

Recommendations on Long-Term Information Management

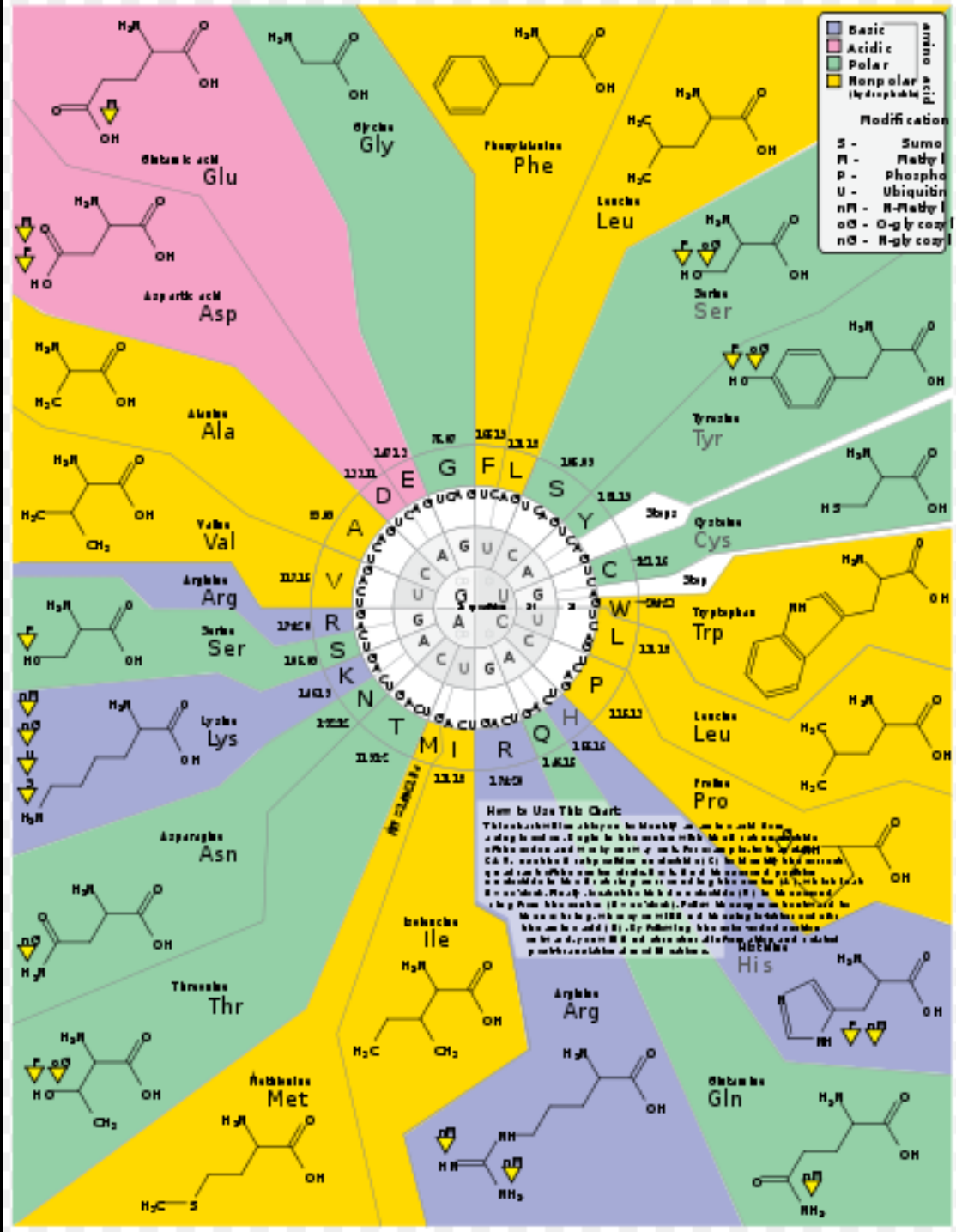
Federal Remediation Technologies
Roundtable Meeting on Data Management
May 2009

