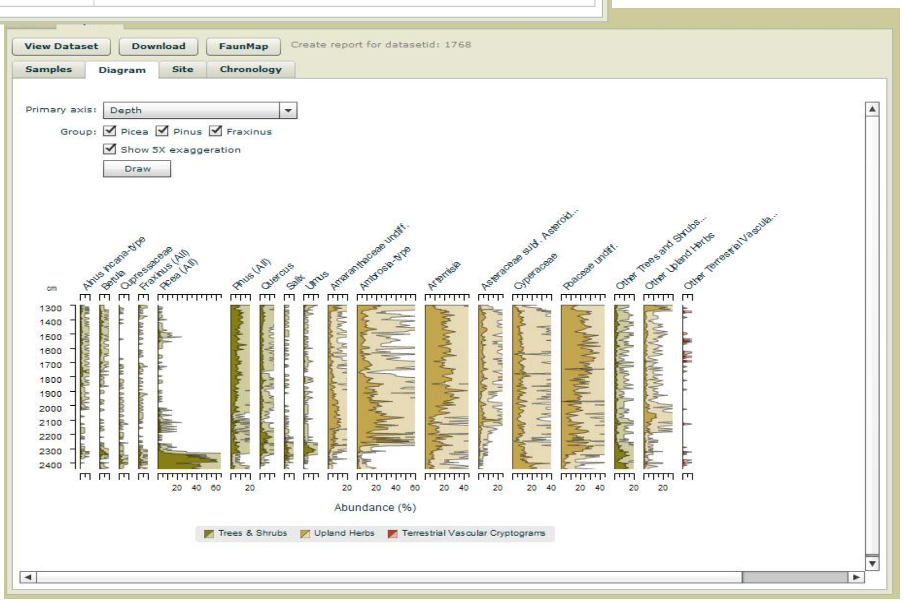


Site Summary

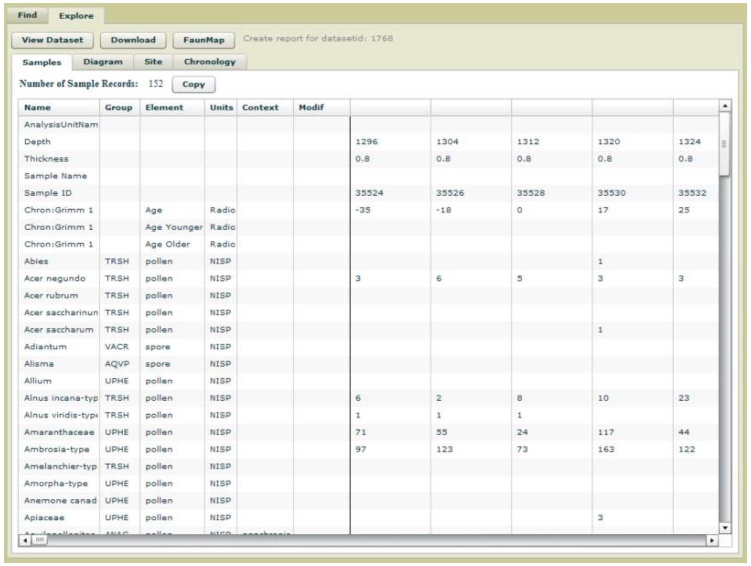


Maps



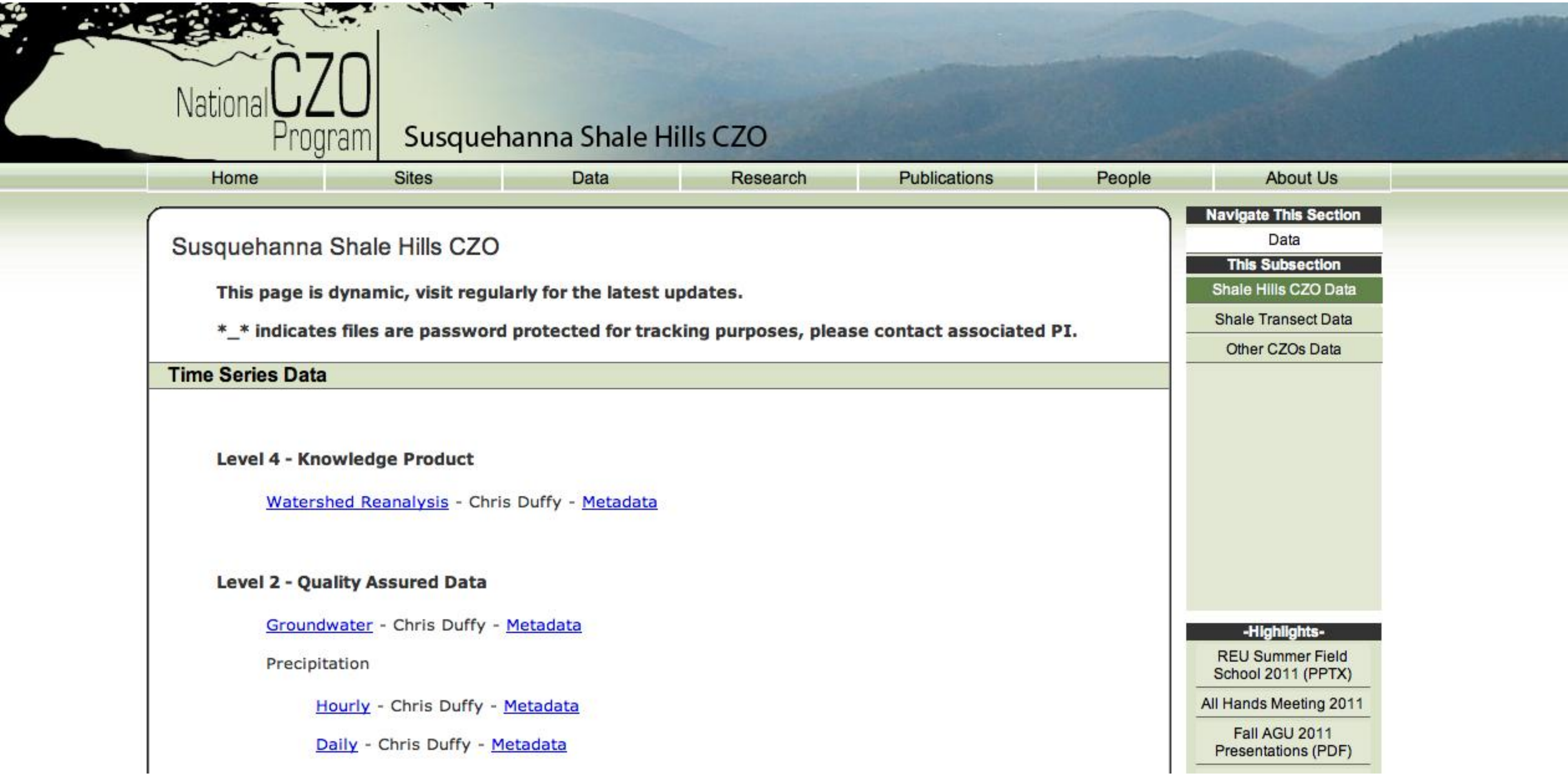


Summary Diagrams (pollen diagram)

