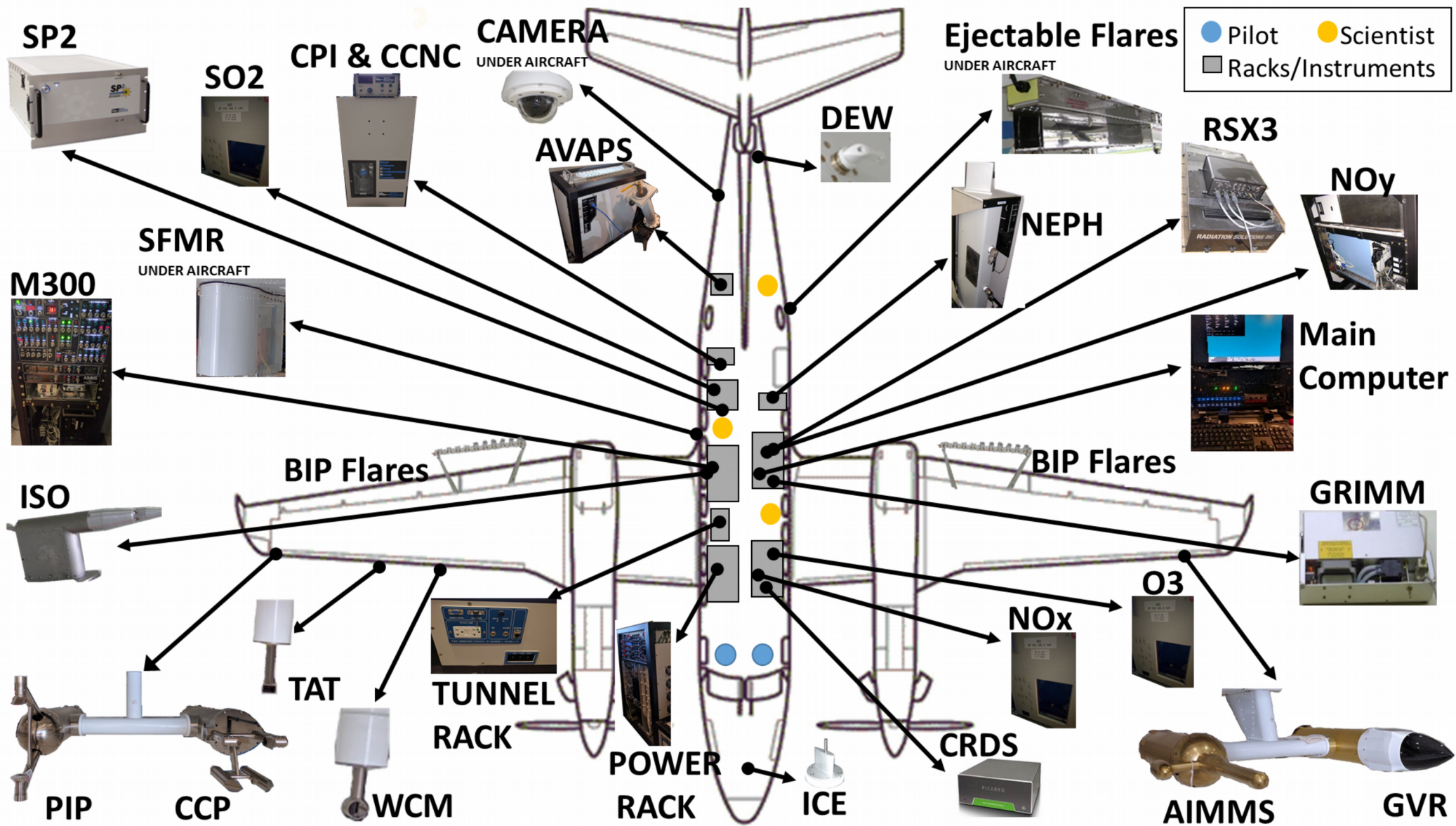


# Data Processing, Analysis, and Visualization



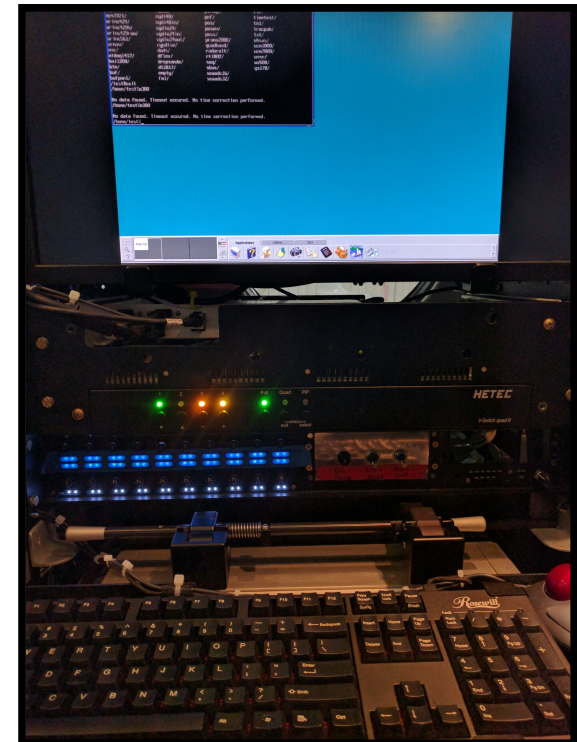
**Dr. David Delene**  
**University of North Dakota**



# Science Engineering Associates Model M300 Data Acquisition System (M300)



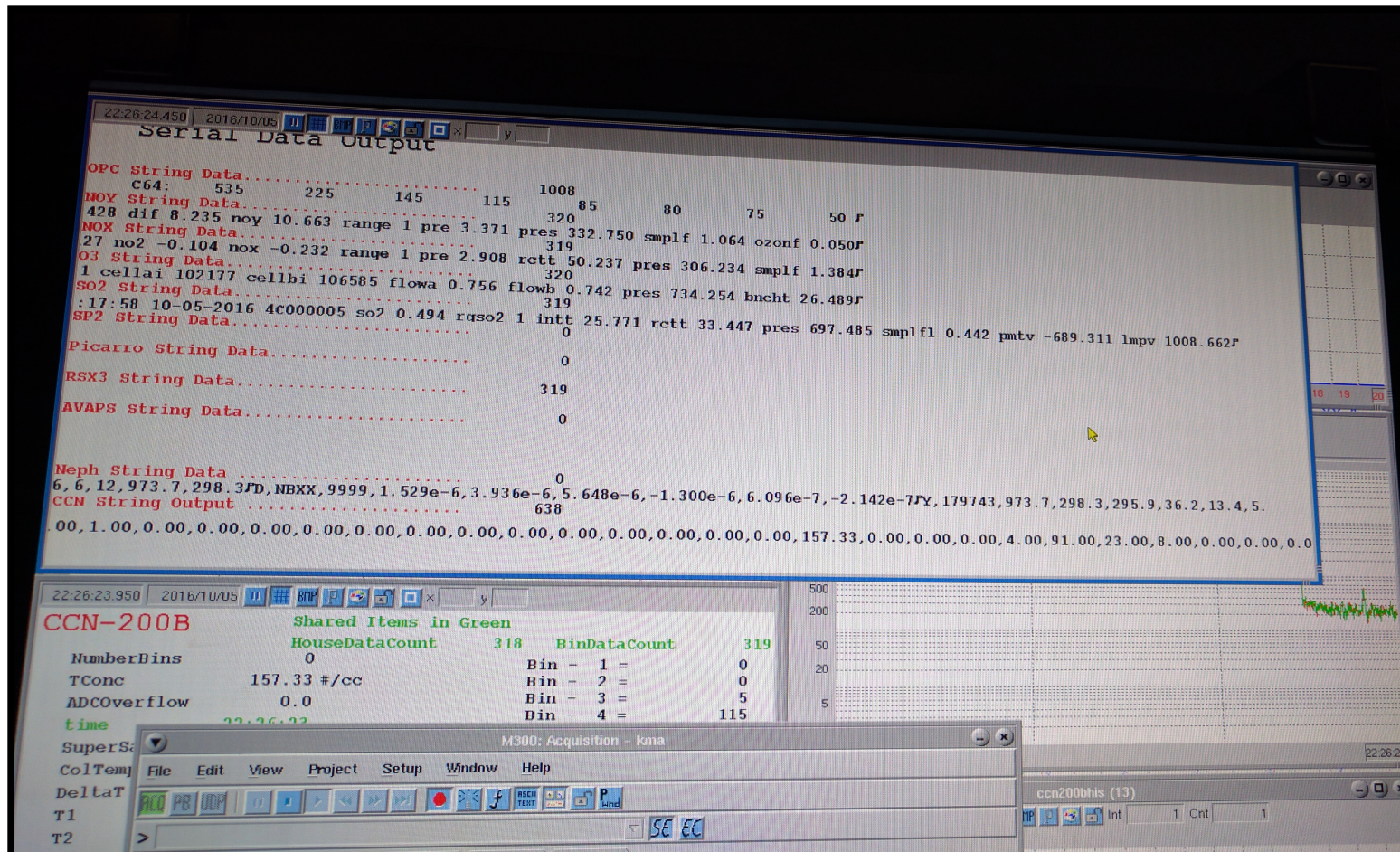
- The M300 acquires all data on the KMA King Air 350 aircraft.
- Some instruments also acquire data.
  - AIMMS
  - CCNC
  - CRDS
  - SFMS
  - SP2





