The BioPortal© global infectious disease information-sharing infrastructure and Cal-X (the California Exchange)

Michael S. Ascher, M.D., FACP

Department of Medicine and Epidemiology
School of Veterinary Medicine
UC Davis
and
California Emergency Management Agency

BioPortal background

- The problem is the lack of a (inter)national system for secure, real-time acquisition and sharing of infectious disease information
- Absolutely critical for the management of a large-scale natural or unnatural outbreak
- Central model (“send us the data” or “use our software”) does not meet local and state needs for information
- In spite of lots of $$ and talk, very little progress has been made in the last ten years to rectify this situation

BioPortal background (2)

- In the fall of 2002, an interagency working group on genomics was formed, chaired by the Director of NSF
- The Infectious Disease Informatics Working (sub)Group (IDIWG) was established to develop specifications of a national infectious disease informatics intrastructure (NIDII)
- In June 2003, the group requested that the subgroup develop, by Christmas, “one or more rapid prototype systems to demonstrate interoperability and innovation across species and jurisdictions” following NIDII specs.

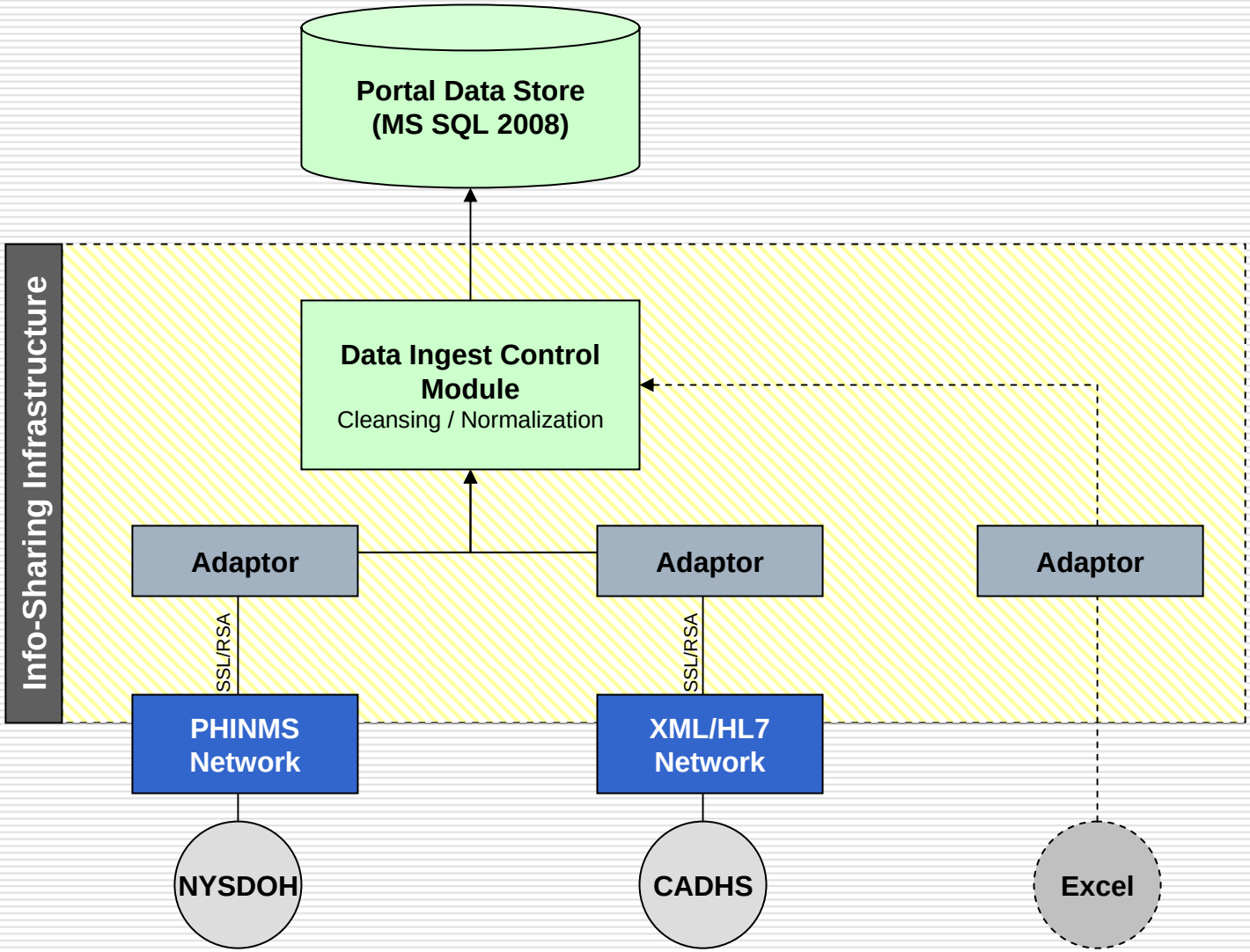
BioPortal background (3)

- Botulism and West Nile virus infection selected as diseases
- New York State and California public health departments selected as partners
- University of Arizona selected as integrator and given supplement to existing COPLINK grant from NSF
- USGS Federal partner

NIDII specifications

- Human and animal data
- Multiple jurisdictions
- Data integration and case management
- Multi-way data sharing
- Data access through “peeks”
- Analytical output
  - Visualization (time and space)
  - Anomaly detection
- Consortium governance
- Federal partners
- International data
- Genomics information

Information Sharing Infrastructure Design



Data Access Infrastructure Design

Browser (IE/Mozilla/…)

Public health professionals, researchers, policy makers, law enforcement agencies & other users

SSL connection

WNV-BOT Portal

Data Search and Query

Spatial - Temporal Visualization

Analysis / Prediction

HAN or Personal Alert Management

Dataset Privileges Management

Web Server (Tomcat 4.21 / Struts 1.2)

Data Store

User Access Control API (Java)

Access Privilege Definition

Data Store (MS SQL 2008)

Bio Portal - Windows Internet Explorer provided by Your Friendly Neighborhood IT

http://www.bioportal.org/

Favorites | NYT | GNews | Citrix | Comcast | UCMail | SuDoku | Amtrak | Live Meeting | Schwab | SharePoint | UCD Lib | USAA | Jumble | CalEMA

Bio Portal

Page | Safety | Tools

BioPortal

BIOPORTAL

Login:

Password:

Login Clear

This site is for public demo purposes. Please use following accounts to login:

- General Access: tEstaCCount/ nY_Ca_Ua
- Detail Access: Sandy/ Sandy
- Aggregated Access: Jack/ Jack

http://www.bioportal.org

Internet | 100%

BioPortal | Artificial Intelligence Lab

The University of Arizona

No new alerts

- About Us
- About the BioPortal
- About the Data
- Publications
- Logged in As: Sandy
- Logout
- New Data Selection
- Help Files
- How to Select Data
- Hotspot Analysis
- STV Result Reviewer
- New Features

Step 1 | Step 2 | Step 3 | Step 4 | Step 5

Begin Search

Choose below the disease databases that you can search and visualize using this portal. Data selection is a 5-step process:

1. Select disease database
2. Select data sets within the database
    - (optional) Advanced: Select subsets of records from each data set (e.g., by species in animal data sets, by type of botulism in human botulism data sets)
3. Select time range and location of data
4. View list of individual records you've chosen and select option to view map and graphs of data in SpatioTemporal Visualizer screen
5. Select map layers (e.g., precipitation data or population data for the area) you'd like to see in the data map