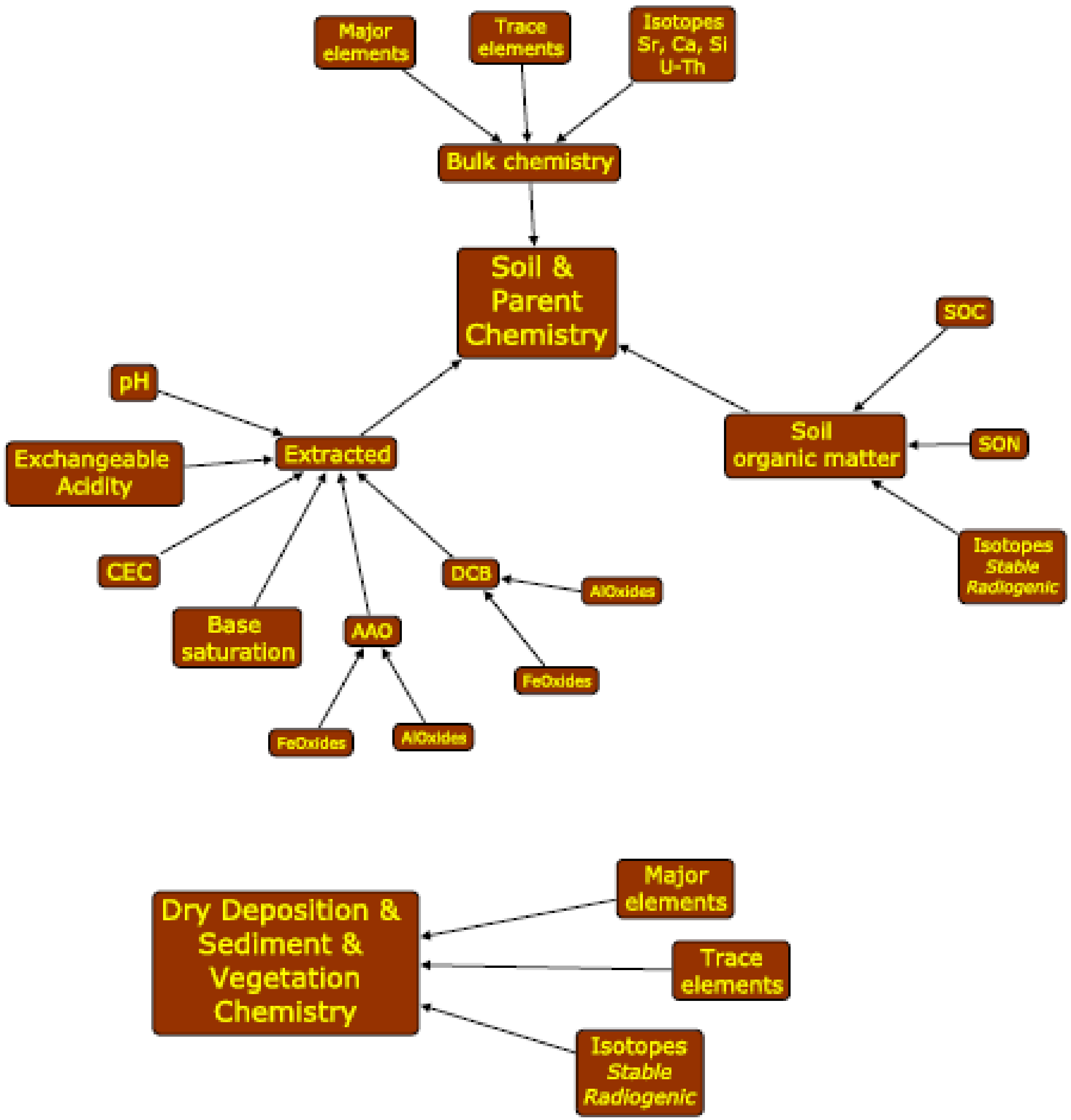


Raw Data

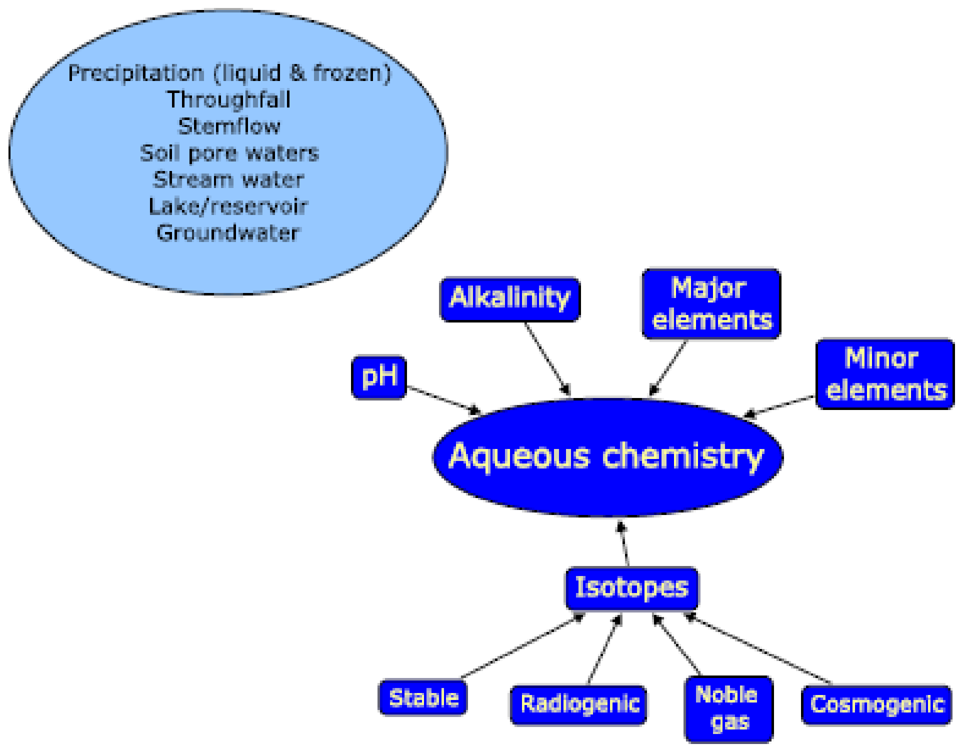
Shale Hills Critical Zone Observatory



Solid chemistry

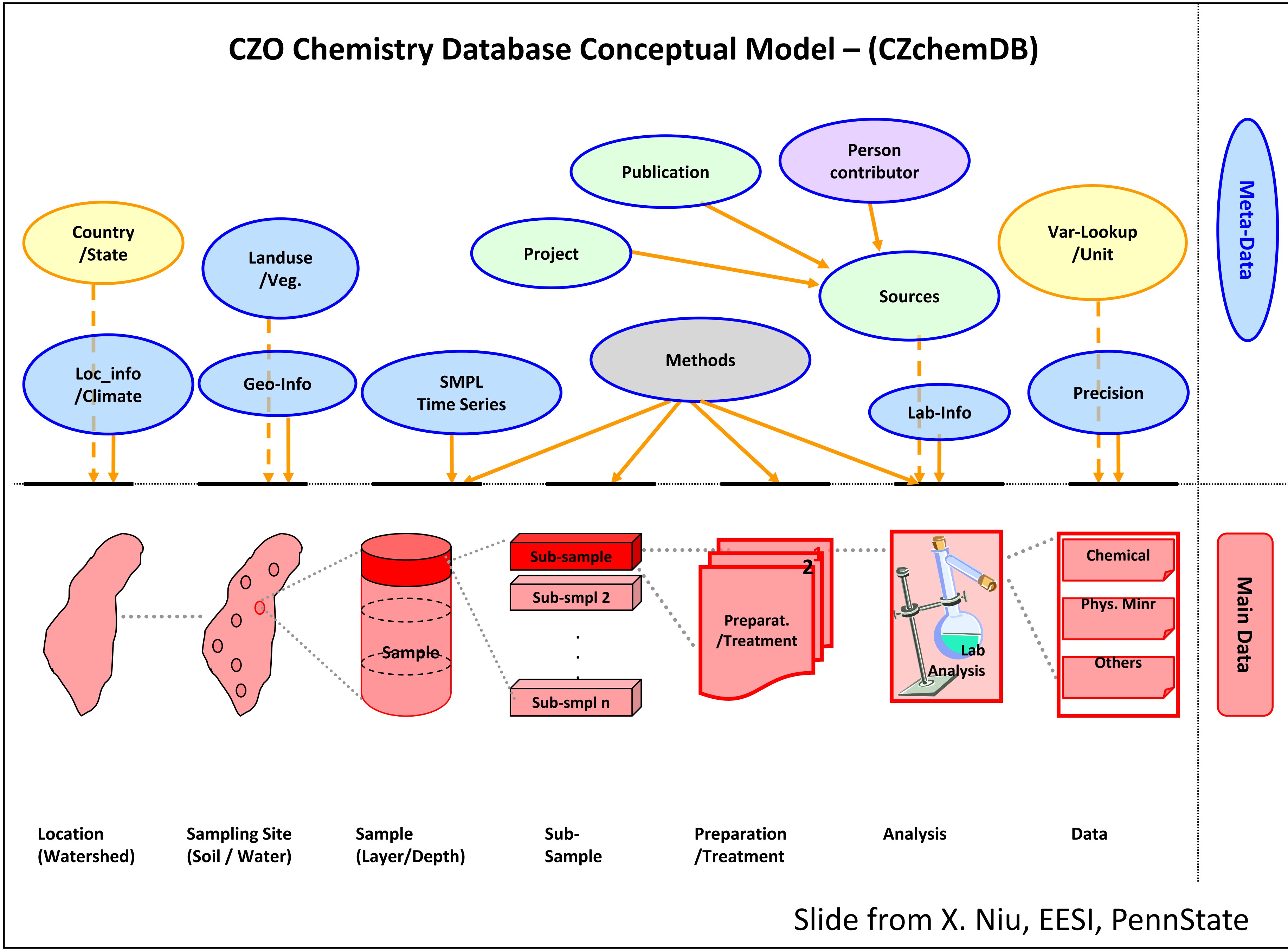


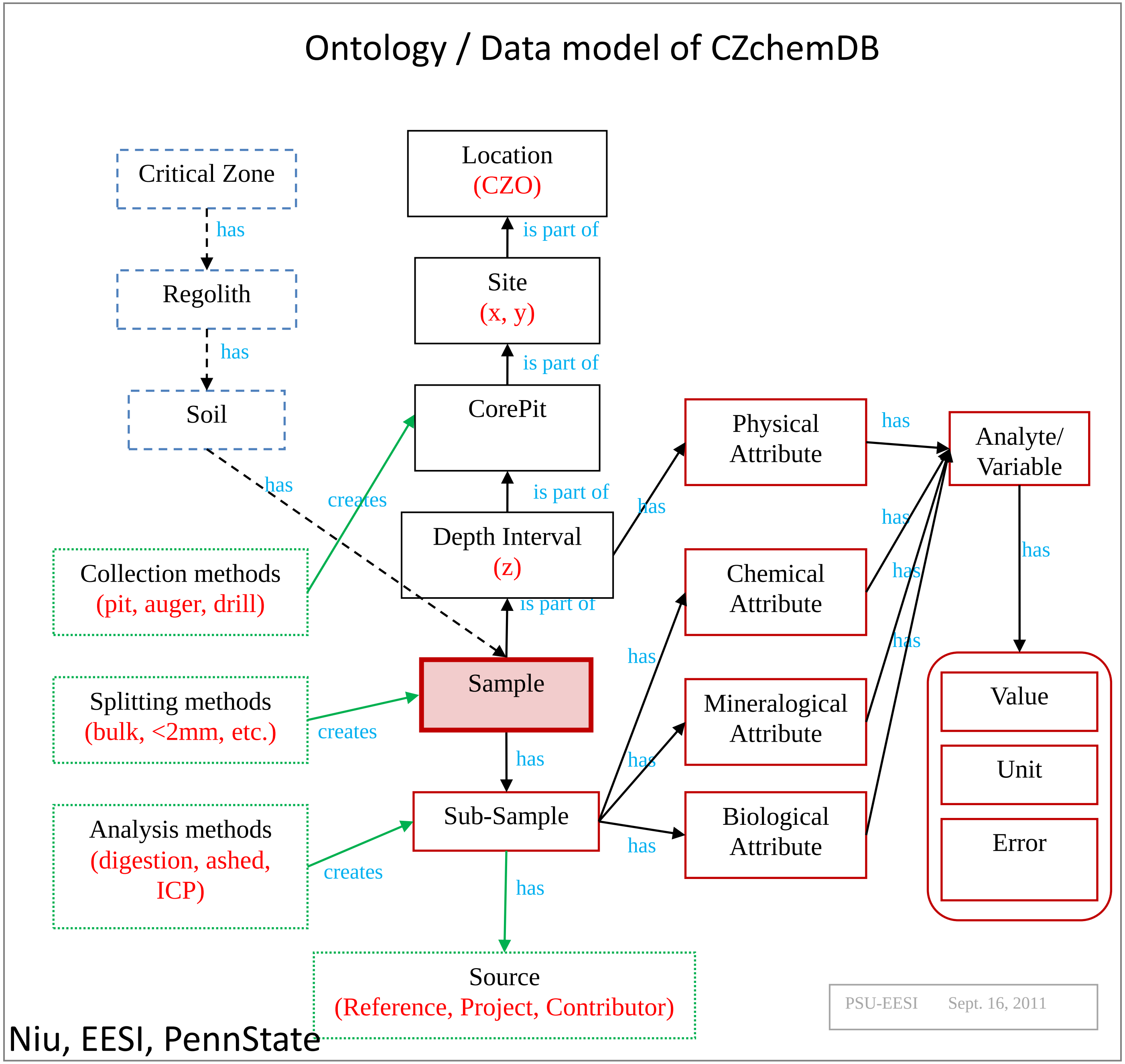
Aqueous & gas chemistry



Major Components of CZchemDB

- Main Data:
  - Location/Site info – Geo-info, climate, landuse, etc.
  - Sampling info – Time, methods, treatment/preparation, etc.
  - Data – Chemical, physical and mineral properties and others
- Meta Data
  - Methods – Sampling, preparation and lab-analysis, etc.
  - Data quality – Precision, StDev, detection-limit, etc.
  - Source – Publication, projects, or contributor etc.
  - Lookup tables (controlled vocabulary) – Variables, units, standard etc.
- Data sources
  - CZO Observations;
  - Other published data of U.S.
  - Data from European Countries;


CZO Chemistry Database Conceptual Model – (CZchemDB)
Publication
Person contributor
Country /State
Landuse /Veg.
Project
Sources
Var-Lookup /Unit
Meta-Data
Methods
Loc_info /Climate
Geo-Info
SMPL Time Series
Lab-Info
Precision
Sub-sample
Sub-smpl 2
Sub-smpl n
Sample
Preparat. /Treatment
Lab Analysis
Chemical
Phys. Minr
Others
Main Data
Location (Watershed)
Sampling Site (Soil / Water)
Sample (Layer/Depth)
Sub-Sample
Preparation /Treatment
Analysis
Data
Slide from X. Niu, EESI, PennState