# M300 Data Acquisition System

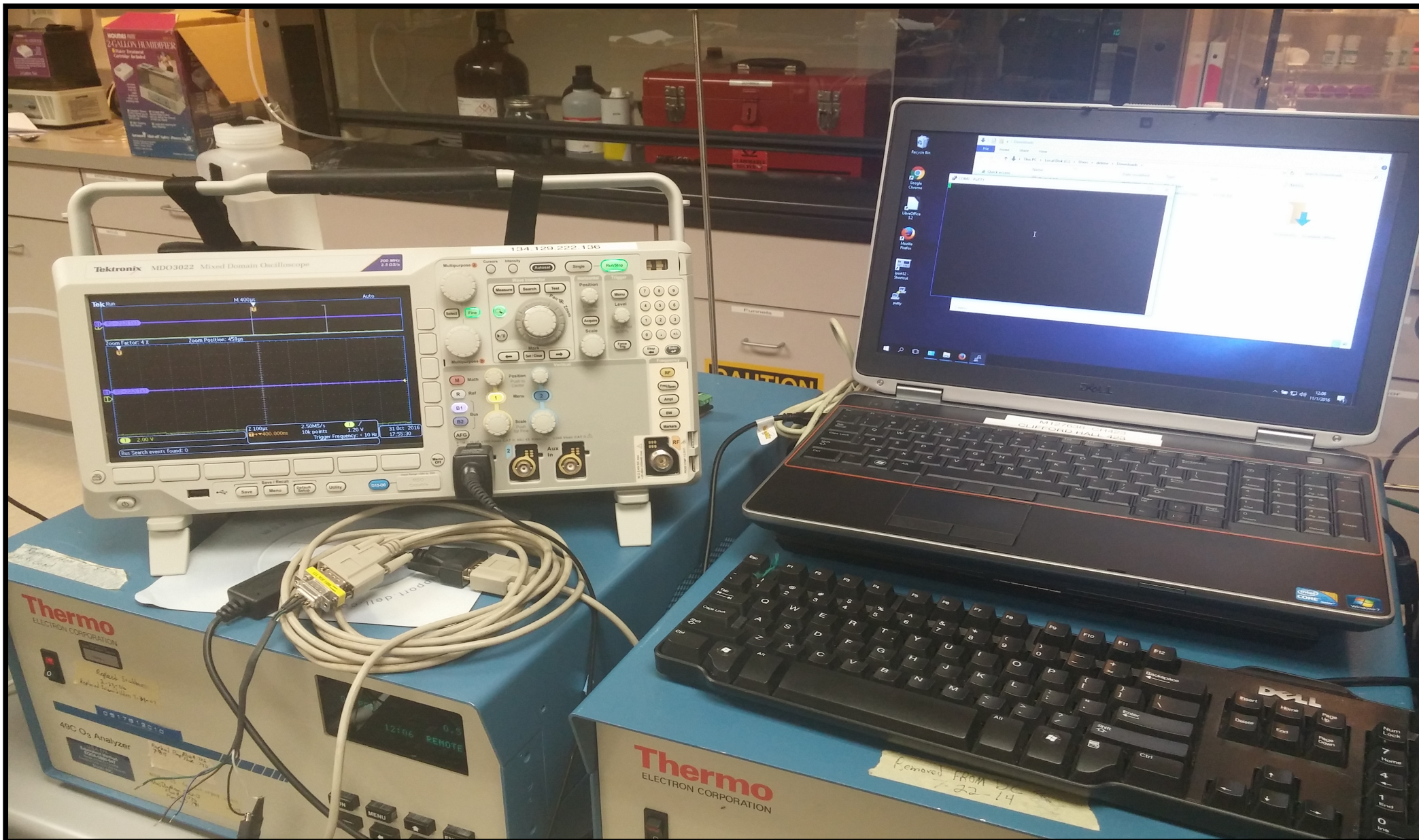
- Real-time Computer System (QNX OS)
- Optimized for Research Aircraft
- Totally Configurable Display System.



