Infrastructure, Curation, and Metadata

Mark A. Parsons

NSF Workshop on Cyberinfrastructure for Polar Sciences
10 September 2013

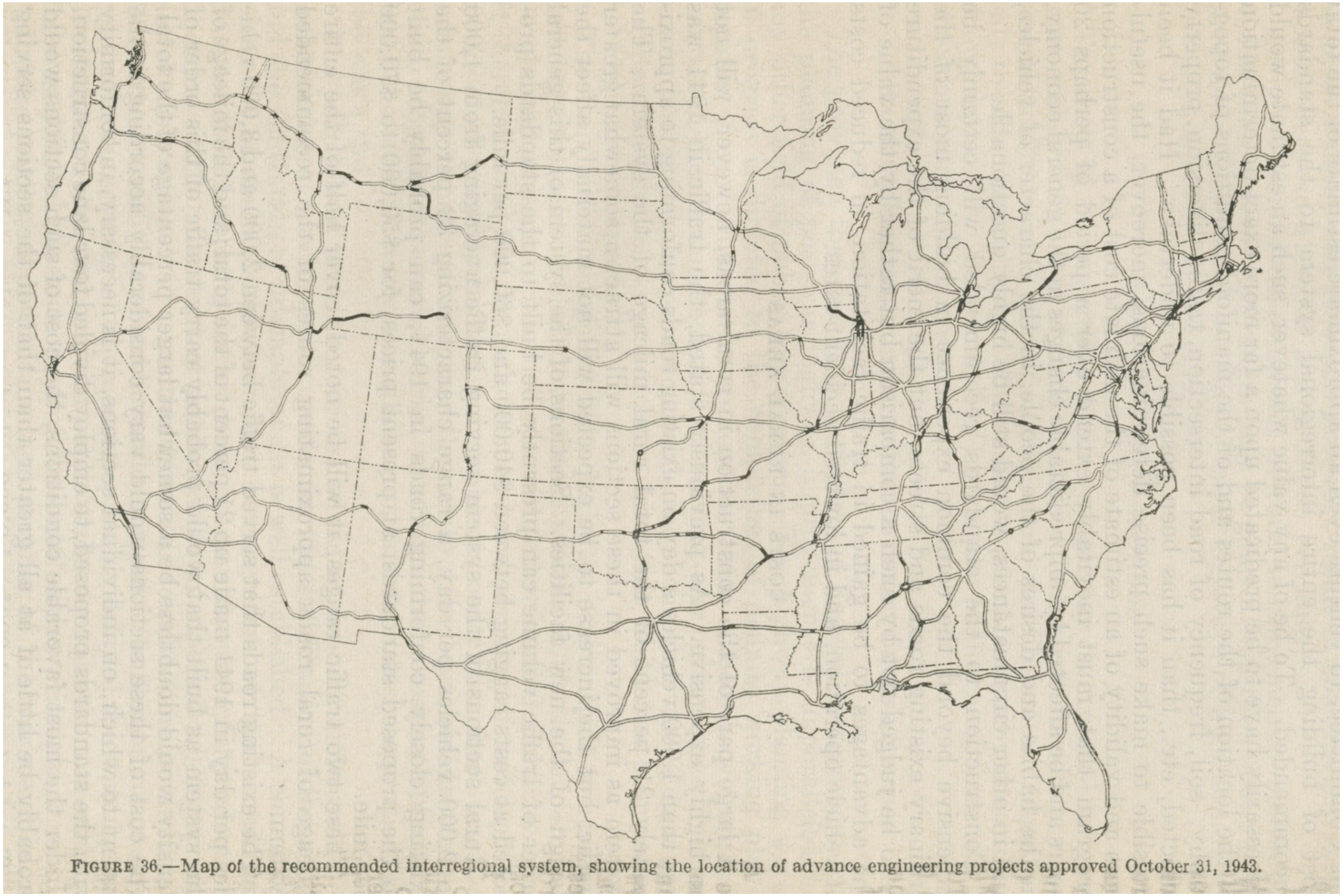


FIGURE 36.—Map of the recommended interregional system, showing the location of advance engineering projects approved October 31, 1943.