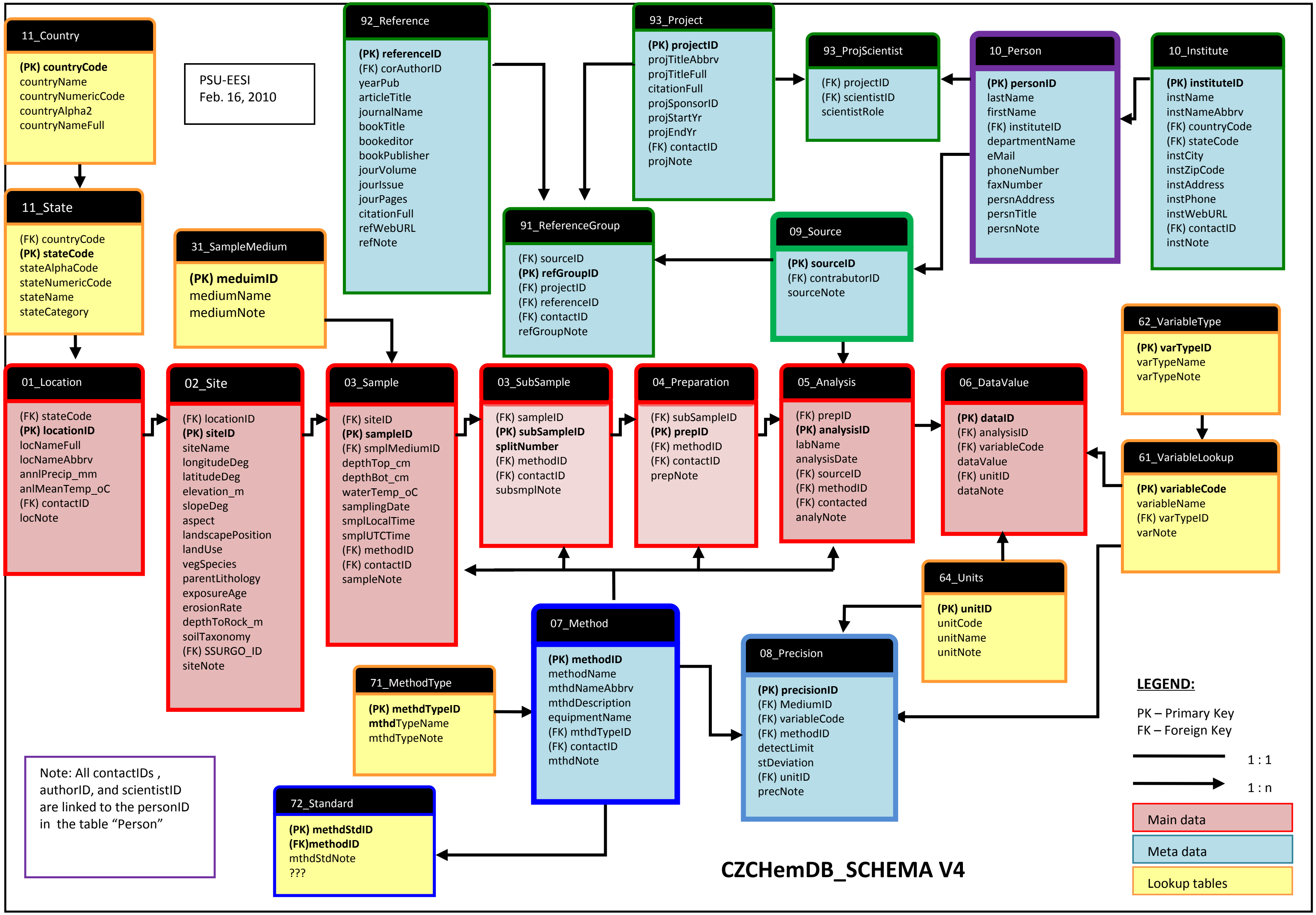


Slide from X. Niu, EESI, PennState

Slide from X. Niu, EESI, PennState

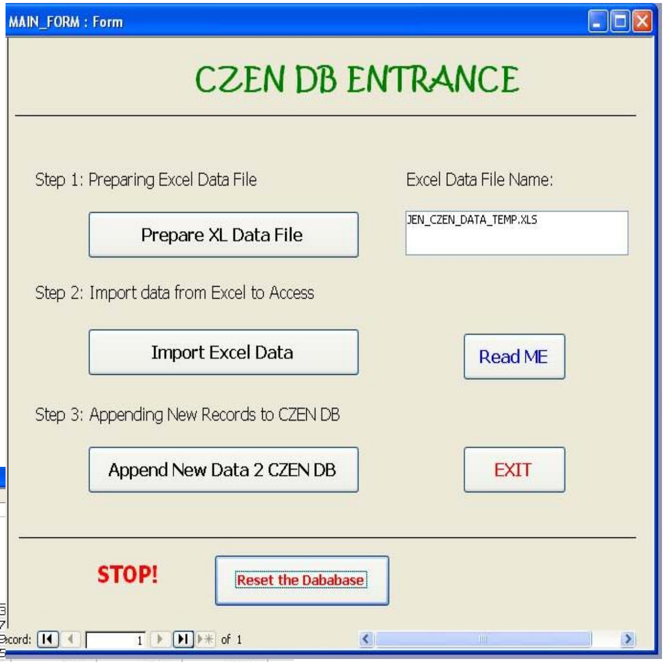
CZchemDB Schema

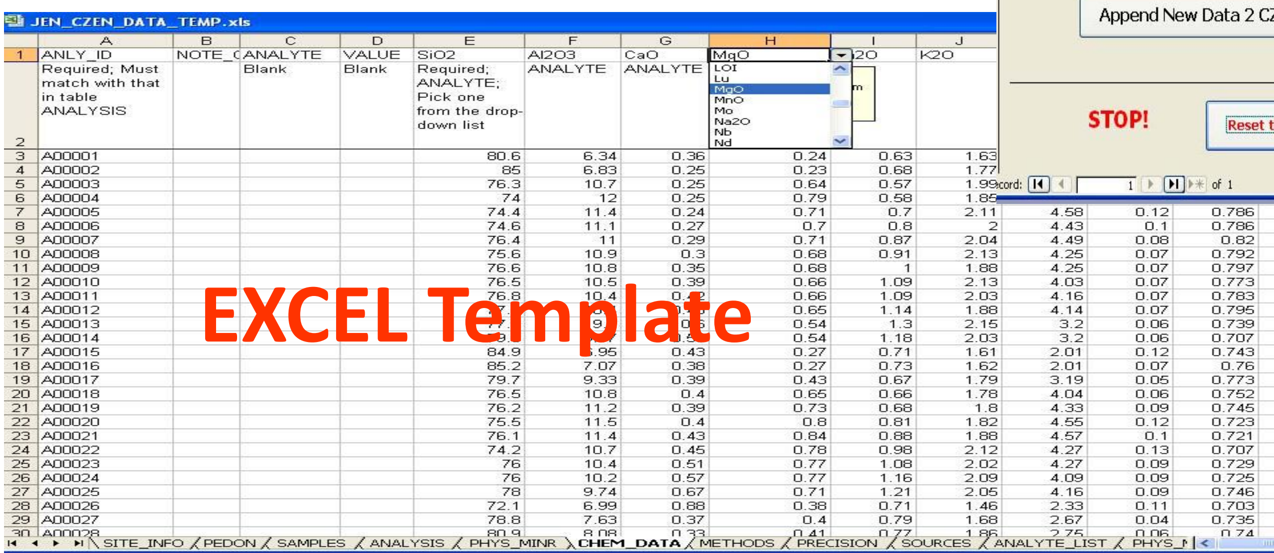
"Dentistry on the brain"



CZchemDB Data Entrance

- An Excel template file was developed for data entrance;
- Tools were developed to prepare, import, and append new data sets to the CZchemDB

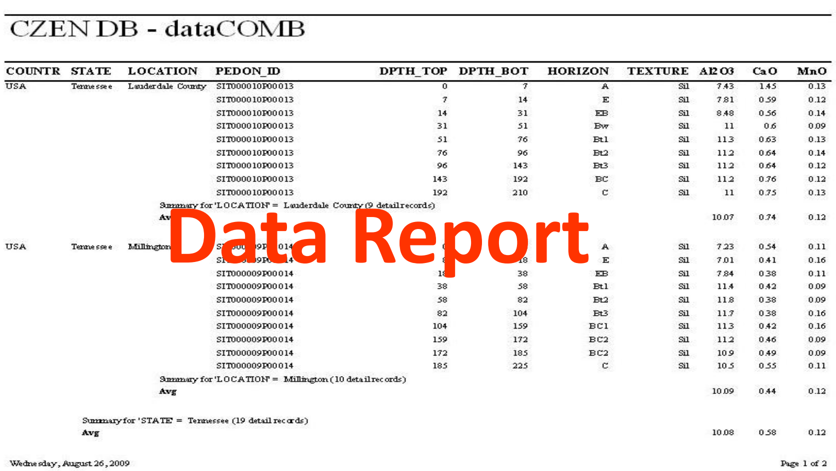




	A	B	C	D	E	F	G	H	I	J			
1	ANLY_ID	NOTE_	ANALYTE	VALUE	SiO2	Al2O3	CaO	MgO	2O	K2O			
2	Required; Must match with that in table ANALYSIS		Blank	Blank	Required; ANALYTE; Pick one from the drop-down list	ANALYTE	ANALYTE	LOI Lu MgO MnO Mo Na2O Nb Nd					
3	A00001				80.6	6.34	0.36	0.24	0.63	1.63			
4	A00002				85	6.83	0.25	0.23	0.68	1.77			
5	A00003				76.3	10.7	0.25	0.64	0.57	1.99			
6	A00004				74	12	0.25	0.79	0.58	1.85			
7	A00005				74.4	11.4	0.24	0.71	0.7	2.11	4.58	0.12	0.786
8	A00006				74.6	11.1	0.27	0.7	0.8	2	4.43	0.1	0.786
9	A00007				76.4	11	0.29	0.71	0.87	2.04	4.49	0.08	0.82
10	A00008				75.6	10.9	0.3	0.68	0.91	2.13	4.25	0.07	0.792
11	A00009				76.6	10.8	0.35	0.68	1	1.88	4.25	0.07	0.797