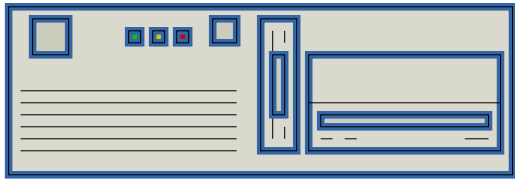


# Digital Data Acquisition

## Serializes Data (RS-232)







# Data Transmission



- **Analog**

- Simple, but Noisy
  - 500,000 Samples per Second

- **Digital**

- **UART (RS-232)**

- Low Speed Communication ( $< 1$  Mbit/sec)
- Point-to-Point

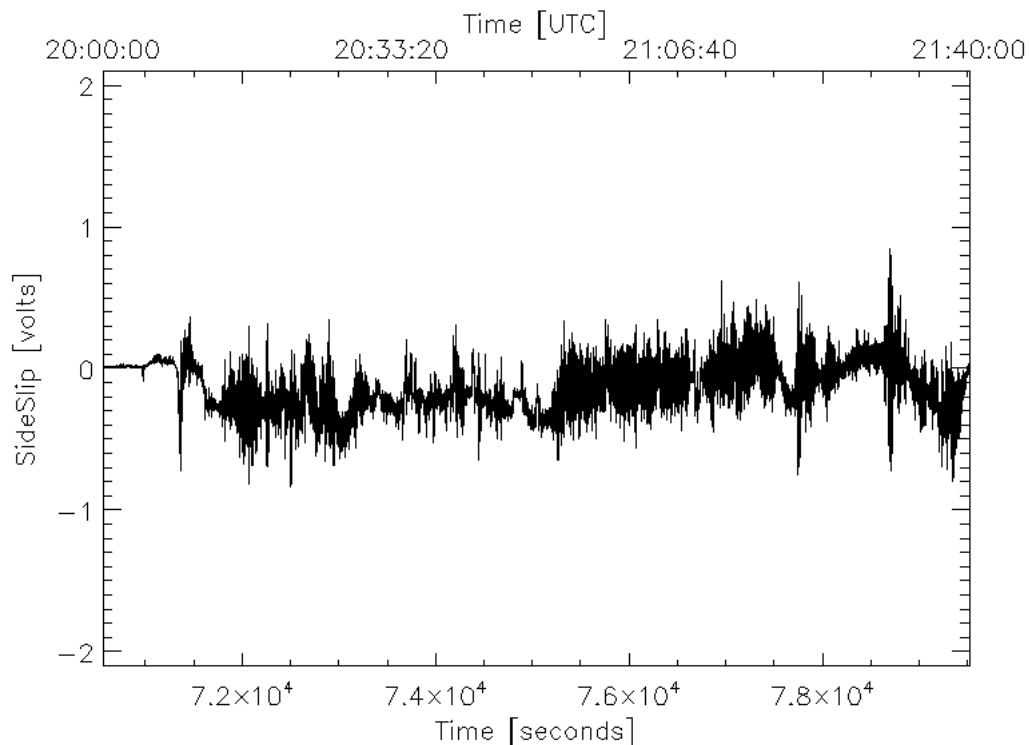
- **Ethernet (TCP/IP)**

- Higher Speed Communication
  - (100 Gbit/sec)
- Interconnect Number of Boxes

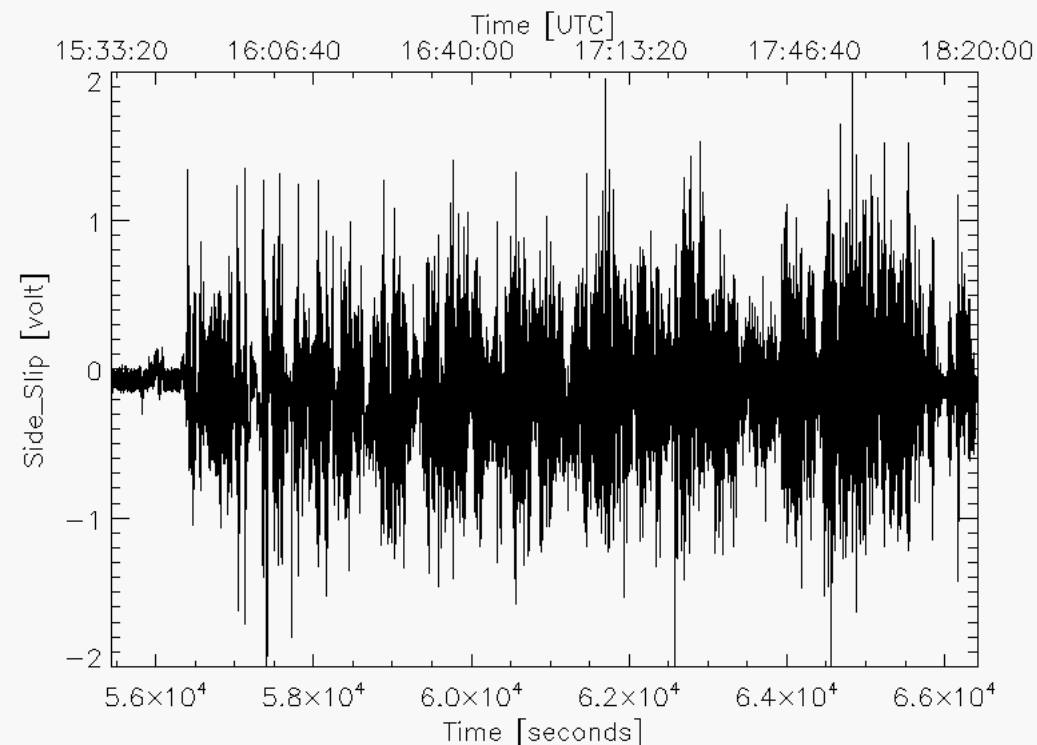


# Analog Voltage Measurement

**29 SEP 2005**



**2 SEP 2012**



Voltage measurement from the side slip pressure transducer on the North Dakota Citation Research Aircraft.



# Measurement Statistical Uncertainty

- To improve our certainty in a measurement, we can make repeated measurement.
- Repeated measurements can be represented as a Histogram that represents how the measurements are distributed.
- What single number best characterizes the complete group of measurements?
  - Mode – The value(s) the come up most often.
  - Median – The value where 50% of the measurements are above the value and 50% below.
  - Mean – Sum of all measurement divided by number of measurements.



# Information Representation

## Binary (Zeros and Ones)

- 01000101
- What is the Decimal Value?
- What is the Hex Value?
- What is the ASCII Text Character?



Esc	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
1B (1B) <1B> Π	0,3B (0,54) <0,5E> Π	0,3C (0,55) <0,5F> Π	0,3D (0,56) <0,60> Π	0,3E (0,57) <0,61> Π	0,3F (0,58) <0,62> Π	0,40 (0,59) <0,63> Π	0,41 (0,5A) <0,64> Π	0,42 (0,5B) <0,65> Π	0,43 (0,5C) <0,66> Π	0,44 (0,5D) <0,67> Π	0,85 (0,) <0,> Π	0,86 (0,) <0,> Π

Print	Scroll Lock	Pause
2A() <0,72> Π		()

CODE (in Extended Code- 0,XX	LEGEND Hexadecimal) Normal- XX Shift- (XX) Control- <XX> Alt- [XX]
---------------------------------------	---

~ 60(7E) <> Π	! 1 31(21) <> Π	@ 2 32(40) <0,3> Π	# 3 33(23) <> Π	\$ 4 34(24) <> Π	% 5 35(25) <> Π	^ 6 36(5E) <1E> Π	& 7 37(26) <> Π	* 8 38(2A) <> Π	( 9 39(28) <> Π	) 0 30(29) <> Π	- 2D(5F) <1F> Π	= 2B(2B) <> Π	BackSpace 08(08) <7F> Π
------------------------	--------------------------	-----------------------------	--------------------------	---------------------------	--------------------------	----------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	------------------------	----------------------------------

Insert	Home	Page Up
0,52 ( ) <>	0,47 ( ) <>	0,49 ( ) <>

Num Lock	/	*	-

Tab	Q	W	E	R	T	Y	U	I	O	P	[	]	Enter
	71(51) <11> Π	77(57) <17> Π	65(45) <05> Π	72(52) <12> Π	74(54) <14> Π	79(59) <19> Π	75(55) <15> Π	69(49) <09> Π	6F(4F) <0F> Π	70(50) <10> Π	5B(7B) <1B> Π	5D(7D) <1D> Π	

Delete	End	Page Down
0,53 ( ) <>	0,4F ( ) <>	0,51 ( ) <>

7 Home	8 ↑	9 Pg Up	
0,47 (37) <0,77>	0,48 (38) <>	0,49 (39) <0,84>	

Caps Lock	A	S	D	F	G	H	J	K	L	;	'	0D(0D) <0A> Π
	61(41) <01> Π	73(53) <13> Π	64(44) <04> Π	66(46) <06> Π	67(47) <07> Π	68(48) <08> Π	6A(4A) <0A> Π	6B(4B) <0B> Π	6C(4C) <0C> Π	3B(3A) <> Π	27(22) <> Π	