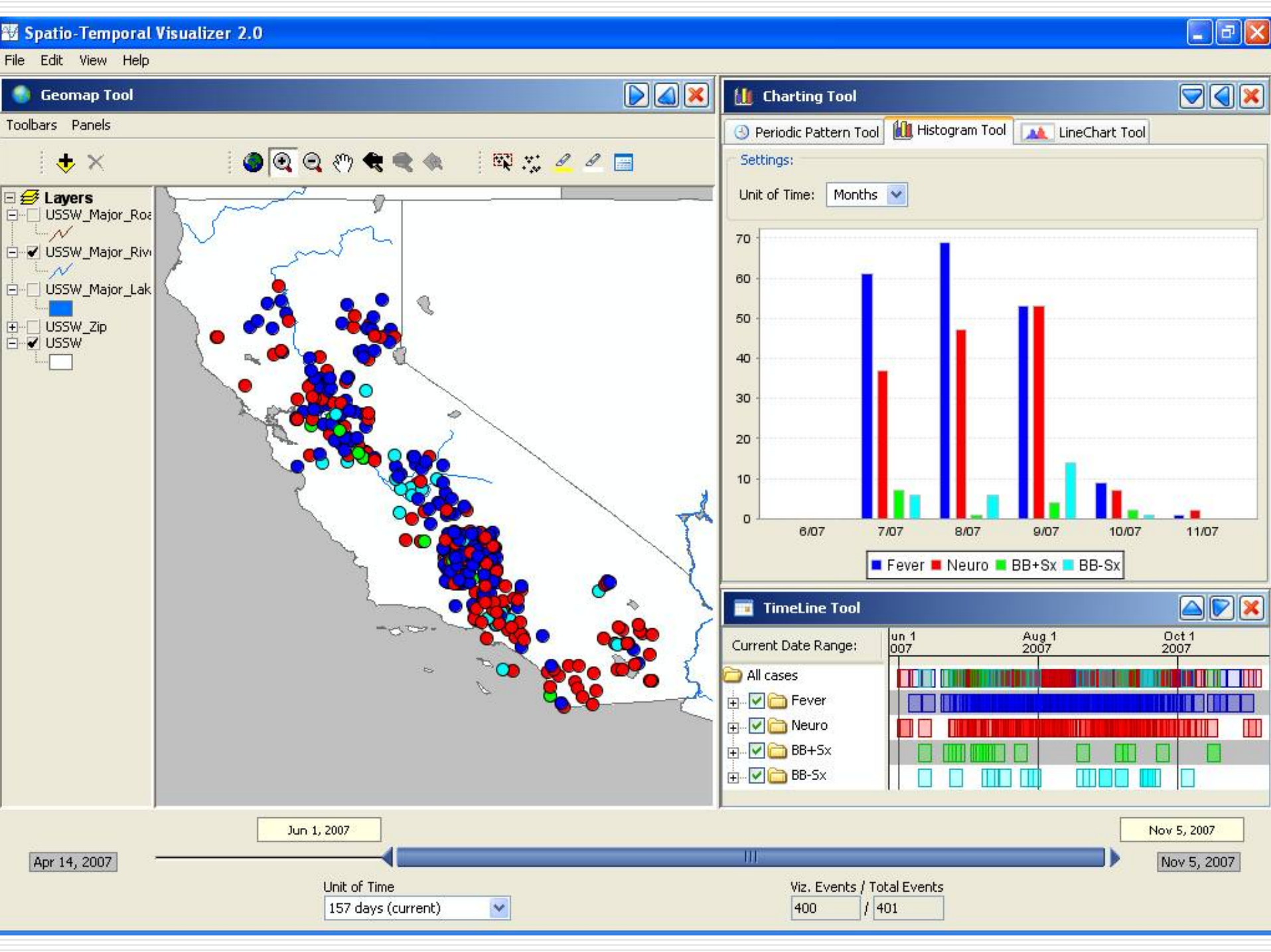
Please choose one of the following to start your search:

West Nile Virus

Botulism

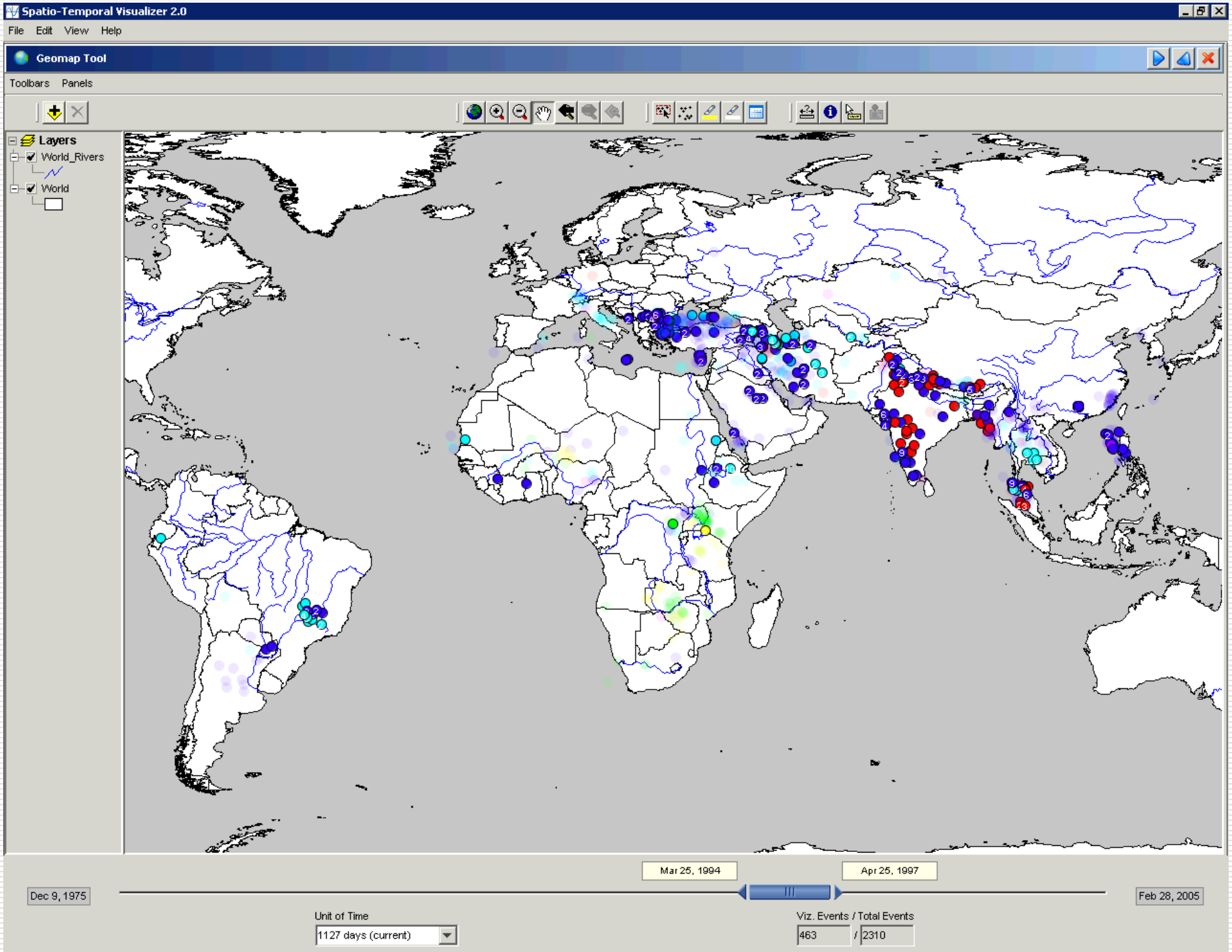
Foot & Mouth Disease

BioWatch


Spatio-Temporal Visualizer 2.0
File
Edit
View
Help
Geomap Tool
Toolbars
Panels
Layers
USSW_Major_Roa
USSW_Major_Riv
USSW_Major_Lak
USSW_Zip
USSW
Charting Tool
Periodic Pattern Tool
Histogram Tool
LineChart Tool
Settings:
Unit of Time:
Months
70
60
50
40
30
20
10
0
6/07
7/07
8/07
9/07
10/07
11/07
Fever
Neuro
BB+Sx
BB-Sx
TimeLine Tool
Current Date Range:
un 1
007
Aug 1
2007
Oct 1
2007
All cases
Fever
Neuro
BB+Sx
BB-Sx
Jun 1, 2007
Nov 5, 2007
Apr 14, 2007
Nov 5, 2007
Unit of Time
157 days (current)
Viz. Events / Total Events
400
/
401