Mindy Vanderford, Ph.D.
GSI Environmental Inc.
www.gsi-net.com





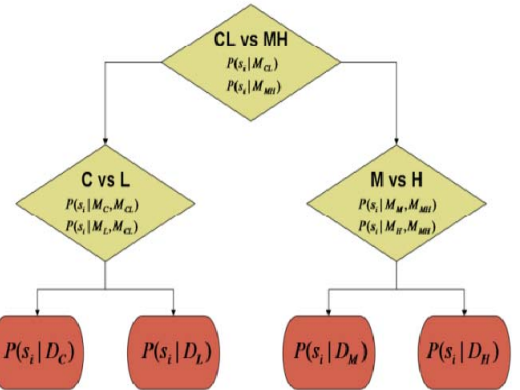




Translate Complex Data to Probability Distributions

RAW

num	sample name	samp rep	day	hour	tx	treatment	wt	ms id	total pM	PLFA
1	C01	1	0	0	C	Control	0.15	P035EC1	111309	TerBtS
2	C02	2	0	0	C	Control	0.27	P035EC2	196198	TerBtS
3	C03	3	0	0	C	Control	0.4	P035EC3	235721	TerBtS
4	C04	4	0	0	C	Control	0.56	P035EC4	537103	TerBtS
5	L01	1	0	0	L	Low	0.34	P035EC5	414699	TerBtS
6	L02	2	0	0	L	Low	0.29	P035EC6	312583	TerBtS
7	L03	3	0	0	L	Low	0.53	P035EC7	563971	TerBtS
8	M01	1	0	0	M	Med	0.42	P035EC10	445174	TerBtS
9	M02	2	0	0	M	Med	0.48	P035EC11	360255	TerBtS
10	M03	3	0	0	M	Med	0.34	P035EC12	389153	TerBtS
11	M04	4	0	0	M	High	0.79	P035EC13	607592	TerBtS
12	H01	1	0	0	H	High	0.38	P035EC14	342172	TerBtS
13	H02	2	0	0	H	High	0.46	P035EC15	284707	TerBtS
14	H03	3	0	0	H	High	0.42	P035EC16	480635	TerBtS
15	H04	4	0	0	H	High	0.54	P035EC17	484433	TerBtS
16	C21	1	1	24	C	Control	0.41	P035EC18	472161	TerBtS
17	C22	2	1	24	C	Control	0.54	P035EC19	445170	TerBtS
18	C23	3	1	24	C	Control	0.56	P035EC20	392334	TerBtS
19	C24	4	1	24	C	Control	0.63	P035EC21	389692	TerBtS
20	L21	1	1	24	L	Low	0.39	P035EC22	434875	TerBtS
21	L22	2	1	24	L	Low	0.37	P035EC24	246544	TerBtS
22	L23	3	1	24	L	Low	0.44	P035EC25	516695	TerBtS
23	L24	4	1	24	L	Low	0.44	P035EC26	516695	TerBtS