Insert	Home	Page Up
0,52 ( ) <>	0,47 ( ) <>	0,49 ( ) <>

4 ←	5	6 →	
0,4B (34) <0,73>	(35) <>	0,4D (36) <0,74>	

Shift	Z	X	C	V	B	N	M	<	>	?	Shift
	7A(5A) <1A> Π	78(58) <18> Π	63(43) <03> Π	76(56) <16> Π	62(42) <02> Π	6E(4E) <0E> Π	6D(4D) <0D> Π	2C(3C) <> Π	2E(3E) <> Π	2F(3F) <> Π	

↑
0,48 ( ) <>

1 End	2 ↓	3 Pg Dn	
0,4F (31) <0,75>	0,50 (32) <>	0,51 (33) <0,76>	

Ctrl	Alt	Macro	Space	 \ <1C> Π	Alt	Ctrl
		() <> Π	20(20) <20> [20]			

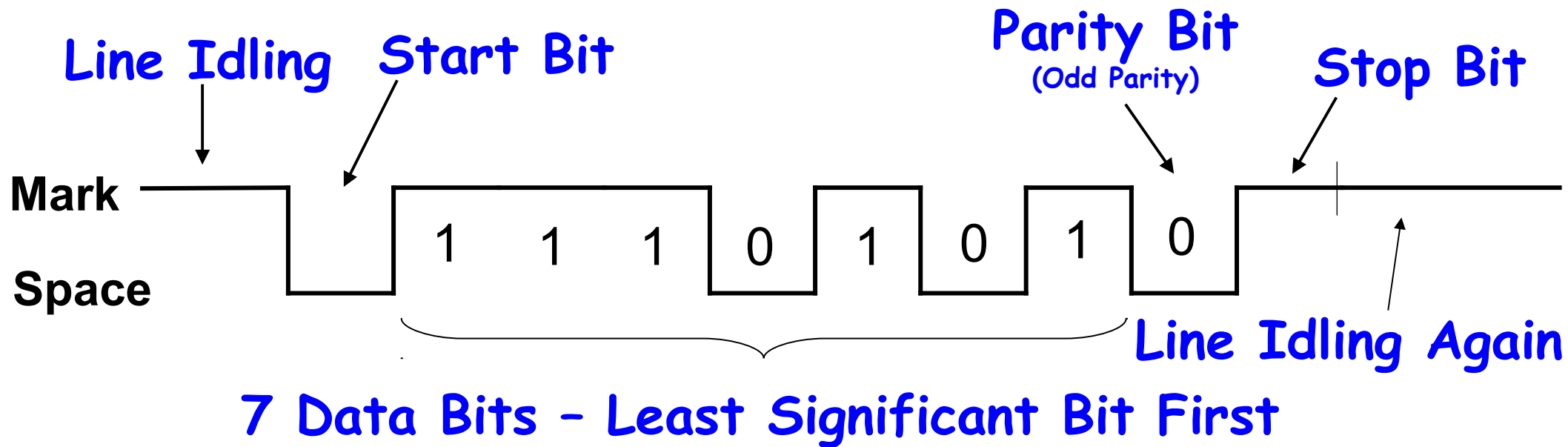
←	↓	→
0,4B ( ) <>	0,50 ( ) <>	0,4D ( ) <>

0 Insert	. Del	
0,52 (30) <>	0,53 (2E) <>	

# Binary Data Transmission



- Universal Asynchronous Receiver /Transmitter (UART) Example
- Send the ASCII letter 'W' (1010111)



How would Least Significant Bit Last be different?

Adapted from BYU UART Lecture,

<http://ece224web.groups.et.byu.net/lectures/UART.pdf>



# RS-232 Summary

- **Common Settings**

- 9600 8-N-1
- 115,200 7-E-1

- **Speed**

- Serial ports use two-level (binary) signalling, so the data rate in bits per second is equal to the symbol rate in baud.

- **Data Bits**

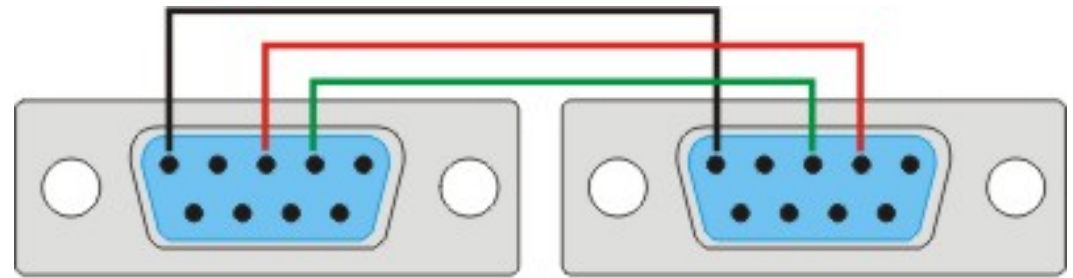
- The number of data bits in each character.
  - 7 for ASCII Data

- **Parity (N-None, O-Odd, E-Even, M-Mark, S-Space)**

- Method of detecting errors in transmission.

- **Stop Bits**

- Detect the end of a character.



5	-	GND	----	GND	-	5
2	-	RX	----	TX	-	3
3	-	TX	----	RX	-	2

# Necessary of Post-processing Data

- Data acquisition systems are a difficult and time consuming programming environment.
- Commercially available data acquisition systems are proprietary development environments that do not use open source software.
- There is a limited ability to test processing software since data acquisition systems are specialized computer systems with limited availability.
- Data acquisition systems are not modular data processing environments.
- Reprocessing of data using a data acquisition system is difficult to automate.
- There is absolutely no way of using measurements at a future point in time in a data processing algorithm when processing data in real time



# Scientific Programming

- Scientist write program because they understand the problem to be solved; however, they are judged/reward for writing papers, not producing data sets or software.
- Software used to collect, process and analyze measurements and models almost never evaluated.
  - Sometimes by supervisors or co-workers
  - No independent source code review. (Class???)
- Mistakes in software result in wrong scientific conclusions!
  - Black Box / White Box Testing