From *Interregional Highways: Message from the President of the United States Transmitting a Report of the National Interregional Highway Committee, Outlining and Recommending a National System of Interregional Highways*, 12 Jan. 1944.

http://www.shockblast.net/aerial-photographs/urban-sprawl-by-christoph-gielen-arizona/

Infrastructure is

Relationships, interactions, and connections
between humans, technologies, and institutions

Interchange

Ranch Exit

Understanding Infrastructure: Dynamics, Tensions, and Design

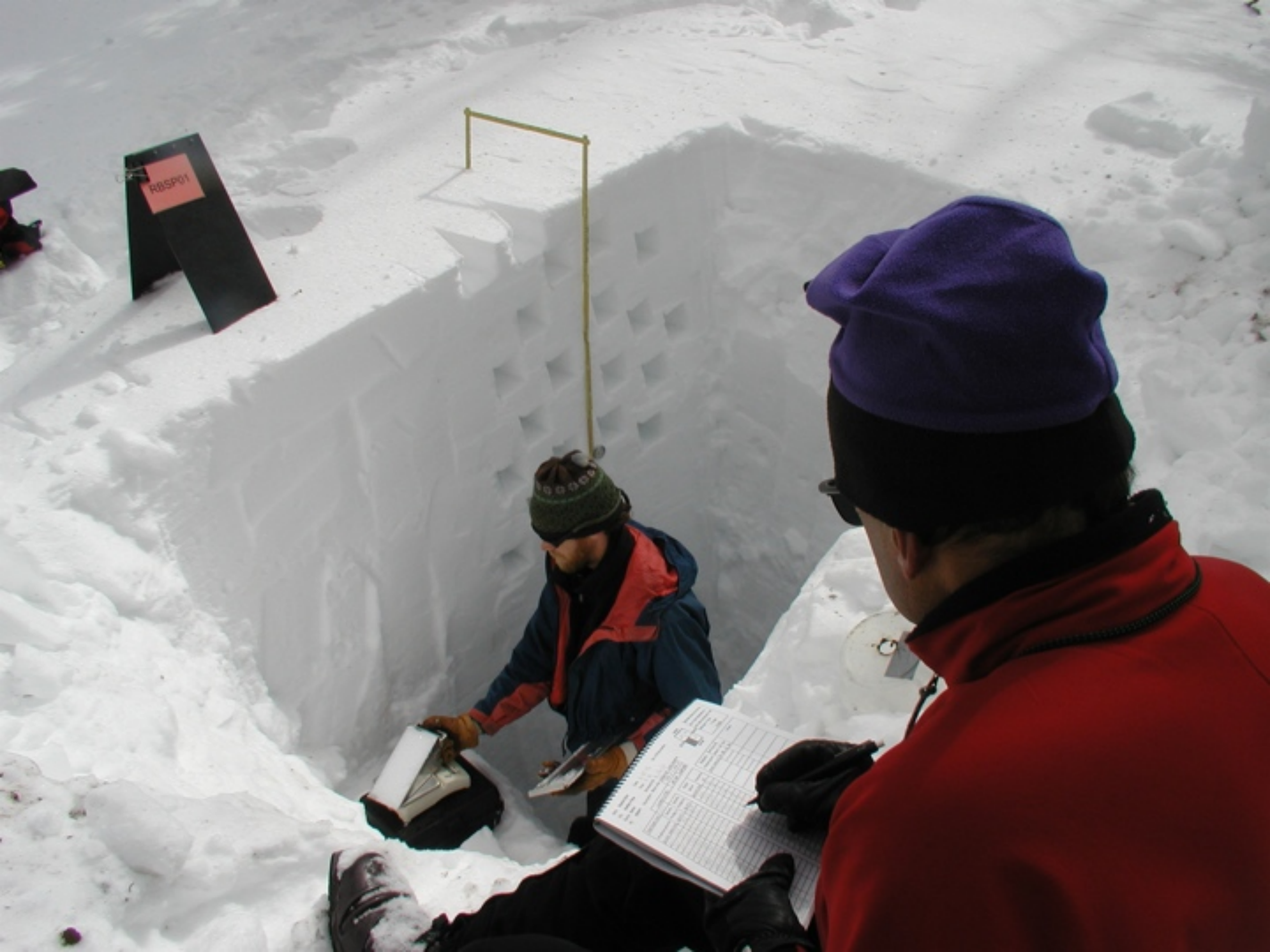
2007

Paul Edwards, Steven Jackson, Geoffrey Bowker,
and Cory Knobel

http://hdl.handle.net/2027.42/49353


JACK'S TIRE SERVICE
GOODYEAR
Budget
ONE WAY
Car and Truck Rental
homestore.com


Budget
ONE WAY
homestore.com


RBSP01


The Nordic Lodge
Pool Area


Elkhorn
Bar & Cafe
LIGHT


Budget
ONE WAY
homestore.com


data
wrangler
Budget
ONE WAY
homestore.com

Example of Snow Pit Data

We start with this:

MSA: **Rabbit Ears** Date: 2-25-03 ISA UTM Boundaries:

ISA: **Buffalo Pass** Time: 10:54

Sector: **Bravo** Surveyed By: HYMSTRA/MULLER/ELDER

Pit: **RBSP5** Description: Black pole, orange top - N end of large clearing

4,488,891 | B | C | 357,076 | 358,076 | A | D | 4,487,891

	UTM - E	UTM - N
Location (UTM):	~~357,414~~ 7451	~~4,488,859~~ 8873

20 M ACCURACY ON GPS

Total Depth (cm): 270

Ht above ground top (cm) - bottom (cm)	Density Profile A kg m-3	Density Profile B kg m-3	Height above ground (cm)	Temperature °C
270 - 260	71	64	270	-3
260 - 250	76	75	260	-8
250 - 240	86	88	250	-10