Synthesis of FMD and BioPortal©

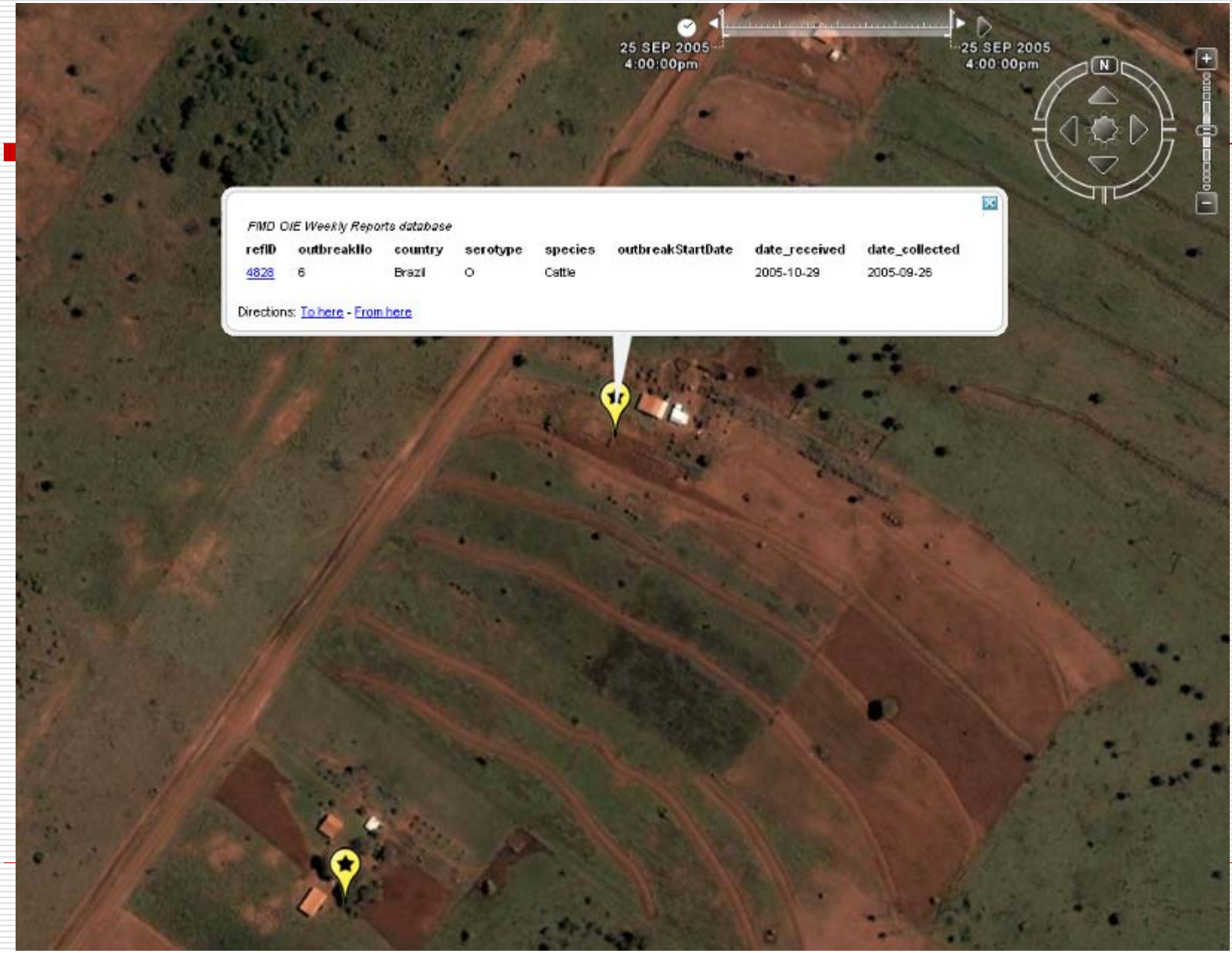
- Animal disease
- International scope
- Genomics component
  - Real-time sequencing


Spatio-Temporal Visualizer 2.0
File
Edit
View
Help
Geomap Tool
Toolbars
Panels
Layers
World_Rivers
World
Mar 25, 1994
Apr 25, 1997
Dec 9, 1975
Feb 28, 2005
Unit of Time
1127 days (current)
Viz. Events / Total Events
463
/ 2310

FMD BioPortal link to Google Earth

OIE Weekly FMD Reports - Brazil

OIE Weekly FMD Reports - Brazil



FMDv genomic/epidemiologic database

Sources: GenBank and FMD literature

Serotype	Isolates
A	585
C	133
O	1046
Asia1	316
SAT1	258
SAT2	310
SAT3	81
TOTAL	2729

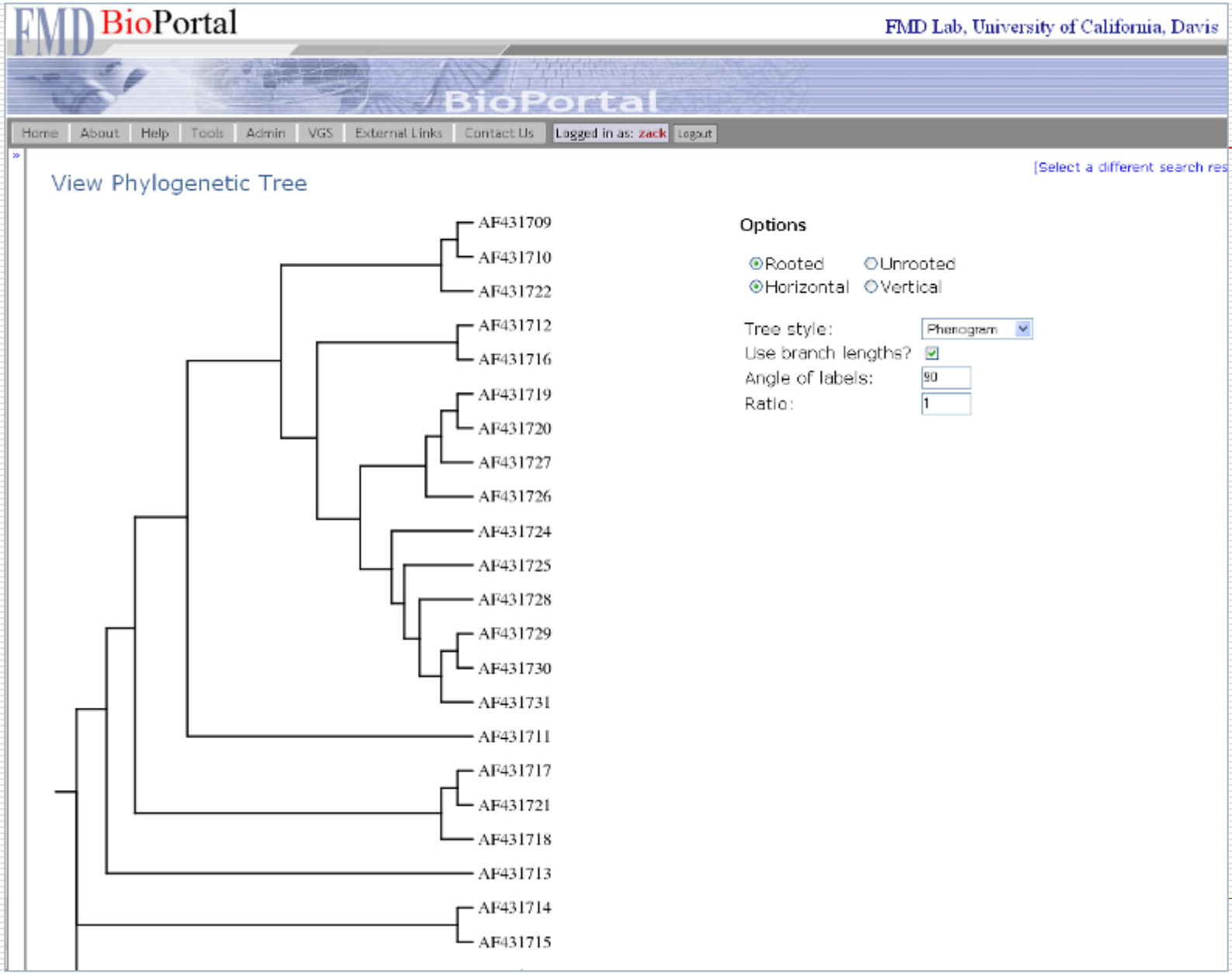
All the Nucleotide sequences in genbank as of july 26/2006
NOTE: NUCLEOTIDE SEQUENCES FILE: Type C nucleotide seq.fas(70 SEQUENCES)
PROTEIN SEQUENCES FILE : TYPE C PROTEIN SEQ.FAS(64 SEQUENCES)