12	A00010				76.5	10.5	0.39	0.66	1.09	2.13	4.03	0.07	0.773
13	A00011				76.8	10.4	[illegible]	0.66	1.09	2.03	4.16	0.07	0.783
14	A00012				[illegible]	[illegible]	[illegible]	0.65	1.14	1.88	4.14	0.07	0.795
15	A00013				[illegible]	[illegible]	[illegible]	0.54	1.3	2.15	3.2	0.06	0.739
16	A00014				[illegible]	[illegible]	[illegible]	0.54	1.18	2.03	3.2	0.06	0.707
17	A00015				84.9	[illegible]	0.43	0.27	0.71	1.61	2.01	0.12	0.743
18	A00016				85.2	7.07	0.38	0.27	0.73	1.62	2.01	0.07	0.76
19	A00017				79.7	9.33	0.39	0.43	0.67	1.79	3.19	0.05	0.773
20	A00018				76.5	10.8	0.4	0.65	0.66	1.78	4.04	0.06	0.752
21	A00019				76.2	11.2	0.39	0.73	0.68	1.8	4.33	0.09	0.745
22	A00020				75.5	11.5	0.4	0.8	0.81	1.82	4.55	0.12	0.723
23	A00021				76.1	11.4	0.43	0.84	0.88	1.88	4.57	0.1	0.721
24	A00022				74.2	10.7	0.45	0.78	0.98	2.12	4.27	0.13	0.707
25	A00023				76	10.4	0.51	0.77	1.08	2.02	4.27	0.09	0.729
26	A00024				76	10.2	0.57	0.77	1.16	2.09	4.09	0.09	0.725
27	A00025				78	9.74	0.67	0.71	1.21	2.05	4.16	0.09	0.746
28	A00026				72.1	6.99	0.88	0.38	0.71	1.46	2.33	0.11	0.703
29	A00027				78.8	7.63	0.37	0.4	0.79	1.68	2.67	0.04	0.735
30	A00028				80.9	8.08	0.33	0.41	0.77	1.86	2.75	0.06	0.74