PROBABILITY

0.473711	0.313666	0.055953	0.287623	0.473711	0.313666	0.055953	0.287623
0.219278	0.295568	0.132782	0.608081	0.219278	0.295568	0.132782	0.608081
0.144396	0.306953	0.079038	0.310847	0.144396	0.306953	0.079038	0.310847
0.040405	0.132858	0.31589	0.378118	0.040405	0.132858	0.31589	0.378118
0.372054	0.158264	0.158264	0.175086	0.372054	0.158264	0.158264	0.175086
0.57432	0.386041	0	0.50014	0.57432	0.386041	0	0.50014
0.511906	0.382946	0	0.255744	0.511906	0.382946	0	0.255744
0.420278	0.339017	0.249848	0.283142	0.420278	0.339017	0.249848	0.283142
0.543464	0.221838	0.143506	0.370555	0.543464	0.221838	0.143506	0.370555
0.588386	0.339716	0.054256	0.367221	0.588386	0.339716	0.054256	0.367221
0.121556	0.469715	0	0.378118	0.121556	0.469715	0	0.378118
0.452489	0.263121	0.116695	0.313551	0.452489	0.263121	0.116695	0.313551
0.308997	0.256766	0.301043	0.391642	0.308997	0.256766	0.301043	0.391642
0.174569	0.22993	0.289089	0.133695	0.174569	0.22993	0.289089	0.133695
0.690154	0.326689	0	0.561341	0.690154	0.326689	0	0.561341
0.808593	0	0.214349	0.438506	0.808593	0	0.214349	0.438506
0.49537	0.440306	0	0.316043	0.49537	0.440306	0	0.316043
0.371778	0.359215	0.105276	0.335633	0.371778	0.359215	0.105276	0.335633
0.144396	0.306953	0.079038	0.310847	0.144396	0.306953	0.079038	0.310847
0.040405	0.132858	0.31589	0.378118	0.040405	0.132858	0.31589	0.378118
0.372054	0.158264	0.158264	0.175086	0.372054	0.158264	0.158264	0.175086
0.57432	0.386041	0	0.50014	0.57432	0.386041	0	0.50014
0.511906	0.382946	0	0.255744	0.511906	0.382946	0	0.255744
0.420278	0.339017	0.249848	0.283142	0.420278	0.339017	0.249848	0.283142
0.543464	0.221838	0.143506	0.370555	0.543464	0.221838	0.143506	0.370555
0.588386	0.339716	0.054256	0.367221	0.588386	0.339716	0.054256	0.367221
0.121556	0.469715	0	0.378118	0.121556	0.469715	0	0.378118
0.452489	0.263121	0.116695	0.313551	0.452489	0.263121	0.116695	0.313551
0.308997	0.256766	0.301043	0.391642	0.308997	0.256766	0.301043	0.391642
0.174569	0.22993	0.289089	0.133695	0.174569	0.22993	0.289089	0.133695
0.690154	0.326689	0	0.561341	0.690154	0.326689	0	0.561341
0.808593	0	0.214349	0.438506	0.808593	0	0.214349	0.438506
0.49537	0.440306	0	0.316043	0.49537	0.440306	0	0.316043
0.371778	0.359215	0.105276	0.335633	0.371778	0.359215	0.105276	0.335633
0.144396	0.306953	0.079038	0.310847	0.144396	0.306953	0.079038	0.310847
0.040405	0.132858	0.31589	0.378118	0.040405	0.132858	0.31589	0.378118
0.372054	0.158264	0.158264	0.175086	0.372054	0.158264	0.158264	0.175086
0.57432	0.386041	0	0.50014	0.57432	0.386041	0	0.50014
0.511906	0.382946	0	0.255744	0.511906	0.382946	0	0.255744
0.420278	0.339017	0.249848	0.283142	0.420278	0.339017	0.249848	0.283142
0.543464	0.221838	0.143506	0.370555	0.543464	0.221838	0.143506	0.370555
0.588386	0.339716	0.054256	0.367221	0.588386	0.339716	0.054256	0.367221
0.121556	0.469715	0	0.378118	0.121556	0.469715	0	0.378118
0.452489	0.263121	0.116695	0.313551	0.452489	0.263121	0.116695	0.313551
0.308997	0.256766	0.301043	0.391642	0.308997	0.256766	0.301043	0.391642
0.174569	0.22993	0.289089	0.133695	0.174569	0.22993	0.289089	0.133695
0.690154	0.326689	0	0.561341	0.690154	0.326689	0	0.561341