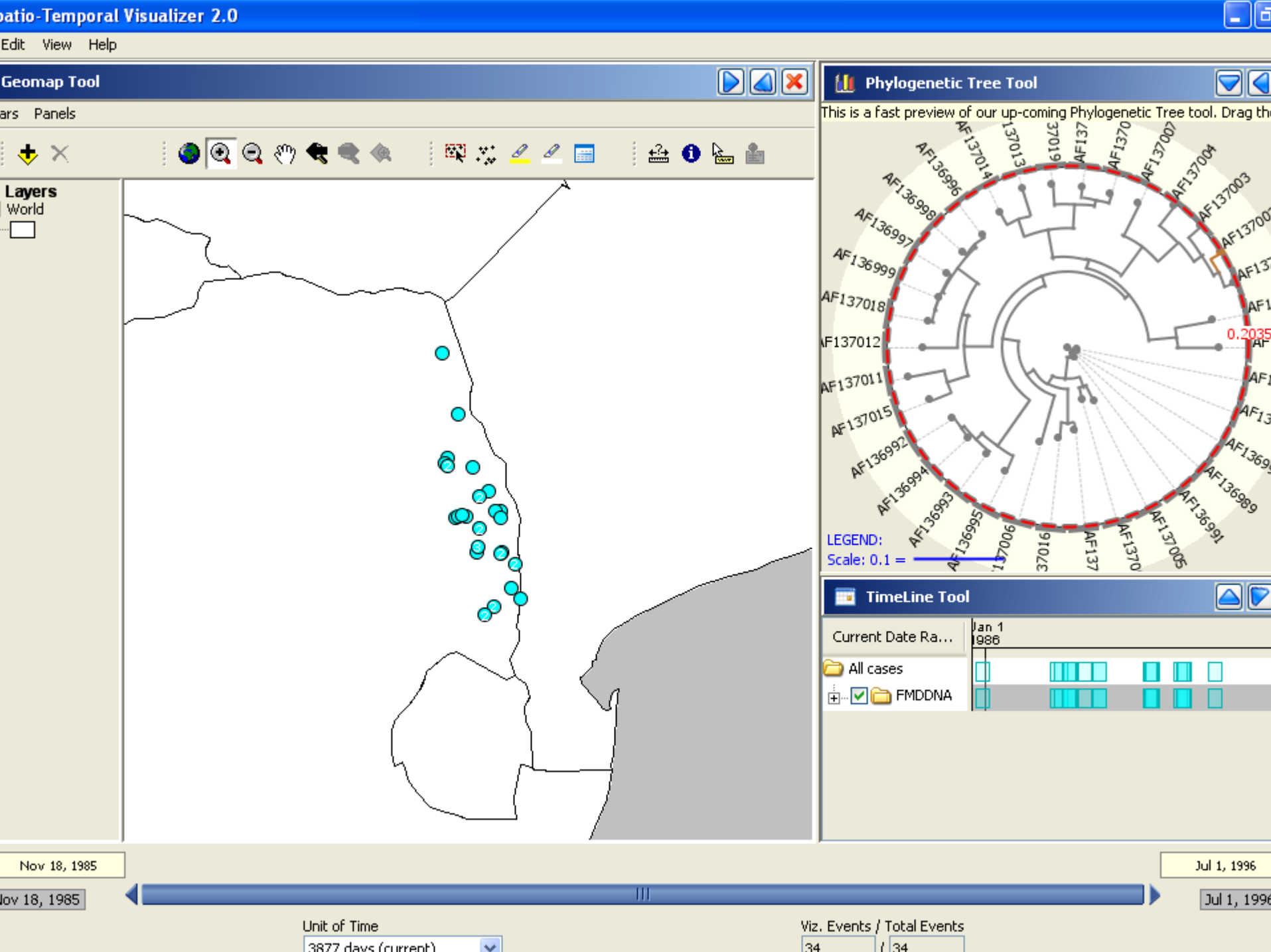
Accession #	Protein GI #	Sequence	Strain	Isolate	Isolation I	Submitted	size	Country	Species	Vaccine Strain	Artificially attenuated
AF207522	11493910	Pro L gene	C/Bombay/64	na	1964	19-Nov-99	603 bp	India: Bombay	na	na	na
AF274010	10334812	polyprotein	C-S8 clone MARLS	MARLS	na	31-May-00	8115 bp	Spain	na	na	na
AF283451	15419318	Pro L gene	na	C1Noville	1965	29-Jun-00	639 bp	Switzerland	na	na	na
AF283452	15419320	Pro L gene	na	C3Resende	1965	29-Jun-00	639 bp	Brazil	Bovine	na	na
AF378301	21434559	Pro L gene	na	C3 Resende	1988	9-May-01	615 bp	Brazil	Cattle	na	na
AF536536	22770784	3D pol	na	Noville/Switzerland/65	1965	8-Aug-02	383 bp	Switzerland	na	na	na
AJ133357	6318188	polyprotein	C	C-s8c1	na	26-Feb-99	8115 bp	Spain	na	na	na
AJ133358	6318190	polyprotein	C	rp99	na	26-Feb-99	8115 bp	Spain	na	na	na
AJ133359	6318192	polyprotein	C	rp146	na	26-Feb-99	8115 bp	Spain	na	na	na
AJ306213	14330393	VP1 (1D) gene	Gral. Lamadrid/Arg/93		1993	24-Jan-01	630 bp	Argentina:Buenos Aires	na	na	na
AJ306214	14330395	VP1 (1D) gene	Gral. Roca/Arg/02/93	Gral. Roca	1993	24-Jan-01	630 bp	Argentina:Cordoba	na	na	na
AJ306215	14330397	VP1 (1D) gene	Gral. Roca/Arg/03/93	Gral. Roca	1993	24-Jan-01	429 bp	Argentina:Cordoba	na	na	na
AJ306216	14330399	VP1 (1D) gene	Rivadavia/Arg/93	Rivadavia	1993	24-Jan-01	543 bp	Argentina:Buenos Aires	na	na	na
AJ306217	14330401	VP1 (1D) gene	Gral. Villegas/Arg/93	Gral. Villegas	1993	24-Jan-01	630 bp	Argentina:Buenos Aires	na	na	na
AJ306218	14330403	VP1 (1D) gene	San Cristobal/Arg/94	San Cristobal	1994	24-Jan-01	516 bp	Argentina:Santa Fe	na	na	na
AJ308703	15130872	VP1 (1D) gene	C3/Salto/Arg/91	Salto	1991	9-Mar-01	630 bp	Argentina:Buenos Aires	na	na	na
AJ308704	15130874	VP1 (1D) gene	C3/San Antonio de Giles/Ar	San Antonio de Giles	1992	9-Mar-01	603 bp	Argentina:Buenos Air	na	na	na
AY026896	17223558	3A	Resende/Brazil/55	Resende/Brazil/55	1955	2-Feb-01	459 bp	Brazil	na	YES	na
AY460598	38606175	VP1 (1D) gene	C/IND/51/79	IND/51/79	1979	7-Nov-03	243 bp	India: Tamilnadu, Ramnathapura	na	na	na
AY460599	38606177	VP1 (1D) gene	C/APN/33/93	APN/33/93	1993	7-Nov-03	243 bp	India: Andhra Pradesh, Nizamat	cattle	na	na
AY460600	38606179	VP1 (1D) gene	C/KAB/10/94	KAB/10/94	1994	7-Nov-03	243 bp	India: Karnataka, Bangalore	cattle	na	na
AY460601	38606181	VP1 (1D) gene	C/KAG/13/93	KAG/13/93	1993	7-Nov-03	243 bp	India: Karnataka, Gulbarga	cattle	na	na
AY460602	38606183	VP1 (1D) gene	C/MAD/31/91	MAD/31/91	1991	7-Nov-03	243 bp	India: Maharashtra, Dhule	cattle	na	na
AY460603	38606185	VP1 (1D) gene	C/MAT/08/93	MAT/08/93	1993	7-Nov-03	243 bp	India: Maharashtra, Th	cattle	na	na
AY460604	38606187	VP1 (1D) gene	C/MAW/33/91	MAW/33/91	1991	7-Nov-03	243 bp	India: Maharashtra, Wardha	cattle	na	na
AY460605	38606189	VP1 (1D) gene	C/MPH/12/92	MPH/12/92	1992	7-Nov-03	243 bp	India: Madhya Pradesh:Hosanga	cattle	na	na
AY460606	38606191	VP1 (1D) gene	C/MPU/32/91	MPU/32/91	1991	7-Nov-03	243 bp	India: Madhya Pradesh, Ujja	cattle	na	na
AY460607	38606193	VP1 (1D) gene	C/TNP/32/93	TNC/32/93	1993	7-Nov-03	243 bp	India: Tamilnadu, Periyar	cattle	na	na
AY460608	38606195	VP1 (1D) gene	C/WBN/04/91	WBN/04/91	1991	7-Nov-03	243 bp	India: West Bengal, Nadia	cattle	na	na