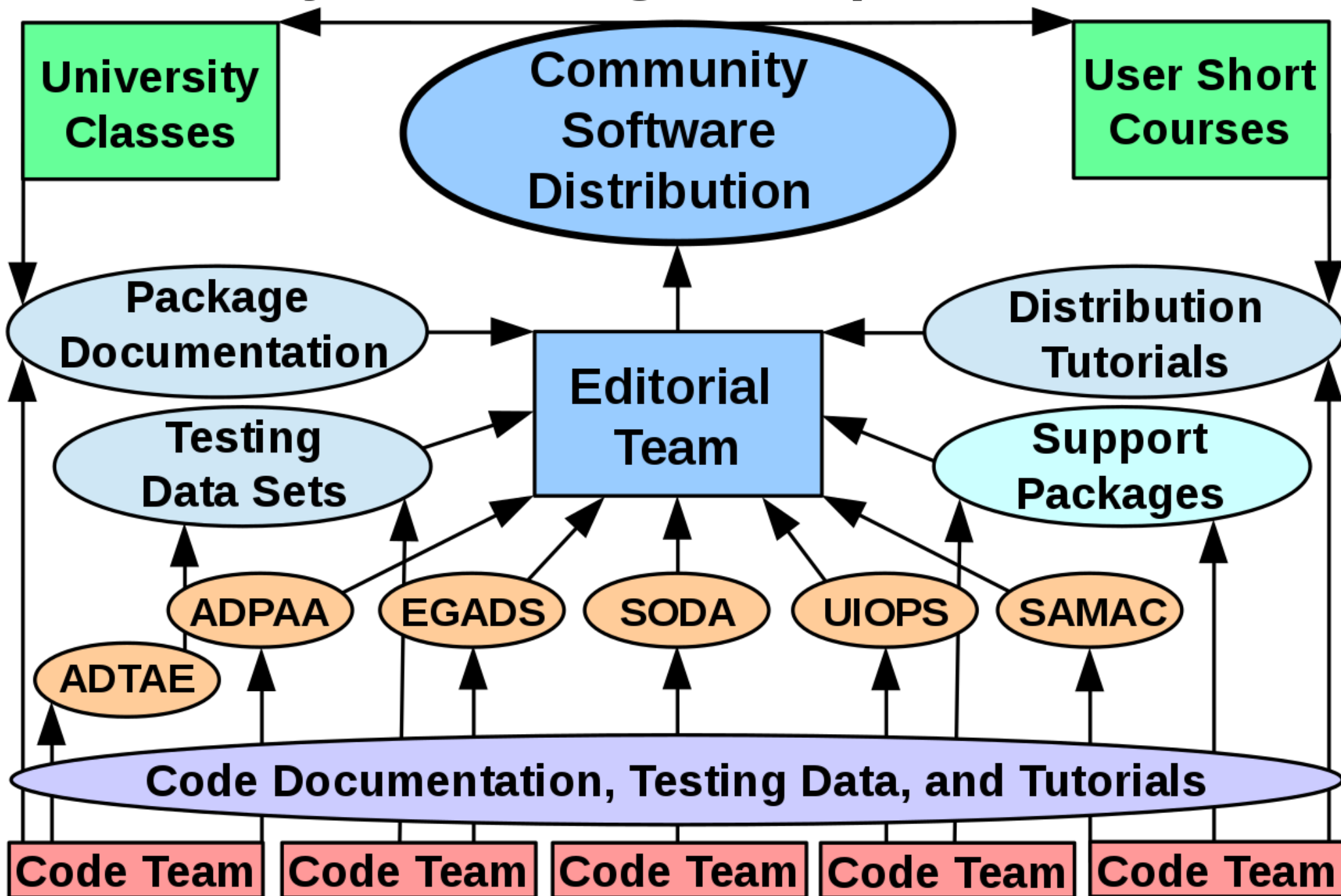
# Accuracy of Scientific Results

- Scientist could only reproduce 6 out of 53 “landmark” articles published by reputable labs in top journals.
  - <http://www.nature.com/nature/journal/v483/n7391/full/483531a.html>
  - <http://www.reuters.com/article/2012/03/28/us-science-cancer-idUSBRE82R12P20120328>
- John P. A Ioannidis explains how “It can be proven that most claimed research findings are false.”
  - <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1182327/>
- Personnel Experience