0.808593	0	0.214349	0.438506	0.808593	0	0.214349	0.438506
0.49537	0.440306	0	0.316043	0.49537	0.440306	0	0.316043
0.371778	0.359215	0.105276	0.335633	0.371778	0.359215	0.105276	0.335633
0.144396	0.306953	0.079038	0.310847	0.144396	0.306953	0.079038	0.310847
0.040405	0.132858	0.31589	0.378118	0.040405	0.132858	0.31589	0.378118
0.372054	0.158264	0.158264	0.175086	0.372054	0.158264	0.158264	0.175086
0.57432	0.386041	0	0.50014	0.57432	0.386041	0	0.50014
0.511906	0.382946	0	0.255744	0.511906	0.382946	0	0.255744
0.420278	0.339017	0.249848	0.283142	0.420278	0.339017	0.249848	0.283142
0.543464	0.221838	0.143506	0.370555	0.543464	0.221838	0.143506	0.370555
0.588386	0.339716	0.054256	0.367221	0.588386	0.339716	0.054256	0.367221
0.121556	0.469715	0	0.378118	0.121556	0.469715	0	0.378118
0.452489	0.263121	0.116695	0.313551	0.452489	0.263121	0.116695	0.313551
0.308997	0.256766	0.301043	0.391642	0.308997	0.256766	0.301043	0.391642
0.174569	0.22993	0.289089	0.133695	0.174569	0.22993	0.289089	0.133695
0.690154	0.326689	0	0.561341	0.690154	0.326689	0	0.561341
0.808593	0	0.214349	0.438506	0.808593	0	0.214349	0.438506
0.49537	0.440306	0	0.316043	0.49537	0.440306	0	0.316043
0.371778	0.359215	0.105276	0.335633	0.371778	0.359215	0.105276	0.335633
0.144396	0.306953	0.079038	0.310847	0.144396	0.306953	0.079038	0.310847
0.040405	0.132858	0.31589	0.378118	0.040405	0.132858	0.31589	0.378118
0.372054	0.158264	0.158264	0.175086	0.372054	0.158264	0.158264	0.175086
0.57432	0.386041	0	0.50014	0.57432	0.386041	0	0.50014
0.511906	0.382946	0	0.255744	0.511906	0.382946	0	0.255744
0.420278	0.339017	0.249848	0.283142	0.420278	0.339017	0.249848	0.283142
0.543464	0.221838	0.143506	0.370555	0.543464	0.221838	0.143506	0.370555
0.588386	0.339716	0.054256	0.367221	0.588386	0.339716	0.054256	0.367221
0.121556	0.469715	0	0.378118	0.121556	0.469715	0	0.378118
0.452489	0.263121	0.116695	0.313551	0.452489	0.263121	0.116695	0.313551
0.308997	0.256766	0.301043	0.391642	0.308997	0.256766	0.301043	0.391642
0.174569	0.22993	0.289089	0.133695	0.174569	0.22993	0.289089	0.133695
0.690154	0.326689	0	0.561341	0.690154	0.326689	0	0.561341
0.808593	0	0.214349	0.438506	0.808593	0	0.214349	0.438506
0.49537	0.440306	0	0.316043	0.49537	0.440306	0	0.316043
0.371778	0.359215	0.105276	0.335633	0.371778	0.359215	0.105276	0.335633
0.144396	0.306953	0.079038	0.310847	0.144396	0.306953	0.079038	0.310847
0.040405	0.132858	0.31589	0.378118	0.040405	0.132858	0.31589	0.378118
0.372054	0.158264	0.158264	0.175086	0.372054	0.158264	0.158264	0.175086
0.57432	0.386041	0	0.50014	0.57432	0.386041	0	0.50014
0.511906	0.382946	0	0.255744	0.511906	0.382946	0	0.255744
0.420278	0.339017	0.249848	0.283142	0.420278	0.339017	0.249848	0.283142
0.543464	0.221838	0.143506	0.370555	0.543464	0.221838	0.143506	0.370555
0.588386	0.339716	0.054256	0.367221	0.588386	0.339716	0.054256	0.367221
0.121556	0.469715	0	0.378118	0.121556	0.469715	0	0.378118
0.452489	0.263121	0.116695	0.313551	0.452489	0.263121	0.116695	0.313551
0.308997	0.256766	0.301043	0.391642	0.308997	0.256766	0.301043	0.391642
0.174569	0.22993	0.289089	0.133695	0.174569	0.22993	0	