ALL SEQUENCES | Complete Genomes | Capsid protein | VP1 | S Gene | 3D pol | 3A | 3C | Pro L gene

Ready

Phylogenetic Tree – FMD serotype SAT1 VP1 Sequences



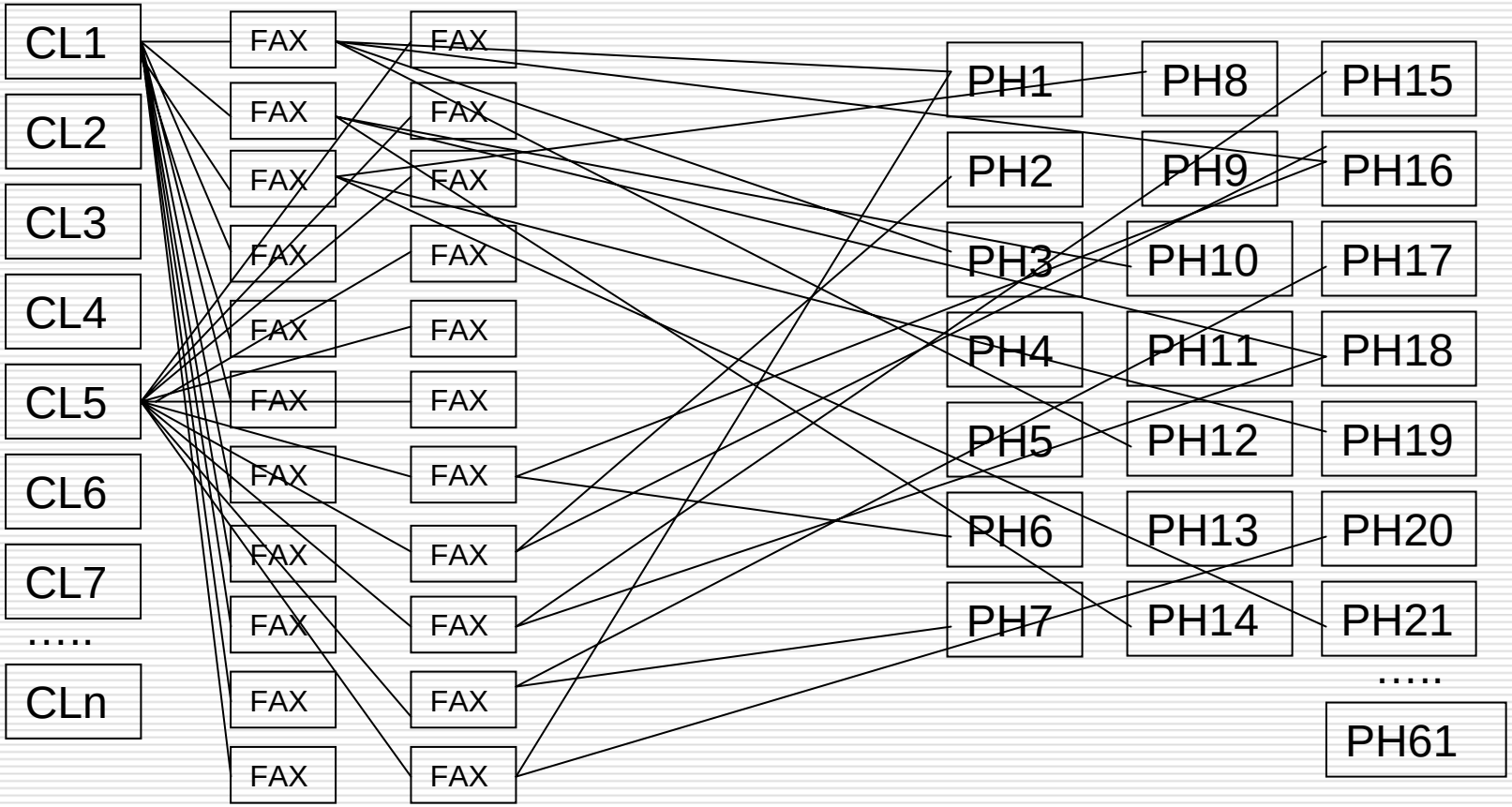

patio-Temporal Visualizer 2.0
Edit View Help
Geomap Tool
ars Panels
Layers
World
Phylogenetic Tree Tool
This is a fast preview of our up-coming Phylogenetic Tree tool. Drag th
AF137014
AF137013
AF137019
AF137
AF1370
AF137007
AF137004
AF137003
AF136996
AF136998
AF136997
AF136999
AF137018
AF137012
AF137011
AF137015
AF136992
AF136994
AF136993
AF136995
AF137016
AF137005
AF136991
AF136989
0.2035
LEGEND:
Scale: 0.1 =
TimeLine Tool
Current Date Ra...
Jan 1
1986
All cases
FMDDNA
Nov 18, 1985
Jul 1, 1996
Unit of Time
3877 days (current)
Viz. Events / Total Events
34
34

Genesis of Cal-X

- Presented BioPortal to public health laboratory directors
- Lots of interest – “no data”
- 80 diseases are “laboratory-reportable” to local HD’s
- How do they do it now?