CZchemDB Application (1) – Data Searching

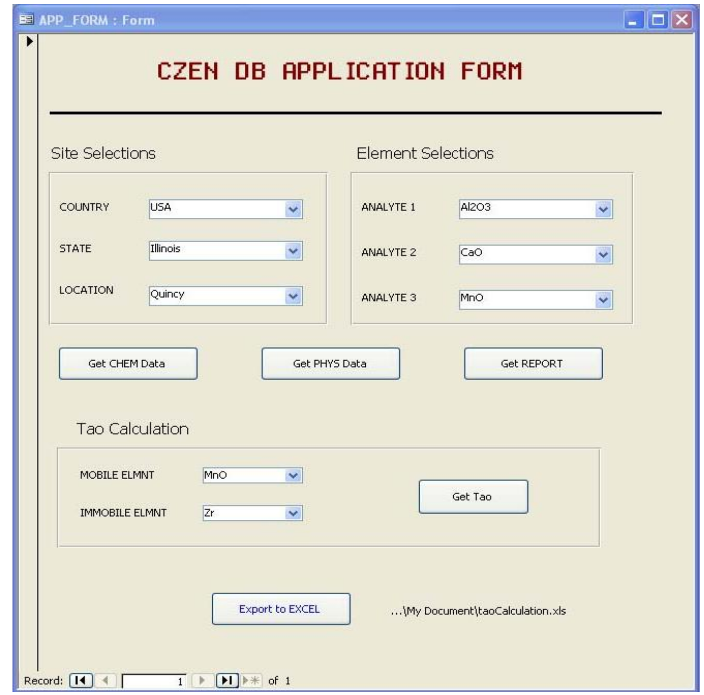
- Data searching
  - Location;
  - Geochemical element;
  - Project
- Data reporting
  - Formatted reports;
  - Excel file or text file



CZEN DB - dataCOMB

COUNTR	STATE	LOCATION	PEDON_ID	DPTH_TOP	DPTH_BOT	HORIZON	TEXTURE	Al2O3	CaO	MnO
USA	Tennessee	Lauderdale County	SIT000010P00013	0	7	A	Sil	7.43	1.45	0.13
			SIT000010P00013	7	14	E	Sil	7.81	0.59	0.12
			SIT000010P00013	14	31	EB	Sil	8.48	0.56	0.14
			SIT000010P00013	31	51	Bw	Sil	11	0.6	0.09
			SIT000010P00013	51	76	Bt1	Sil	11.3	0.63	0.13
			SIT000010P00013	76	96	Bt2	Sil	11.2	0.64	0.14
			SIT000010P00013	96	143	Bt3	Sil	11.2	0.64	0.12
			SIT000010P00013	143	192	BC	Sil	11.2	0.76	0.12
			SIT000010P00013	192	210	C	Sil	11	0.75	0.13
		Summary for 'LOCATION' = Lauderdale County (9 detail records)								
		Av[illegible]						10.07	0.74	0.12
USA	Tennessee	Millington	[illegible]	[illegible]	[illegible]	A	Sil	7.23	0.54	0.11
			[illegible]	[illegible]	[illegible]	E	Sil	7.01	0.41	0.16
			SIT000009P00014	[illegible]	38	EB	Sil	7.84	0.38	0.11
			SIT000009P00014	38	58	Bt1	Sil	11.4	0.42	0.09
			SIT000009P00014	58	82	Bt2	Sil	11.8	0.38	0.09
			SIT000009P00014	82	104	Bt3	Sil	11.7	0.38	0.16
			SIT000009P00014	104	159	BC1	Sil	11.3	0.42	0.16
			SIT000009P00014	159	172	BC2	Sil	11.2	0.46	0.09
			SIT000009P00014	172	185	BC2	Sil	10.9	0.49	0.09
			SIT000009P00014	185	225	C	Sil	10.5	0.55	0.11
		Summary for 'LOCATION' = Millington (10 detail records)								
		Avg						10.09	0.44	0.12
	Summary for 'STATE' = Tennessee (19 detail records)									
	Avg							10.08	0.58	0.12