Challenges

Prioritization Relevance

Curation Archiving

Quality

Access

Integration

Communication

Solutions?

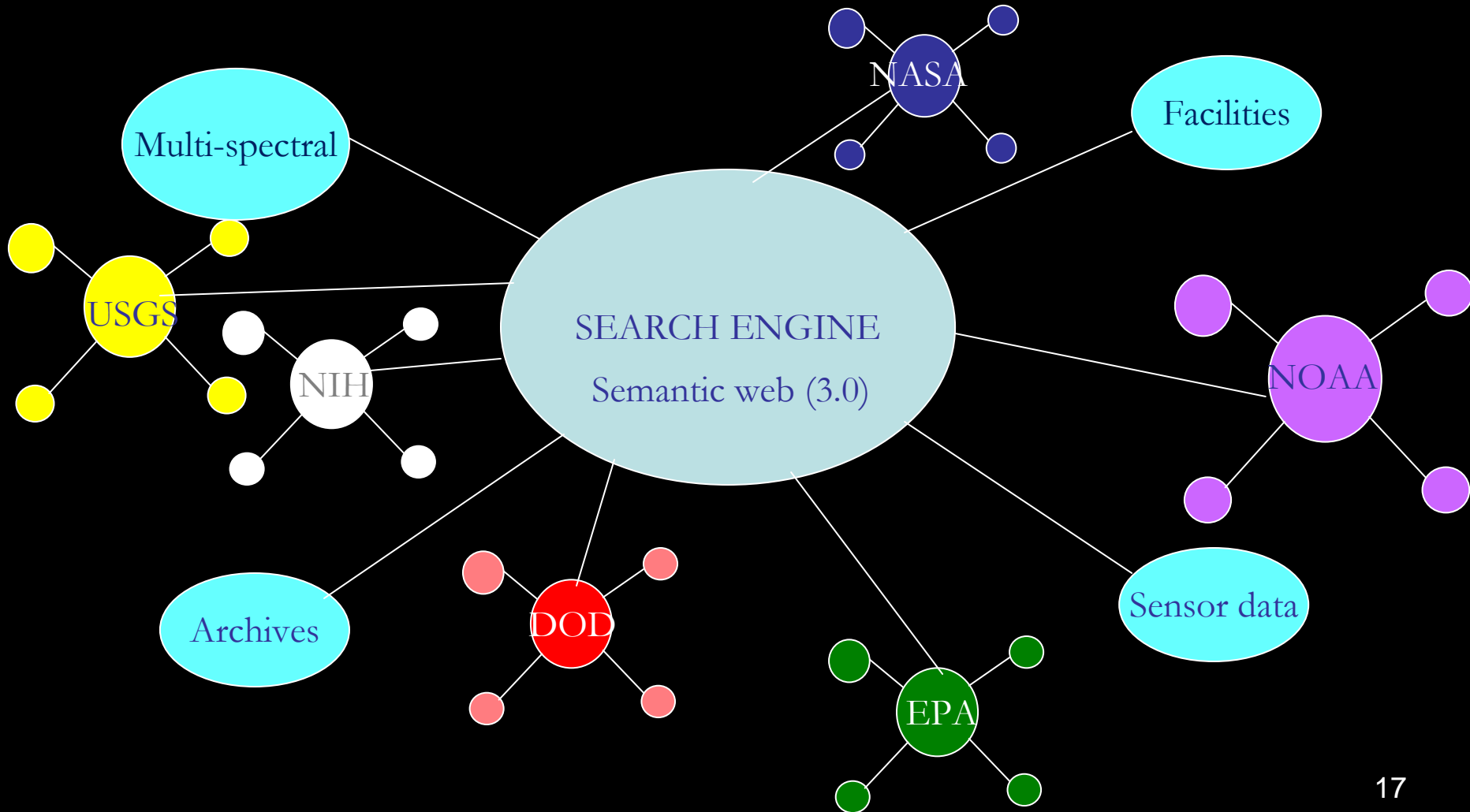
Existing infrastructure

STANDARDS

- **Review and revise governing DOE Orders**
 - **Contractor Standards** —Correct Incentives for contractors – no information left behind;
 - **Software Standards**—Preference for existing and open source code;
 - **Data Standards**—Need a guideline, or multi-agency approach, for establishing basic parameters.

Expanded Cyber-infrastructure (Semantic Web)

DOEgle



Information is costly to obtain

Information in physical form is costly to store

Information in physical form is stored using fixed ontology

Information in physical form is costly to retrieve and manipulate

Information in physical form is costly to disseminate

We must re-think information beyond material constraints

Digital information has no fixed physical form

Digital storage is easy

Digital preservation is difficult

Preservation of digital information is best done through redundancy, linking and degeneracy.