Commercial labs

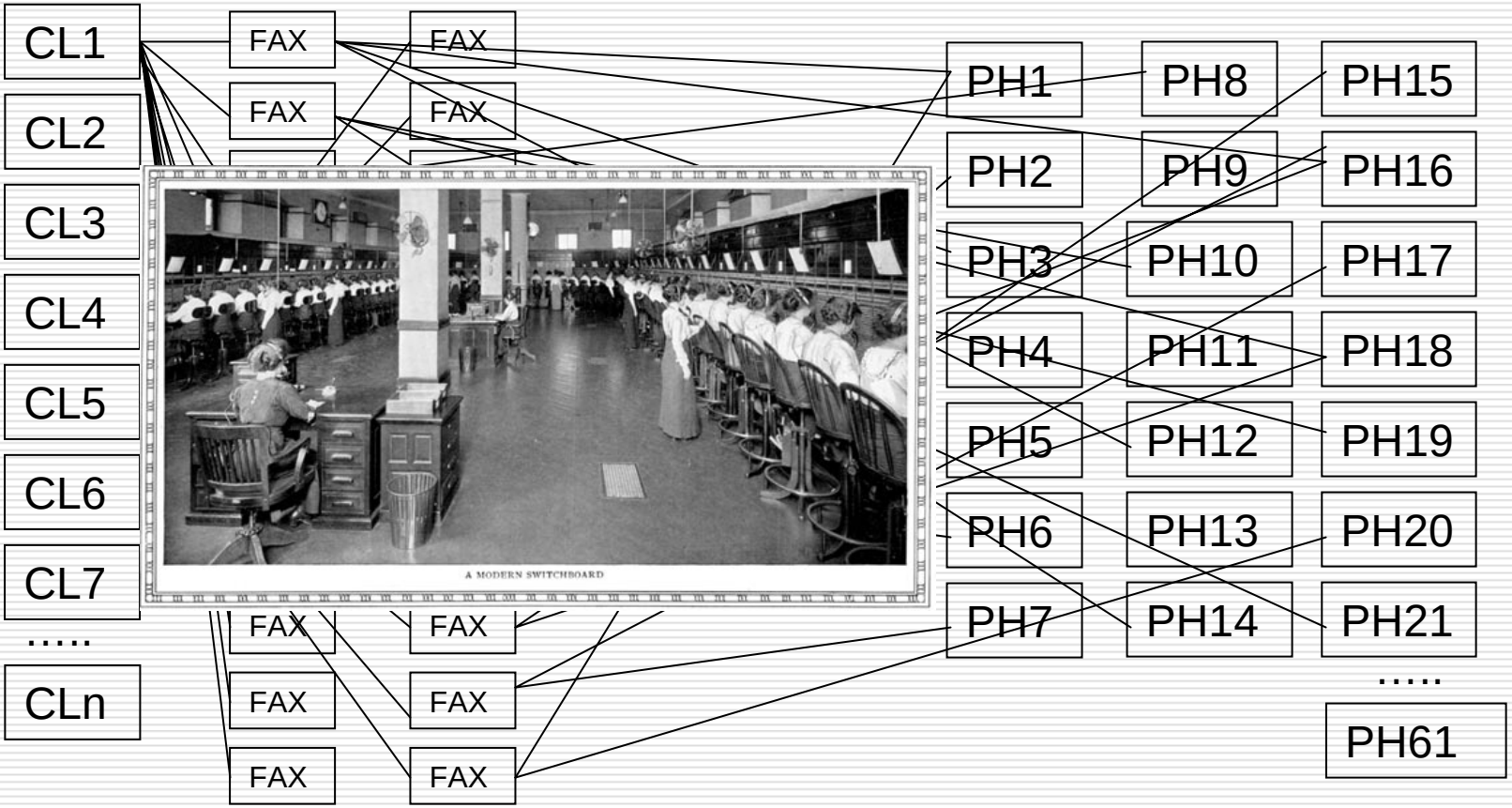
Public health jurisdictions



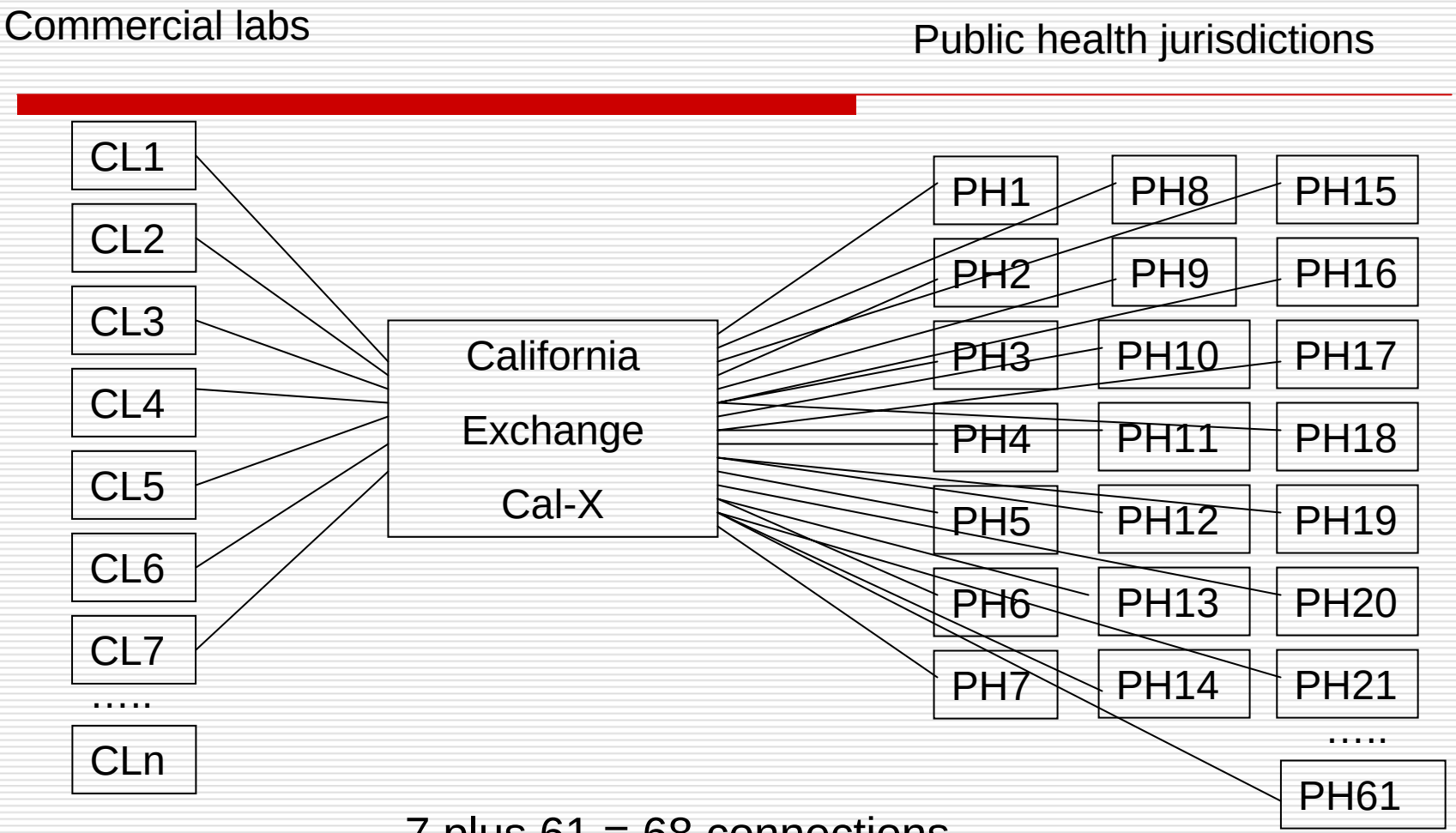
7 x 30 + 7 x 61 = 637 connections!

Commercial labs

Public health jurisdictions



7 x 30 + 7 x 61 = 637 connections!



7 plus 61 = 68 connections

Commercial labs

Public health jurisdictions

CL1
CL2
CL3
CL4
CL5
CL6
CL7
…..
CLn

PH1 PH8 PH15
PH2 PH9 PH16
PH3 PH10 PH17
PH4 PH11 PH18
PH5 PH12 PH19
PH6 PH13 PH20
PH7 PH14 PH21
…..
PH61

7 plus 61 = 68 connections

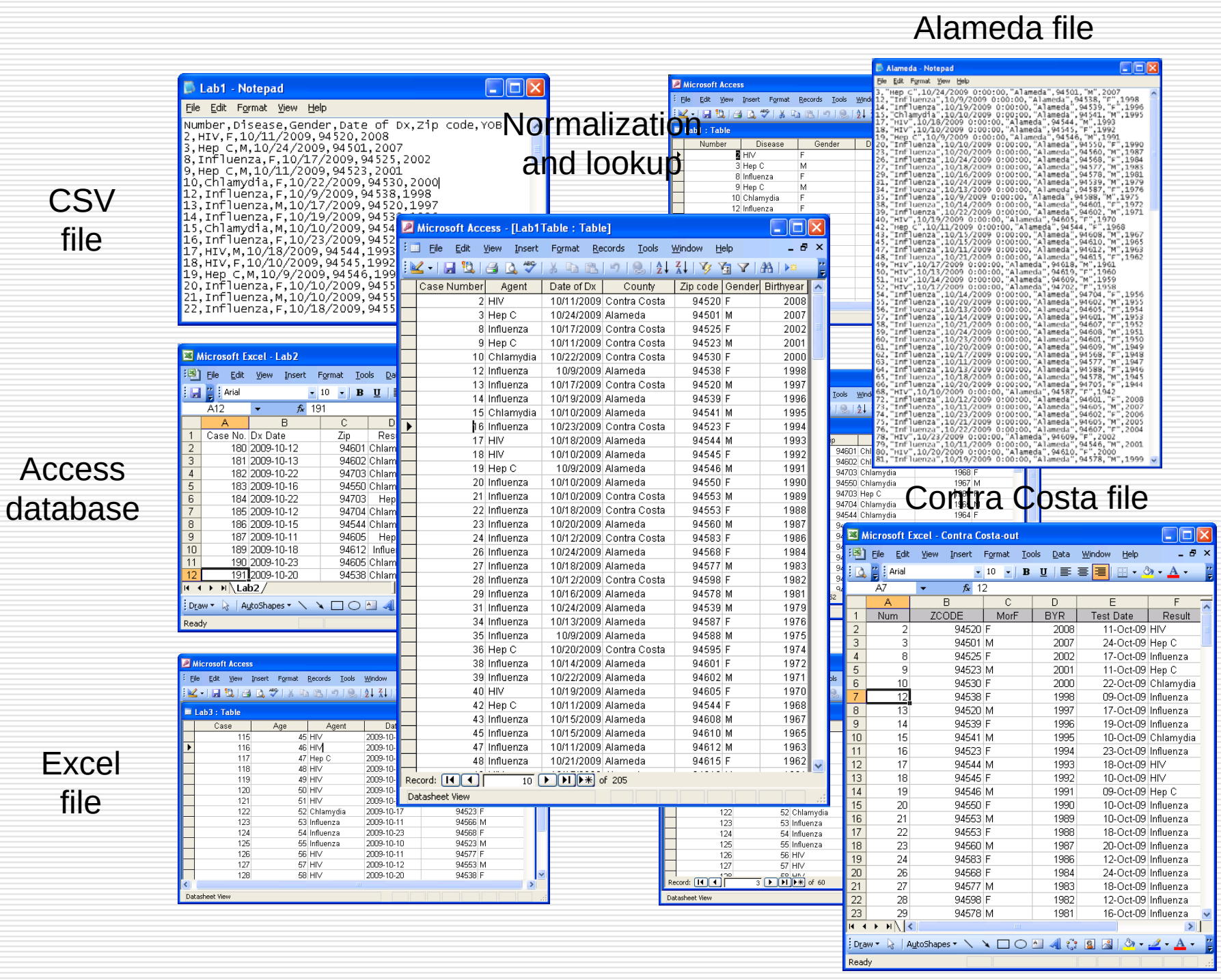
Genesis of Cal-X (2)