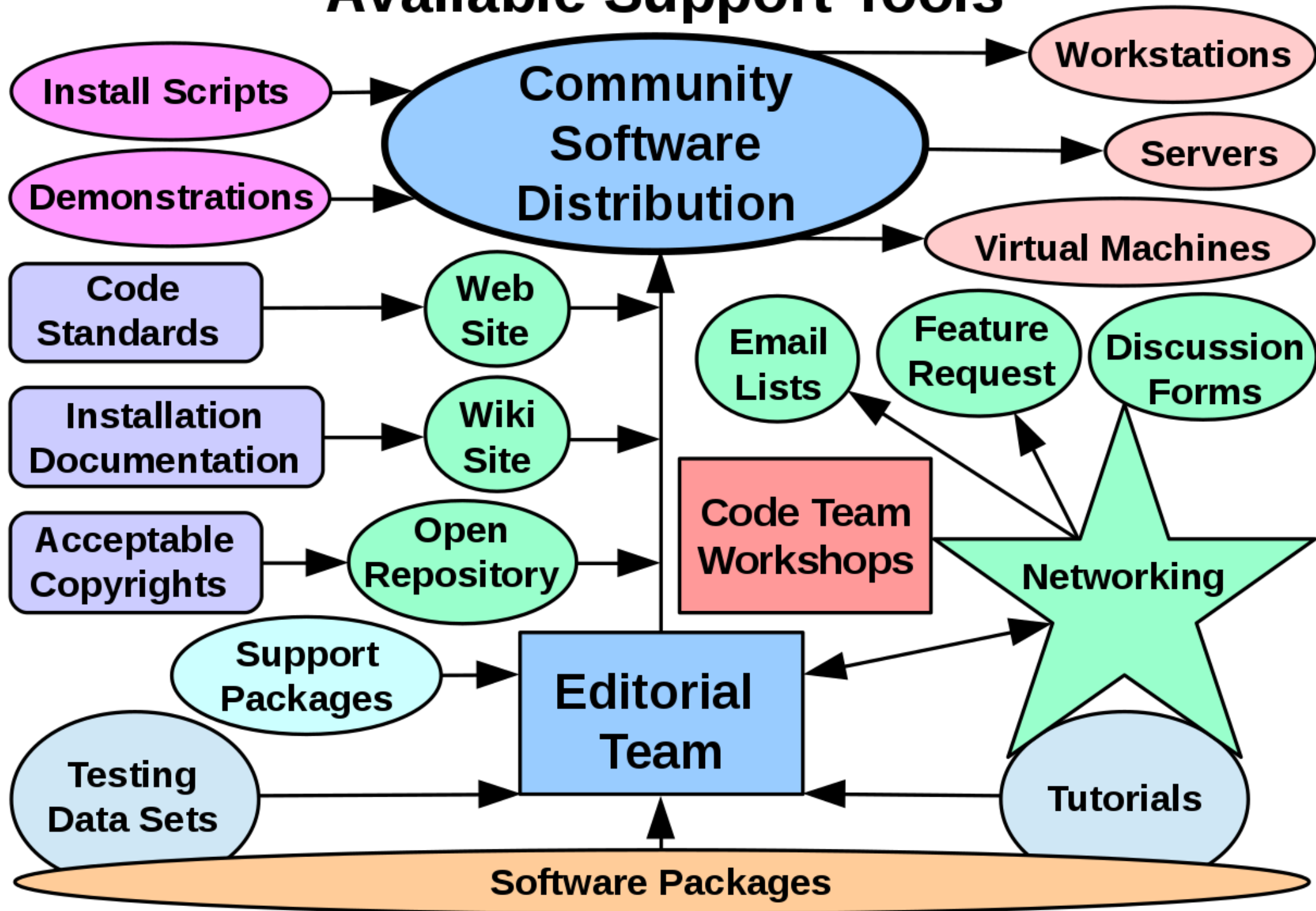




# System Design Components

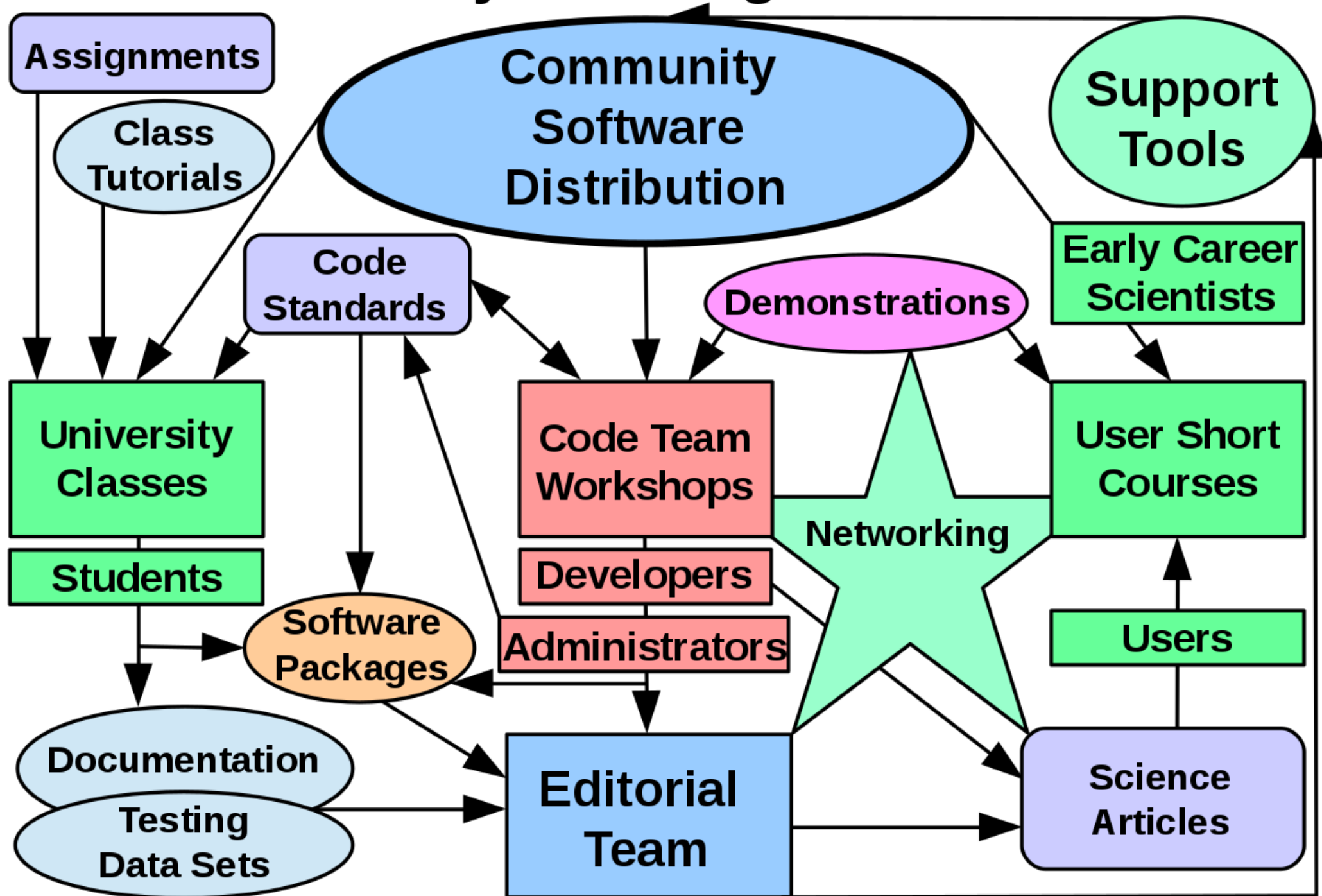


# Available Support Tools





# Community Learning Environment



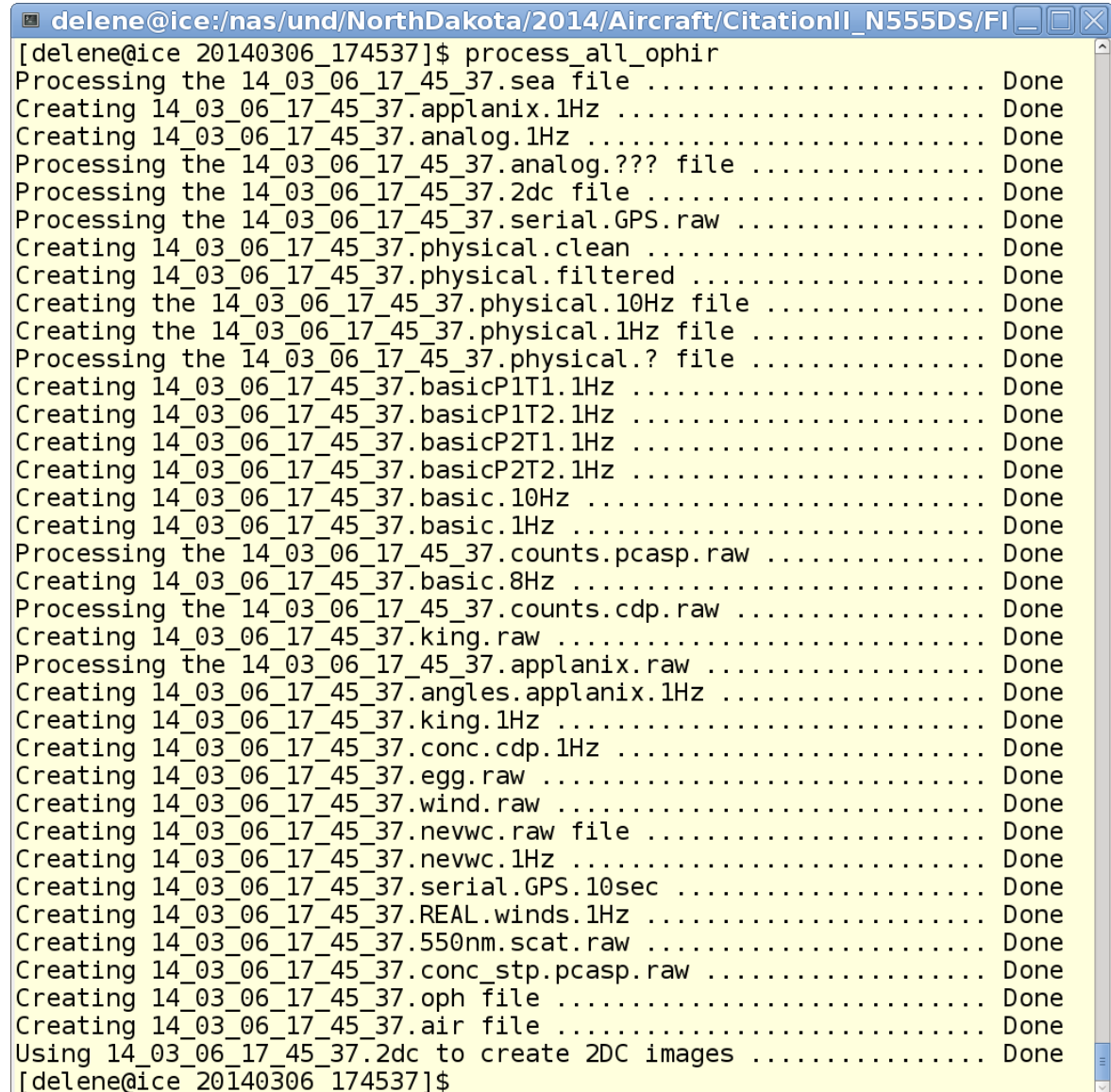
# ADPAA Computer Requirements

- Data Processing
  - Flight and Project Processing Scripts
    - Linux (Redhat, Fedora, Ubuntu, Mint)
- Analysis
  - Cplot/Cplot2 Visualization Packages
    - Linux, Mac, Windows



# Package Programming Language(s)

- Main Language
  - IDL (GDL) – Binary Version Available (No License Fee)
- Additional Languages
  - **Python**
  - Perl
  - Bash
  - Csh
  - FORTRAN
  - C
  - Matlab
  - Scilab