240 - 230	93	109	240	-10
230 - 220	110	112	230	-10
220 - 210	117	118	220	-10
210 - 200	142	138	210	-10
200 - 190	154	154	200	-9
190 - 180	178	183	190	-8
180 - 170	214	224	180	-8
170 - 160	263	249	170	-7
160 - 150	273	275	160	-7
150 - 140	297	291	150	-6
140 - 130	291	296	140	-6
130 - 120	320	320	130	-5
120 - 110	327	310	120	-4
110 - 100	338	320	100	-4
100 - 90	364	356	90	-3
90 - 80	369	376	80	-3
80 - 70	380	363	70	-3
70 - 60	340	347	60	-2
60 - 50	333	344		

Grain Size: Measure short and long axis of 3 grains representative of a small, medium and large grain in layer

Grain Type: N = New, R = Rounds, F = Facets, M = Mixed

Stratigraphy Ht above ground top (cm) - bottom (cm)	Axis type	Grain Size (mm) nearest 0.1 mm Sm	Med	Lg	Grain Type	Comments
270-223	short	.1	.2	.2	N	NEW SNOW
	long	.1	.3	.5		
223-170	short	.1	.2	.2	R	THIS WEEKS SNOW
	long	.1	.3	.4		
170-142	short	.1	.2	.3	F	PREVIOUS SNOWS MELT CRUST @ 142
	long	.2	.2	.4		
142-80	short	.2	.3	.5	R	TRENDING TOWARD FACETS
	long	.2	.5	.8		
80-35	short	.4	.7	.10	M	95% ROUNDED 5% FACETS
	long	.6	.8	1.0		
35-0	short	.4	.6	1.2	M	85% ROUNDS 15% FACETED
	long	.4	1.0	2.0		

DIRECTIONS:

Surface Wetness	D
Surface Roughness	Y
Soil Sample A	18
Soil Sample B	18
Canopy	N

Wetness: Dry (D), Moist (M), Wet (W), Very Wet (VW)

Roughness: Photo Taken? Yes (Y) or No (N)

Soil Samples: Depth of Core (cm)

2 samples, label as **RBSP5A** and **RBSP5B**

Canopy: Point underneath canopy? No (N), Yes/Coniferous (YC), Yes/Coniferous with snow (YCS), or Yes/Deciduous (YD)

Comments: PIT BETWEEN 2 ROWS OF TALL TREES.