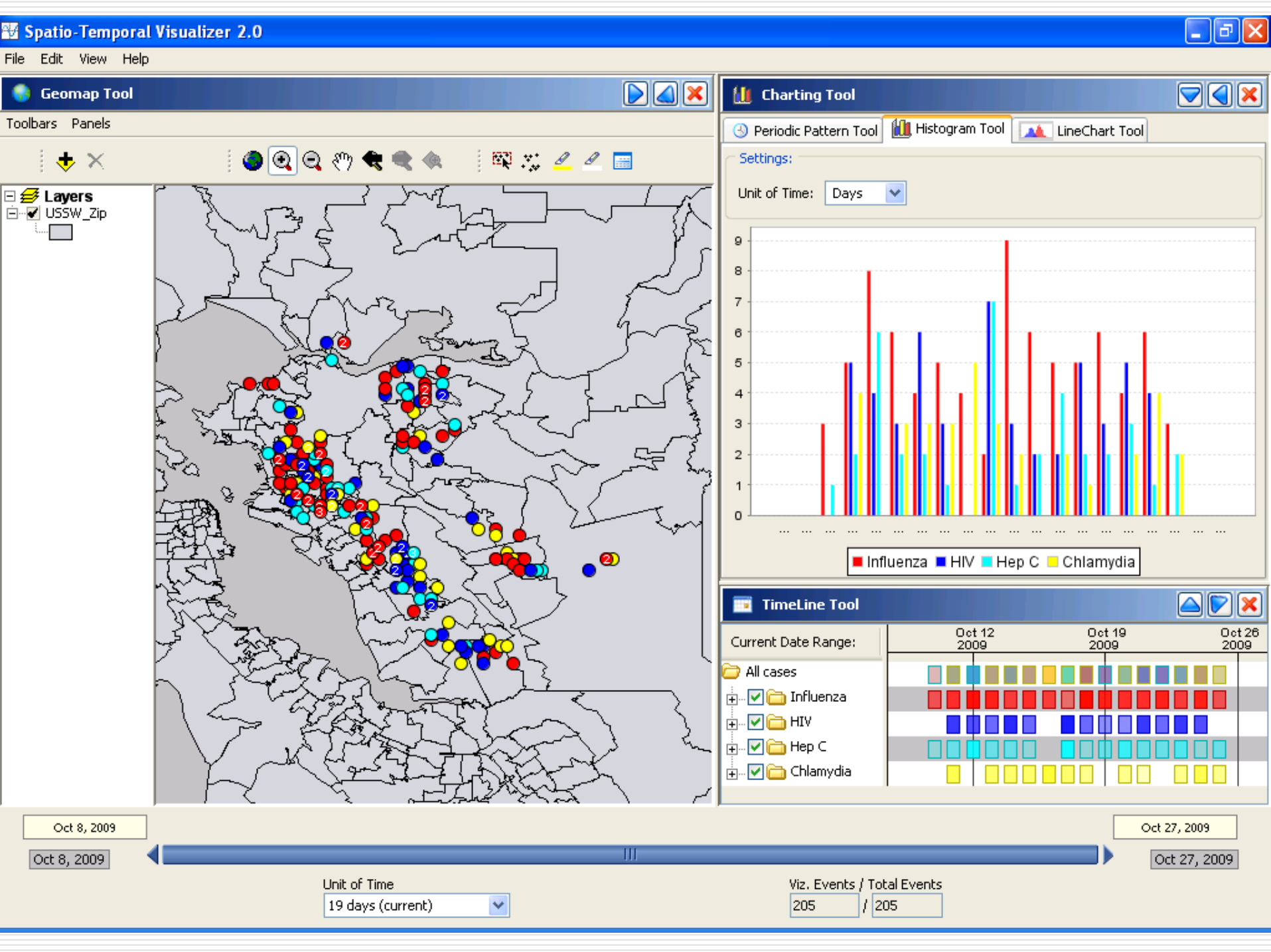
- Public health laboratory directors funded proof-of-concept of Cal-X
- Started with reportable diseases

- Received CalEMA grant to extend to medical surge and all-hazard situational awareness

