# ROSETTA - UNIDATA'S WEB-BASED DATA TRANSLATION TOOL

Sean C. Arms <sup>1</sup>
Jennifer Oxelson Ganter <sup>1</sup>
Jeff Weber <sup>1</sup>
Mohan K. Ramamurthy <sup>1</sup>









- What is Rosetta?
  - Vision: General Purpose Data Format Translator
  - Goal: Get data into standard format









# What Many Users Like to Use

	<b>♠</b> Home	Layou	t Tab	les	Charts	SmartAi	rt	Formulas	Data	Review			^	- 30	ž.
E	dit :	F	ont		lignment	Numbe	er	Form	at	Cells	Then	nes			
2	- Ari	al	▼ 10	•	- □	General	•	¥ ≦š		<b>-</b>	Aa -				
Pa	ste E	I U	<u></u>	A v	Align	₩ 7 %	,	Conditional Formatting	Styles	Actions	Themes	Aa∗			
	D5	<b>‡</b>	8 0 (	fx											Ī
4	A	В	С	D	E	F	G	H		J	K	L	M	N	1
L		ature at differ	rent depths					RC-0632400 ar	nd ARC 085	6864					J
2	Ilu, Greenla					ntact =Vladim	ir E. Ron	nanovsky							
3		69.2390 W 5	51.0623		Professor										4
1_	Elevation (n	neters):				al Institute UA		: (907)474-745							4
_	Slope: flat				903 Koyuk		FAX	(907)474-729	90						4
7	Aspect: flat				P.O.Box 75										4
						AK 99775-73		ail: veromano							4
3					Data provi	ded byThoma	s Ingema	an-Nielsen, De	partment of	Civil Engineer	ing Technic	cal Univers	ity of Denmari	ĸ	4
)	Data data	00 100 10007 0	710410040												4
0		09/03/2007-0			007.6-14-				- 1 10						_
1	WARNING:	"999" fields n	nean not vali	d data; *-9	99" fields m	ean data are	absent; a	all temperature	s in grad C						4
2			Tomorotus	Tomoroto	Tomoroto	Townsont w	Tomorol	hu Tomonomitus	Tomorostus	Tomorotus					4
4	YEAR	DATE	remperatur 0	0.25				tur Temperatur 1 2	1emperatur 3	1emperature					+
5	2007	9/3/07	2.10,	2.32,			-0.3		-3.04.	-3.27					4
5 6	2007	9/4/07	3.63.	2.32,			-0.3		-3.04,	-3.27					1
7	2007	9/5/07	4.21,	2.66.	1.44,		-0.3		-3.08,	-3.31					4
8	2007	9/6/07	2.00.	2.28.	1.49,		-0.3		-3.08,	-3.31					4
9	2007	9/7/07	4.57.	2.40.			-0.3		-3.08,	-3.31					-
0	2007	9/8/07	5.32.	3.11.			-0.3		-3.08.	-3.31					+
1	2007	9/9/07	4.27,	3.13,			-0.3		-3.08,	-3.31					7
ź	2007	9/10/07	3.38.	2.68.	1.70,		-0.3		-3.08,	-3.31					1
3	2007	9/11/07	2.29.	2.39,			-0.2		-3.08,	-3.31					+
4	2007	9/12/07	0.26.	1.48.	1.31.		-0.3		-3.08.	-3.31					1
5	2007	9/13/07	1.21,	1.09,	0.92		-0.3		-3.08,	-3.31					1
6	2007	9/14/07	0.03,	0.91,			-0.3		-3.08,	-3.31					1
7	2007	9/15/07	0.15,	0.42,			-0.3		-3.07,	-3.31					1
8	2007	9/16/07	2.47,	0.48,	0.30,		-0.3		-3.06,	-3.31					1
9	2007	9/17/07	3.16,	1.34,	0.43,	0.06,	-0.3	9, -2.17,	-3.05,	-3.31					1
0	2007	9/18/07	2.53,	1.39,	0.66,		-0.3		-3.05,	-3.31					1
1	2007	9/19/07	1.45,	1.33,	0.77,	0.16,	-0.3	9, -2.16,	-3.05,	-3.31					1
-	2007	9/20/07	-0.38,	0.82,	0.70,	0.17,	-0.3	9, -2.16,	-3.05,	-3.31					1
2		0.004.007	-1.41.	0.26.	0.44.	0.13.	-0.3	9, -2.16,	-3.05,	-3.31					1
3	2007	9/21/07	-1.41,	0.20,	0.44,	0.13,	-0.3	9, -2.10,	-3.05,	-3.31					J.