Example of Snow Pit Data

And create this:

Summary Table

PIT	IOP	SECTOR	DATE	TIME	UTME	UTMN	SWET	SRUF	...	CNHT	DEPTH	SWE	DNS_LYRS	DNS_AVG	...
										cm	cm	mm		kg/m^3	
rbsp01	iop1	alpha	20020224	1115	357156	4487926	d	y		-999	214	637	22	298	
rbsp02	iop1	alpha	20020224	930	357211	4488306	d	y		-999	200	587	20	293	
rbsp03	iop1	bravo	20020224	928	357197	4488466	d	y		-999	180	514	18	286	
rbsp04	iop1	bravo	20020224	1330	357188	4488887	d	y		-999	225	667	23	298	
rbsp05	iop1	bravo	20020224	1528	357414	4488859	d	y		-999	212	679	22	320	

Stratigraphy Table

PIT_NAME	IOP#	TOP	BOT	SM-SHT	MD-SHT	LG-SHT	SM-LNG	MD-LNG	LG-LNG	GRN-TYPE	COMMENT
		cm	cm	mm	mm	mm	mm	mm	mm		
rbsp01	iop1	214	205	0.2	0.5	0.7	0.4	0.6	0.8	n	QC(000)
rbsp01	iop1	205	163	0.4	1	1	0.5	1	1.5	m	QC(000)
rbsp01	iop1	163	119	0.2	0.5	0.7	0.4	0.6	1.3	m	QC(000) Ice lens @119 cm.
rbsp01	iop1	119	19	0.3	0.3	0.5	0.3	0.7	1.2	m	QC(000)
rbsp01	iop1	19	0	0.4	0.5	2	0.5	0.8	2	f	QC(000)

Density Table

PIT_NAME	IOP#	TYPE	TOP	BOT	DENSITY-A	DENSITY-B	DENSITY-AVG	QC
			cm	cm	kg/m^3	kg/m^3	kg/m^3	
rbsp01	iop1	C	214	204	104	101	102.5	QC(000)
rbsp01	iop1	C	204	194	147	140	143.5	QC(000)
rbsp01	iop1	C	194	184	142	143	142.5	QC(000)
rbsp01	iop1	C	184	174	161	158	159.5	QC(000)

Temperature Table

PIT_NAME	IOP#	HEIGHT	TEMPERATURE	QC
		cm	deg-C	
rbsp01	iop1	214	-2	QC(000)
rbsp01	iop1	204	-6	QC(000)
rbsp01	iop1	194	-6	QC(000)
rbsp01	iop1	184	-7	QC(000)

Curation is

continual and conscious improvement of data.

Curation matters

- Embedding “data wranglers” in the field significantly improved data quality and completeness. (~20% of the data sheets had issues corrected in the field that would not have been correctable later).
- Reveals and documents tacit knowledge. Can even uncover science questions.
- Integrated much more than the field data with common grids and formats
    - “Standard” formats often do not exist and need to be created by the community. Sometimes you need multiple formats.
- A mediated relationship between user and collector.
- Need to include curators in data collection even if they are not in the field.
- Get to know your local curator.

Metadata are

Everything necessary to make data independently understandable by a designated community.

Infrastructure is comprised of relationships.

Curators create and enhance relationships.

Metadata is often the information shared (often tacitly) in those relationships.

A standard sensor format?

- Temporally varying sensors (e.g. a borehole)
- Spatially varying sensors (e.g. a transect)
- Temporally and spatially varying sensors (e.g. a drifting buoy)

Thank You

parsom3@rpi.edu

NOAA