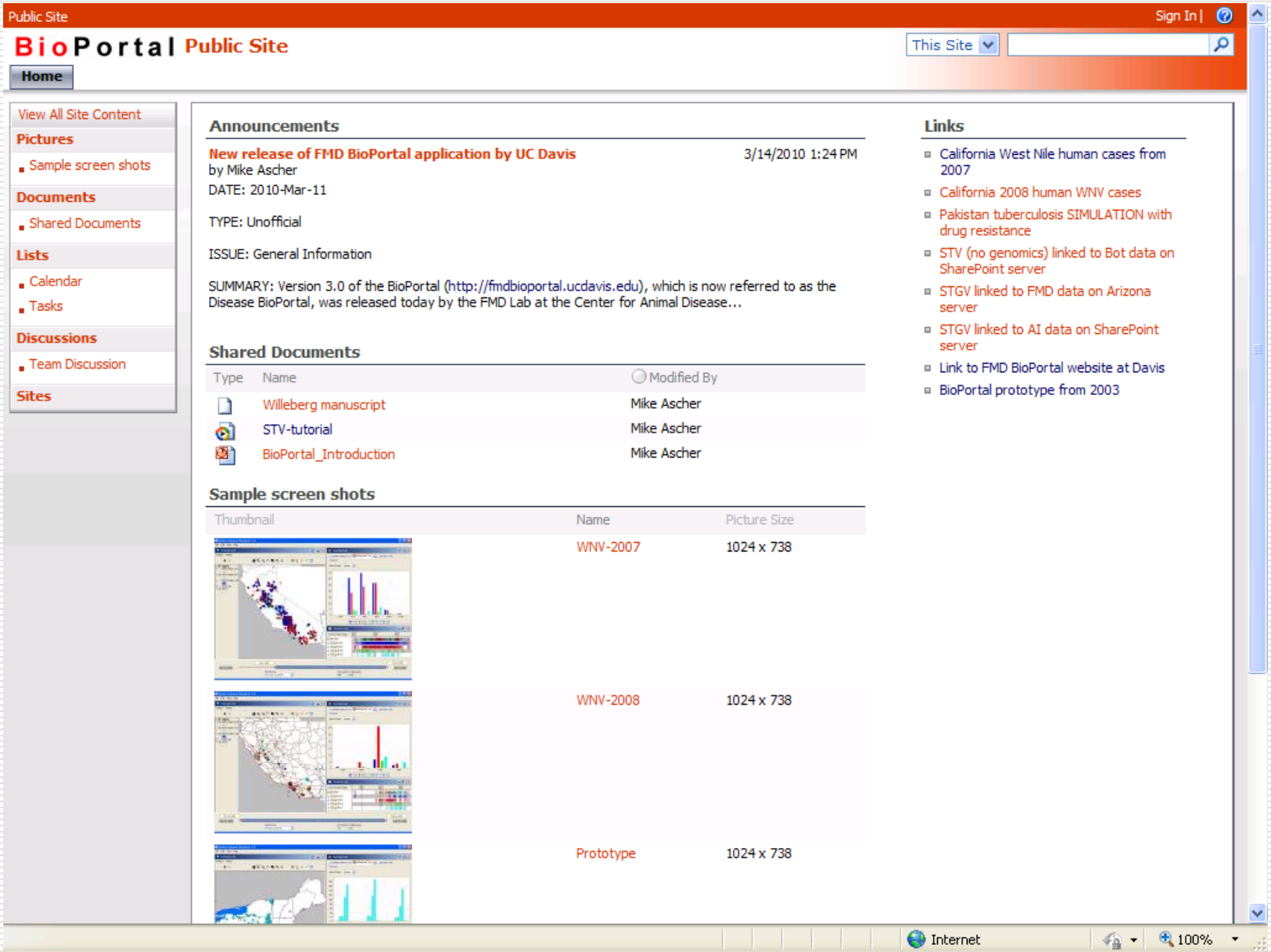
Working example of Cal-X process

- 205 cases of reportable infections
- Divided into three files representing different labs and data formats
- Translation to common format
- Normalization to map common fields
- Output to two counties in different formats


Alameda file
Normalization and lookup
CSV file
Access database
Excel file
Contra Costa file
Lab1 - Notepad
File Edit Format View Help
Number,Disease,Gender,Date of Dx,Zip code,YOB
2,HIV,F,10/11/2009,94520,2008
3,Hep C,M,10/24/2009,94501,2007
8,Influenza,F,10/17/2009,94525,2002
9,Hep C,M,10/11/2009,94523,2001
10,Chlamydia,F,10/22/2009,94530,2000
12,Influenza,F,10/9/2009,94538,1998
13,Influenza,M,10/17/2009,94520,1997
14,Influenza,F,10/19/2009,94539,1996
15,Chlamydia,M,10/10/2009,94541,1995
16,Influenza,F,10/23/2009,94523,1994
17,HIV,M,10/18/2009,94544,1993
18,HIV,F,10/10/2009,94545,1992
19,Hep C,M,10/9/2009,94546,199
20,Influenza,F,10/10/2009,9455
21,Influenza,M,10/10/2009,9455
22,Influenza,F,10/18/2009,9455
Alameda - Notepad
Microsoft Access - [Lab1Table : Table]
Case Number Agent Date of Dx County Zip code Gender Birthyear
2 HIV 10/11/2009 Contra Costa 94520 F 2008
3 Hep C 10/24/2009 Alameda 94501 M 2007
8 Influenza 10/17/2009 Contra Costa 94525 F 2002
9 Hep C 10/11/2009 Contra Costa 94523 M 2001
10 Chlamydia 10/22/2009 Contra Costa 94530 F 2000
12 Influenza 10/9/2009 Alameda 94538 F 1998
13 Influenza 10/17/2009 Contra Costa 94520 M 1997
14 Influenza 10/19/2009 Alameda 94539 F 1996
15 Chlamydia 10/10/2009 Alameda 94541 M 1995
16 Influenza 10/23/2009 Contra Costa 94523 F 1994
17 HIV 10/18/2009 Alameda 94544 M 1993
18 HIV 10/10/2009 Alameda 94545 F 1992
19 Hep C 10/9/2009 Alameda 94546 M 1991
20 Influenza 10/10/2009 Alameda 94550 F 1990
21 Influenza 10/10/2009 Contra Costa 94553 M 1989
22 Influenza 10/18/2009 Contra Costa 94553 F 1988
23 Influenza 10/20/2009 Alameda 94560 M 1987
24 Influenza 10/12/2009 Contra Costa 94583 F 1986
26 Influenza 10/24/2009 Alameda 94568 F 1984
27 Influenza 10/18/2009 Alameda 94577 M 1983
28 Influenza 10/12/2009 Contra Costa 94598 F 1982
29 Influenza 10/16/2009 Alameda 94578 M 1981
31 Influenza 10/24/2009 Alameda 94539 M 1979
34 Influenza 10/13/2009 Alameda 94587 F 1976
35 Influenza 10/9/2009 Alameda 94588 M 1975
36 Hep C 10/20/2009 Contra Costa 94595 F 1974
38 Influenza 10/14/2009 Alameda 94601 F 1972
39 Influenza 10/22/2009 Alameda 94602 M 1971
40 HIV 10/19/2009 Alameda 94605 F 1970
42 Hep C 10/11/2009 Alameda 94544 F 1968
43 Influenza 10/15/2009 Alameda 94608 M 1967
45 Influenza 10/15/2009 Alameda 94610 M 1965
47 Influenza 10/11/2009 Alameda 94612 M 1963
48 Influenza 10/21/2009 Alameda 94615 F 1962
Record: 10 of 205
Datasheet View
Microsoft Excel - Lab2
A12 191
Case No. Dx Date Zip
180 2009-10-12 94601 Chlam
181 2009-10-13 94602 Chlam
182 2009-10-22 94703 Chlam
183 2009-10-16 94550 Chlam
184 2009-10-22 94703 Hep
185 2009-10-12 94704 Chlam
186 2009-10-15 94544 Chlam
187 2009-10-11 94605 Hep
189 2009-10-18 94612 Influe
190 2009-10-23 94605 Chlam
191 2009-10-20 94538 Chlam
Lab2
Ready
Microsoft Access
Lab3 : Table
Case Age Agent
115 45 HIV
116 46 HIV
117 47 Hep C
118 48 HIV
119 49 HIV
120 50 HIV
121 51 HIV
122 52 Chlamydia 2009-10-17 94523 F
123 53 Influenza 2009-10-11 94566 M
124 54 Influenza 2009-10-23 94568 F
125 55 Influenza 2009-10-10 94523 M
126 56 HIV 2009-10-11 94577 F
127 57 HIV 2009-10-12 94553 M
128 58 HIV 2009-10-20 94538 F
Datasheet View
Microsoft Excel - Contra Costa-out
A7 12
Num ZCODE MorF BYR Test Date Result
2 94520 F 2008 11-Oct-09 HIV
3 94501 M 2007 24-Oct-09 Hep C
8 94525 F 2002 17-Oct-09 Influenza
9 94523 M 2001 11-Oct-09 Hep C
10 94530 F 2000 22-Oct-09 Chlamydia
12 94538 F 1998 09-Oct-09 Influenza
13 94520 M 1997 17-Oct-09 Influenza
14 94539 F 1996 19-Oct-09 Influenza
15 94541 M 1995 10-Oct-09 Chlamydia
16 94523 F 1994 23-Oct-09 Influenza
17 94544 M 1993 18-Oct-09 HIV
18 94545 F 1992 10-Oct-09 HIV
19 94546 M 1991 09-Oct-09 Hep C
20 94550 F 1990 10-Oct-09 Influenza
21 94553 M 1989 10-Oct-09 Influenza
22 94553 F 1988 18-Oct-09 Influenza
23 94560 M 1987 20-Oct-09 Influenza
24 94583 F 1986 12-Oct-09 Influenza
26 94568 F 1984 24-Oct-09 Influenza
27 94577 M 1983 18-Oct-09 Influenza
28 94598 F 1982 12-Oct-09 Influenza
29 94578 M 1981 16-Oct-09 Influenza
Ready


Spatio-Temporal Visualizer 2.0
File
Edit
View
Help
Geomap Tool
Toolbars
Panels
Layers
USSW_Zip
Charting Tool
Periodic Pattern Tool
Histogram Tool
LineChart Tool
Settings:
Unit of Time:
Days
9
8
7
6
5
4
3
2
1
0
Influenza
HIV
Hep C
Chlamydia
TimeLine Tool
Current Date Range:
Oct 12
2009
Oct 19
2009
Oct 26
2009
All cases
Influenza
HIV
Hep C
Chlamydia
Oct 8, 2009
Oct 27, 2009
Oct 8, 2009
Oct 27, 2009
Unit of Time
19 days (current)
Viz. Events / Total Events
205
/
205

http://bioportal.phfe.net



End of story

- For assistance or to provide feedback contact:
- michael.ascher@calema.ca.gov
- 916-342-1493