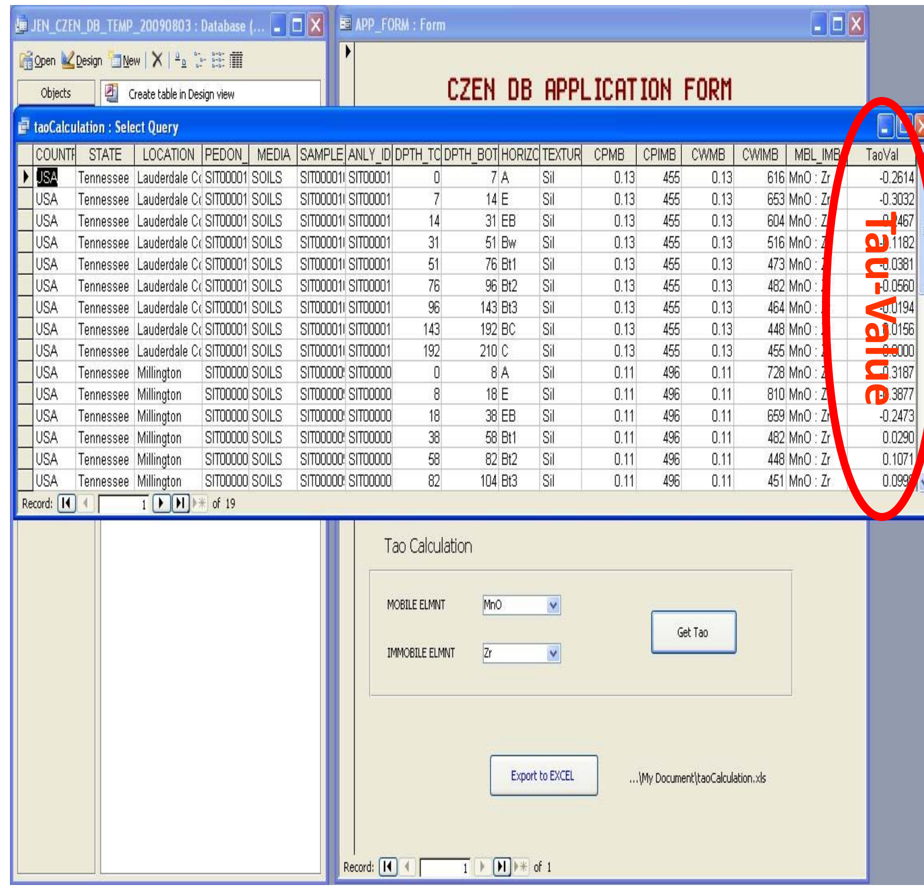
Wednesday, August 26, 2009 — Page 1 of 2



Slide from X. Niu, EESI, PennState

CZchemDB Application (2) – Data mining and modeling

- Data application
  - Tau calculations;
  - Other calculations;
- Future applications:
  - Data mining;
  - CZ modeling;



czo CHEM DB

Positives and Negatives of Online Data

Positives

- Access to data is more egalitarian
- Access to data is easier
- Sensors deployed in the environment are cheaper than human workers
- Sensors can be made to measure identically at different places for comparison
- Sensing can happen all the time
- We can measure what was unmeasurable before

Negatives

- Data is easy to mis-use: just because it is accessible doesn't mean it will be used correctly
- Sensors can replace workers
- Sensors break down or make poor measurements when they are not maintained
- Just because something can be measured (or data can be stored) doesn't mean it should be measured (stored)
- Sensing can happen all the time (loss of privacy)
- Online communications cannot be kept private
- Speed of communication becomes fast…is anyone listening?
- Huge volumes of data – storing, manipulating, understanding, and modelling such data volumes is difficult

Hypotheses: Education

- University-level education will be radically changed by the electronic revolution in the next 10 years. (*This will be disruptive unless we think about it and get out ahead of it*).
- Tenure line faculty will become increasingly involved. (*This will be for resident and non-resident students*).
- U.P. departments will accelerate in finding ways to provide online courses, online degrees, blended degrees, hybrid courses, adaptive learning. (*Even laboratories will be offered online as virtual labs and field trips as virtual trips*).

Hypotheses: Education

- Some courses can be taught better online than in person. (*If this is true, we should figure out what these are.*)
- Since quality of education often depends heavily upon the person in charge of the class, and each person has a “best mode of delivery”, the quality of education may not depend on whether the class is delivered online or in-person. (*We need to continue to promote diversity in teaching styles.*)

Hypotheses: Data

- Types and capabilities of sensors will proliferate. (*This will emphasize the need to choose what to measure as opposed to how to measure it*).
- Increasing numbers of scientists will store data and model output online. (*This will increase the importance of search strategies, data assessment strategies, and modelling strategies.*)
- Eventually we will decide that not all data belongs online and that not everything must be measured. (*We will make choices about what data to save online, and what data not to save.*)

Hypotheses: Data

- We may begin to allow students to use books/online search tools/mobile devices/etc. at all times for everything. (*Why not? It is with us always*).
- As more and more information is available online, branding of data, models, courses, etc. will become more and more important because choices will be harder. (*Your reputation, the reputation of a publication venue, and your institution's reputation will become increasingly important.*)

Hypotheses: Data

- The more metadata, the harder it is to store on line. (*Disciplines that organize their data online earlier will have a stronger voice.*)

- A need exists for the training of students in two intersections: i) sensors + domain science and ii) informatics + domain science. (*Interdisciplinary programs will self-organize to provide these*).

So what do you think we should we do?

Lawrence Lowell (when he was President of Harvard University) said*:

"…institutions are rarely murdered; they meet their end by suicide…They die because they have outlived their usefulness, or fail to do the work that the world wants done."

Thanks to Lee Kump, Seth Blumsack, Andy Nyblade, Ann Taylor, Chris Duffy, Doug Miller, Bill Brune, Bill Easterling, Karl Zimmerer, others.

ACKNOWLEDGEMENTS

*As quoted in *The Economist* in the Schumpeter blog, Dec 10 2011)