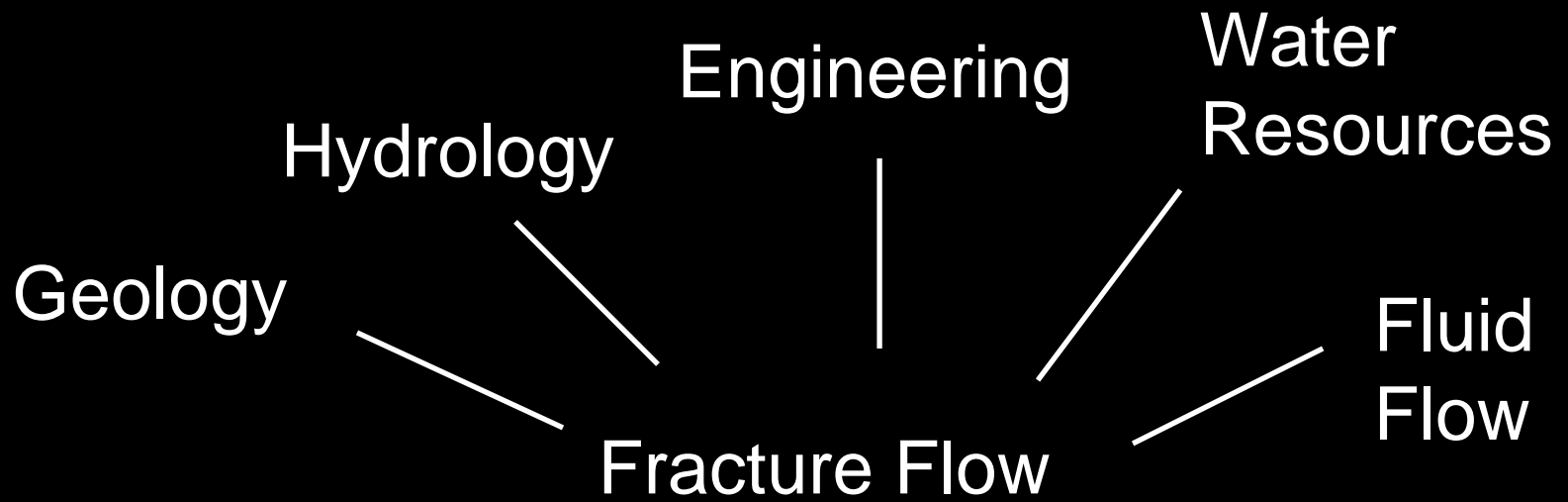
Redundancy

Unlikely X Many = Probable

Linking

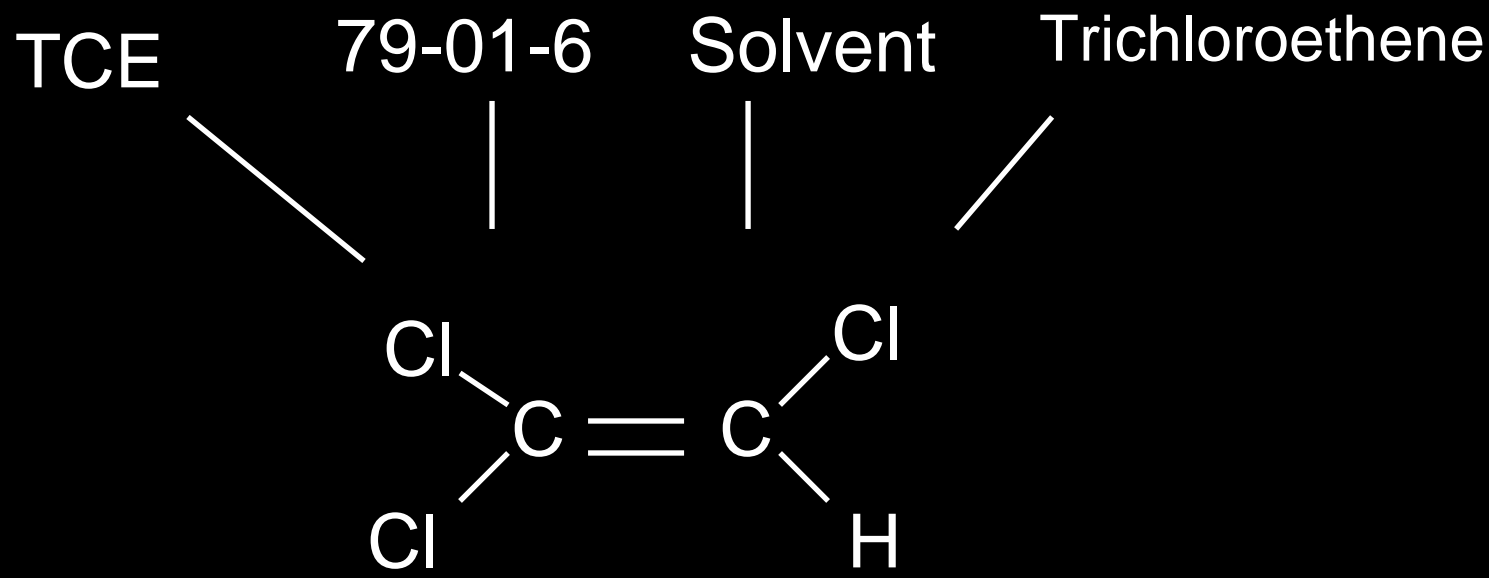
Metadata

(Sugary-delicious robot food)



Degeneracy





Transparency

Security

*to improve and provide transparency for
environmental decision making,*

facilitate long-term responsible care,

*and foster private-sector technology innovation
and growth.*