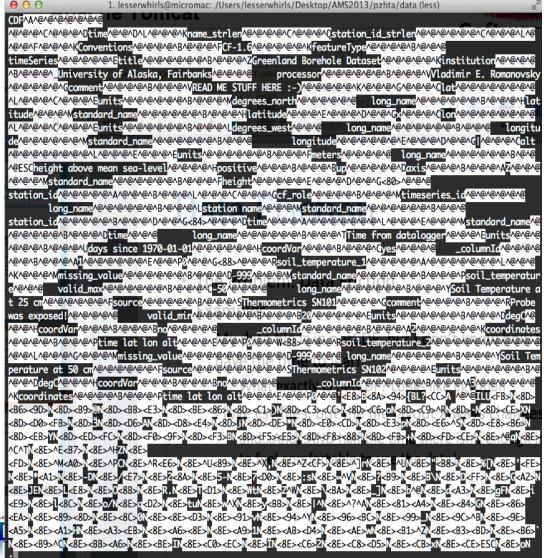








#### What a Standard Format "Looks" Like



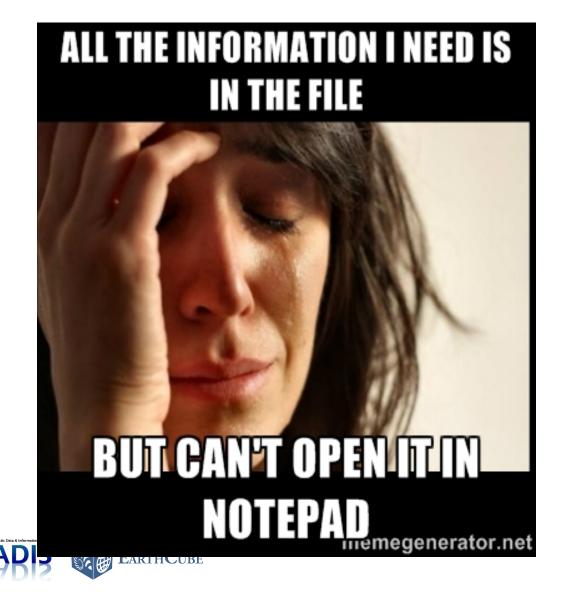








#### What a Standard Format "Looks" Like







- What is Rosetta?
  - Vision: General Purpose Data Format converter
  - Goal: Get data into standard format while still providing users with a format they want to use through the use of standard web services.









- Why bother translating data into a standard format?
  - What works for the data provider may not work well for the data users scientific workflow.









- What does translating data into a standard format buy us?
  - Standard format -> a variety of output formats (programmatically!)
  - Enable subsetting, aggregation, visual data "previews", enhancing search through automated rich metadata extraction, remote access, etc.

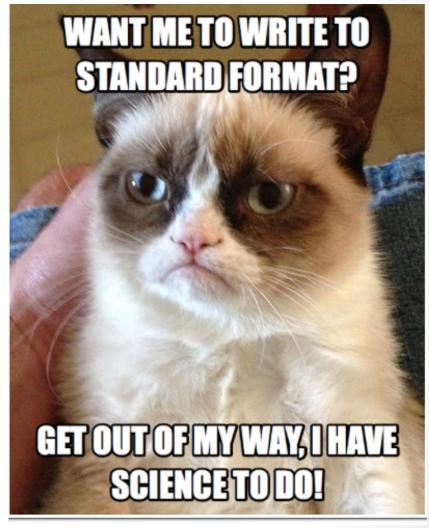








It's really all about the end users of the data (which could be you!)... but we need to get out of the data providers way as quickly as possible.











#### Wizard Based Format Translation

Rosetta		Enter Variable Attributes
Select Observation Platform Upload File	Specify Variable Attributes	What would you like to do with this column of data?  Assign a variable name  Do not use this column of data time
Specify Header Lines	# column 0 column	1
Specify Delimiters		Is this variable a coordinate variable? (examples: latitude, longitude, time)
Specify Variable Attributes	1 2002/10/01 00:30	i i i i i i i i i i i i i i i i i i i
	2 2002/10/01 01:00	What type of coordinate variable?
	3 2002/10/01 01:30	Time only (hour, minute, second, and/or millisecond)   \$
	4 2002/10/01 02:00	Specify variable data type:
Publish	5 2002/10/01 02:30	○ Integer ○ Float (decimal) ◎ Text
	6 2002/10/01 03:00	Required Metadata:
	7 2002/10/01 03:30	Variable Description   Time from Datalogo
Previous Next	8 2002/10/01 04:00	
	9 2002/10/01 04:30	Units • YYYY/MM/DD
	10 2002/10/01 05:00	show unit builder
	11 2002/10/01 05:30	What type of data are we building units for? date only
		unit prefix:   unit: (YYYY/MM/DD t) + -  Recommended Metadata:
Quick Save		CF Name time
Version : 0.2-SNAPSHOT Build Date: 20140101.101	oout Rosetta can be sent to: suppor	Additional Metadata:  (+) [-] (Calendar Type   \$)