A terminal window titled 'delene@ice:/nas/und/NorthDakota/2014/Aircraft/CitationII\_N555DS/FI' showing the execution of a script named 'process\_all\_ophir'. The script processes various data files and creates corresponding output files, all of which are marked as 'Done'. The files include sea, applanix, analog, 2dc, serial.GPS, physical, basicP1T, basicP1T2, basicP2T, basicP2T2, basic, counts, king, angles, conc, egg, wind, nevwc, REAL, 550nm, conc\_stp, oph, and air files.

```
delene@ice:/nas/und/NorthDakota/2014/Aircraft/CitationII_N555DS/FI$ process_all_ophir
Processing the 14_03_06_17_45_37.sea file ..... Done
Creating 14_03_06_17_45_37.applanix.1Hz ..... Done
Creating 14_03_06_17_45_37.analog.1Hz ..... Done
Processing the 14_03_06_17_45_37.analog.??? file ..... Done
Processing the 14_03_06_17_45_37.2dc file ..... Done
Processing the 14_03_06_17_45_37.serial.GPS.raw ..... Done
Creating 14_03_06_17_45_37.physical.clean ..... Done
Creating 14_03_06_17_45_37.physical.filtered ..... Done
Creating the 14_03_06_17_45_37.physical.10Hz file ..... Done
Creating the 14_03_06_17_45_37.physical.1Hz file ..... Done
Processing the 14_03_06_17_45_37.physical.? file ..... Done
Creating 14_03_06_17_45_37.basicP1T1.1Hz ..... Done
Creating 14_03_06_17_45_37.basicP1T2.1Hz ..... Done
Creating 14_03_06_17_45_37.basicP2T1.1Hz ..... Done
Creating 14_03_06_17_45_37.basicP2T2.1Hz ..... Done
Creating 14_03_06_17_45_37.basic.10Hz ..... Done
Creating 14_03_06_17_45_37.basic.1Hz ..... Done
Processing the 14_03_06_17_45_37.counts.pcaspc.raw ..... Done
Creating 14_03_06_17_45_37.basic.8Hz ..... Done
Processing the 14_03_06_17_45_37.counts.cdp.raw ..... Done
Creating 14_03_06_17_45_37.king.raw ..... Done
Processing the 14_03_06_17_45_37.applanix.raw ..... Done
Creating 14_03_06_17_45_37.angles.applanix.1Hz ..... Done
Creating 14_03_06_17_45_37.king.1Hz ..... Done
Creating 14_03_06_17_45_37.conc.cdp.1Hz ..... Done
Creating 14_03_06_17_45_37.egg.raw ..... Done
Creating 14_03_06_17_45_37.wind.raw ..... Done
Creating 14_03_06_17_45_37.nevwc.raw file ..... Done
Creating 14_03_06_17_45_37.nevwc.1Hz ..... Done
Creating 14_03_06_17_45_37.serial.GPS.10sec ..... Done
Creating 14_03_06_17_45_37.REAL.winds.1Hz ..... Done
Creating 14_03_06_17_45_37.550nm.scap.raw ..... Done
Creating 14_03_06_17_45_37.conc_stp.pcaspc.raw ..... Done
Creating 14_03_06_17_45_37.oph file ..... Done
Creating 14_03_06_17_45_37.air file ..... Done
Using 14_03_06_17_45_37.2dc to create 2DC images ..... Done
delene@ice 20140306_174537]$
```



# Data Formats Supported

- Input from any \*.sea acquisition file from the Science Engineering Associates (SEA) Data System
  - Addition Level 1 models could additional data system.
  - Conversion scripts support instrument recorded data.
- Main Data Format
  - Gzipped (compressed) NASA/UND ASCII 1001 Format
- Output Format
  - NASA/UND ASCII 1001 Format
  - 2013 ICARTT NASA Format (ict)
  - NetCDF
  - KML (Google Earth)
- Data File Import Routines
  - IDL, Python, Matlab, Scilab, Igor
- Conversion Routines
  - Too many to list.



# Availability and Copyright

- SourceForge Repository (SVN)

- Current version 3,174
- 2 Admins
- 12 Active Developers
- 6 Inactive Developers

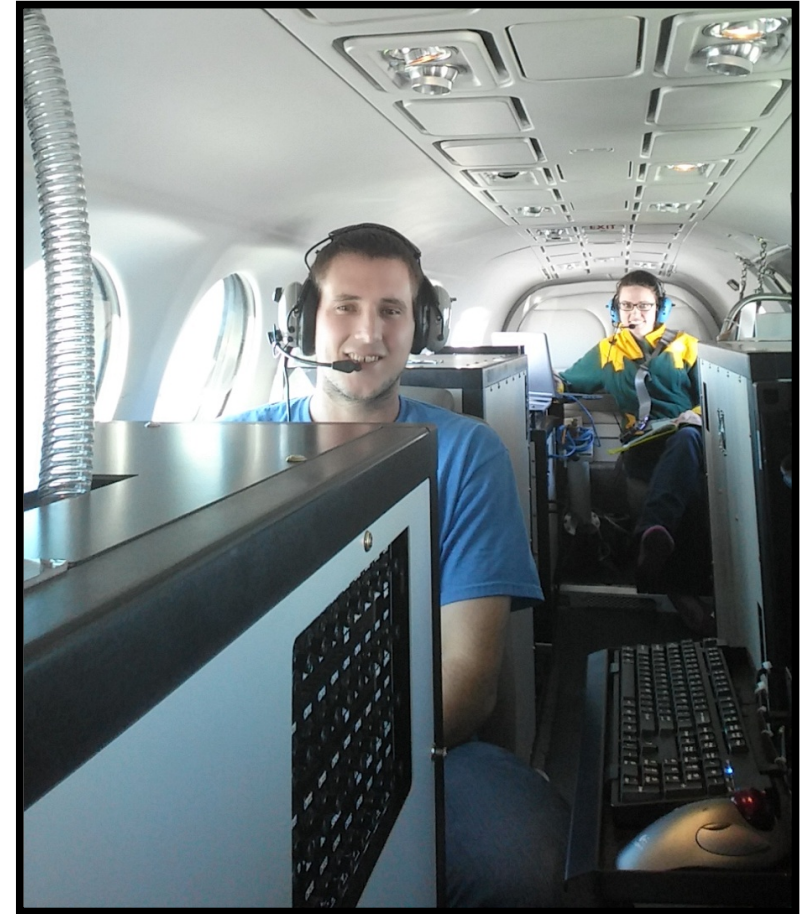
- Download

- svn checkout

`svn://svn.code.sf.net/p/adpaa/code/trunk adpaa-code`

- GNU/GPL v3 Licensed

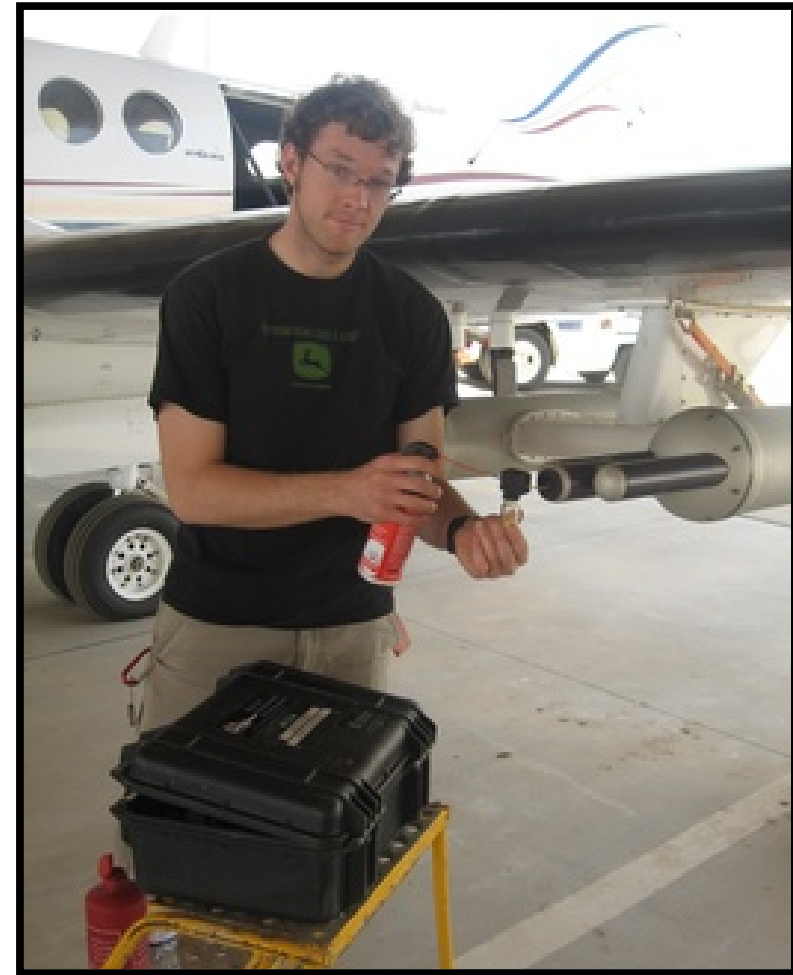
- Have to remain open, even when forked.