## **Technology Stack**

- Java WebApp
- **€** Java<sup>™</sup>



- Java
- Apache Tomcat
- Spring MVC
- netCDF-Java(CDM)
- JavaScript
  - jQuery, SlickGrid, **jWizard**





{js}





**Explore GitHub** 



**Explore GitHub** 

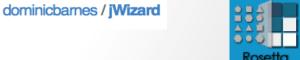
10

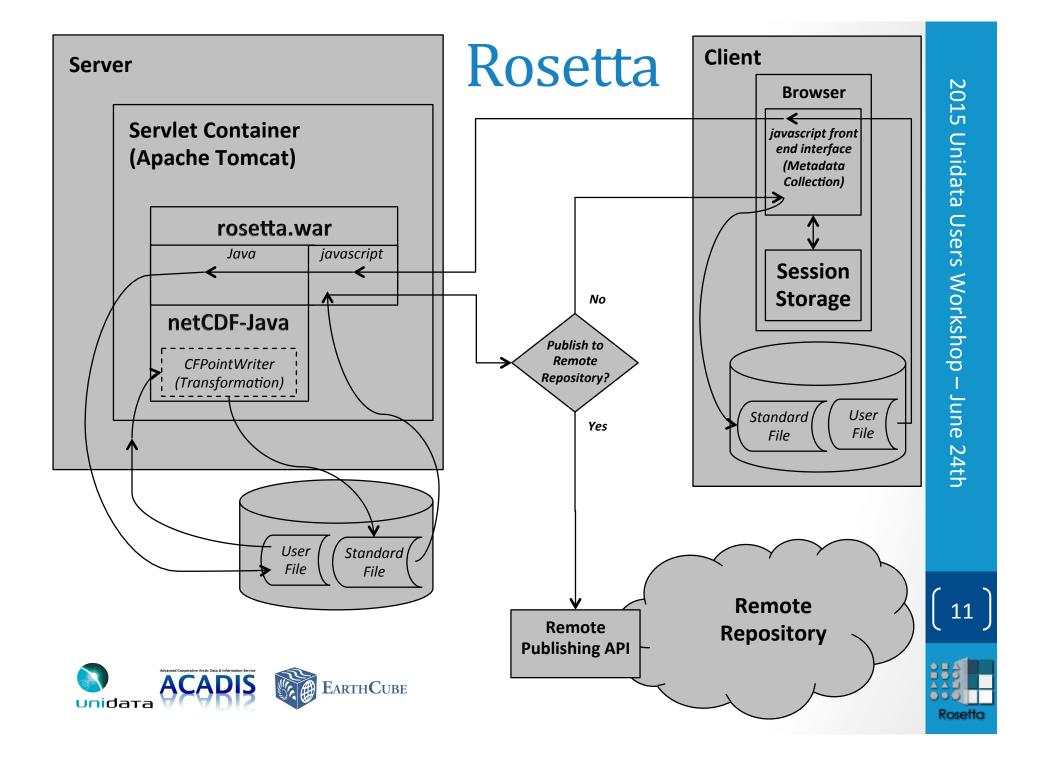












- Open to collaboration!
- Always looking for 'badASCII' example files
- Email:

support-rosetta@unidata.ucar.edu









Home page:

http://www.unidata.ucar.edu/software/rosetta/

Live testing site:

http://rosetta.unidata.ucar.edu

Source code on GitHub:

http://www.github.com/Unidata/rosetta









Development supported by:

NSF Award 1344155 (**GEO**): Unidata 2018: Transforming Geoscience through Innovative Data Services

NSF Award 0833450 (**AGS**): Unidata 2013: A Transformative Community Facility for the Atmospheric and Related Sciences

NSF Award 1016034 (ARC): The Advanced Cooperative Arctic Data and Information Service (ACADIS)

NSF Award 1343785 (ICER): EarthCube Building Blocks: Integrating Discrete and Continuous Data