# Data Processing

(From 20170404\_190617 directory, process\_all\_kma)

- Data Quality Control
  - Calibration Checks
  - Data Missing Values Codes
  - Levels of Data Processing
  - Raw Recorded Data
  - Engineering to Physical Units
  - Single Instrument Data Files
  - Combined Instrument Data File
- Data Quality Assurance
  - Scientific Data Review
  - Scripts Search for Unrealistic Values

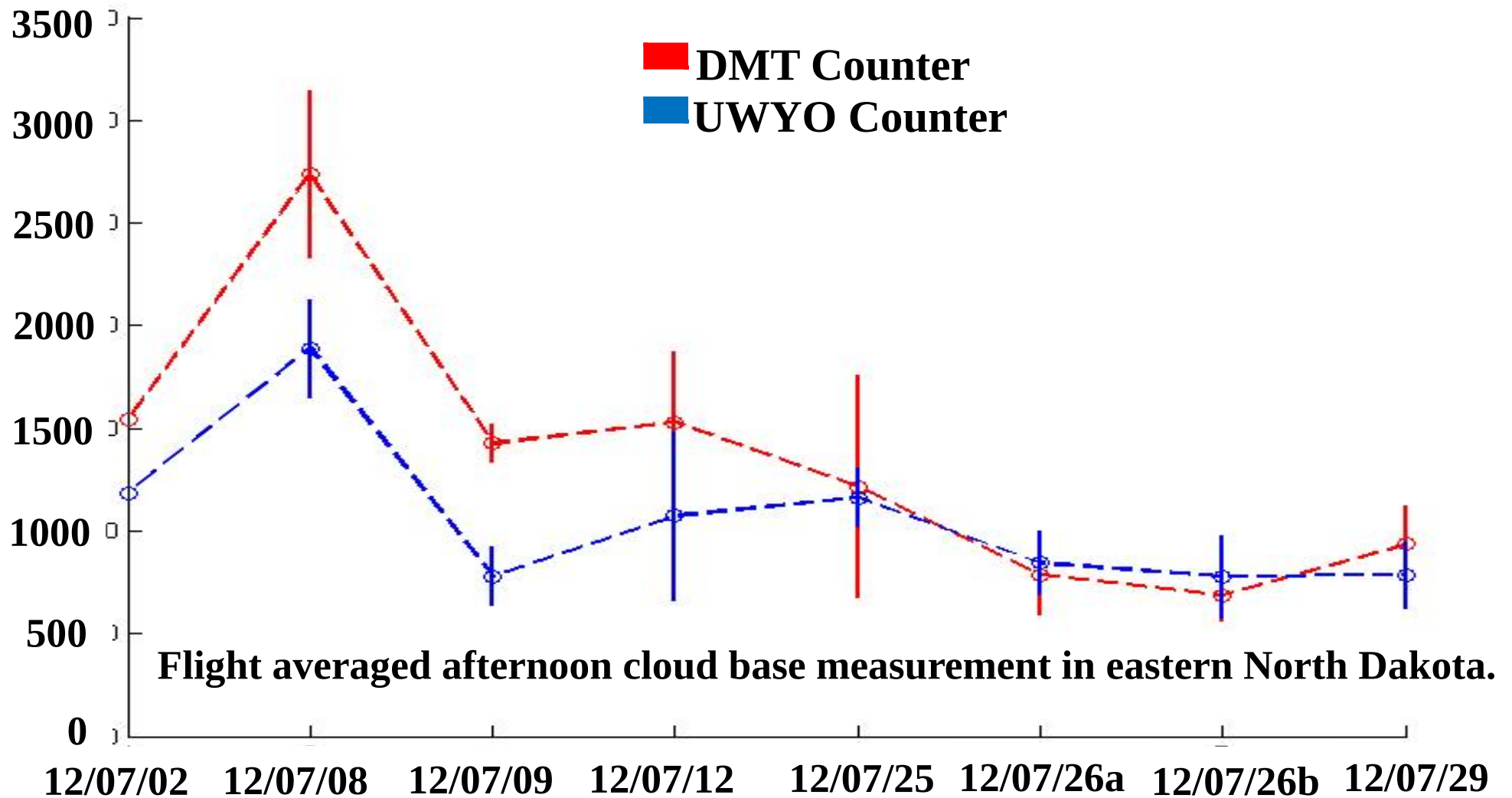




# Comments on Scientific Data

- Quick Visualization of data is very Important.
  - Create a preliminary version of the data using automated processing scripts.
  - Create a final data set after the project is over by applying manual edits to the “raw” data files which replace “bad” data with missing value codes.
- Archive the raw data and any editing files.
- Work with ASCII data as much as possible.
  - Compress ASCII files.
- Use a standard data format, which includes Meta data in all data files.

# Scientist Know Calibrations are Important; However, Software Processing can be just as Important



# References: Digital Data



On-line Binary Hex Converters

- <http://www.binaryhexconverter.com/>

Wikipedia – Serial Port

- [https://en.wikipedia.org/wiki/Serial\\_port#Settings](https://en.wikipedia.org/wiki/Serial_port#Settings)





# References: Data Processing

- Delene, D. J., Andrea Neumann, Alexei Korolev, Matt Freer, Olivier Henry, Jonathan Crosie, Stephanie Gagne, Landan MacDonald, Aaron Bansemer, Andrew Heymsfield, Colin Gurganus, Ted Fisher, Wei Wu, Greg McFarquhar, Towards Community Software Development to Process and Analyze Cloud Physics In-situ Aircraft Data, Bulletin of the American Meteorological Society, in preparation, 2017.
- Delene, D. J., 2016: Airborne Data Processing and Analysis. Source Forge, <http://sourceforge.net/projects/adpaa> (Accessed January 28, 2016).
- Delene, D. J., Airborne Data Processing and Analysis Software Package, Earth Science Informatics, 4(1), 29-44, 2011, URL: <http://dx.doi.org/10.1007/s12145-010-0061-4>, DOI: 10.1007/s12145-010-0061-4.
- Delene, D. J., Suitability of North Dakota for Conducting Effective Hygroscopic Seeding, Journal of Weather Modification, 48, 43-67, 2016.
- Web Site - <https://sourceforge.net/projects/adpaa/>
- Wiki - [http://adpaa.sourceforge.net/wiki/index.php/Main\\_Page](http://adpaa.sourceforge.net/wiki/index.php/Main_Page)

# Flight Data Post-processing

- Conduct all post-flight processing data starting with the M300 flight data file (\*.sea).
  - From 20170404\_190617 directory, process\_all\_kma
- View analog data file in PostProcessing directory,  
cplot2 17\_04\_04\_19\_06\_17.analog2.raw

