CTS-IT

Facilitating science through technology

Christopher P. Barnes (cpb@ufl.edu | @senrabc) , Director
Philip B. Chase (pbc@ufl.edu), Assistant Director
Kevin Hanson (kshanson@ufl.edu), Assistant Director

Team

- We are a team

Informatics Consulting with CTS-IT

CTS-IT staff pictured in 2017: Top row (standing): Christopher Barnes, Tiago Bember Simeão, Prasad Lanka, Lesley Kao, Patrick White, Michael Buchholz, and Jeryl Johnston. Bottom row (sitting): Naomi Braun, Kevin Hanson, Jenny Martinez, Matthew McConnell, Philip Chase, Taryn Stoffs, Marly Cormar, Amber Allen, Miriam Pinedo, and Shikha Mehta

What is our mission

Deliver quality data.

Help researchers with the tools, techniques and project management needed to achieve the highest quality data acquisition, processing, analysis and visualization possible.

Data Management

We provide tools for data mining, analysis and visualization that will best support your research. To manage and integrate your findings and results with data, we make use of existing tools such as content and document management systems. Your results will be compatible with a variety of systems from the moment you collect them.

Grants & Compliance

- We craft informatics components and data sharing plans to strengthen grant proposals by leveraging not only our own strengths, but also those of UF Health. We work with both internal and external-facing web teams to propose data sharing mechanisms that are compliant with all regulations, and as a national leader in the informatics community, we share our work and the work of our clients at national conferences and to collaborating peer institutions.

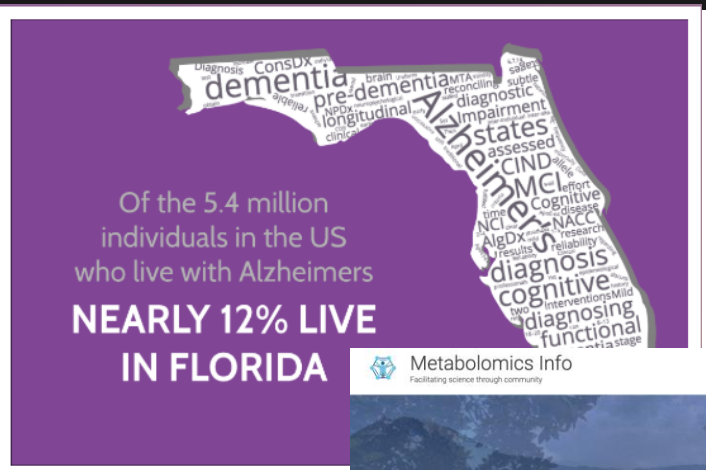
Software & Algorithms

- Our development team designs, implements and maintains custom software and algorithms to meet your scientific requirements. These products are designed to manage and analyze the information you need for your research.

Source: https://www.onespan.com/blog/five-ways-artificial-intelligence-and-machine-learning-can-fight-financial-fraud

Who are some our collaborators

- **Dr. Christopher Cogle** - HemOnc: Early Cancer Detection
- **Dr. Steve DeKosky** - McKnight : Neuroscience Research
- **Dr. Timothy Garrett**- Pathology: Mass Spectroscopy and Metabolomics
- **Dr. Todd Golde** - CTRND : 1 Florida ADRC: Clinical Research in Alzheimer's
- **Ms. Gigi Lipori** – UF IDR : Clinical Data Warehousing
- **Ms. Holly Morris**- CRC/CTSI - Clinical Research Tools.
- **Dr. Krista Vandenborne** - PHHP: DMD/Eli Lilly : Muscular Dystrophy MRI
- **Dr. Rick Yost**- Chemistry : M3C : Metabolomics Consortium



Guiding Principles

- Customer Service
- Collaboration
- Open Standards
- Open Science

Source: https://www.udemy.com/course/machine-learning-masterclass/

Open Standards

LOINC - For lab tests

Logical Observation Identifiers Names and Codes

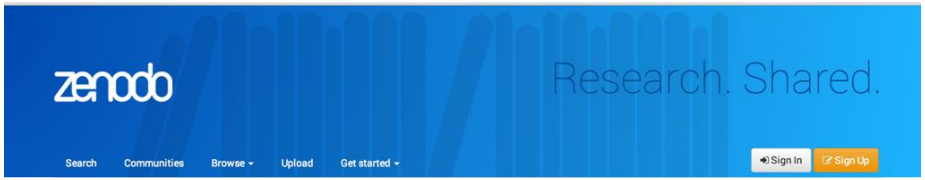
MeSH - Medical Subject Headings

What did you mean by "cat"?

Open Science

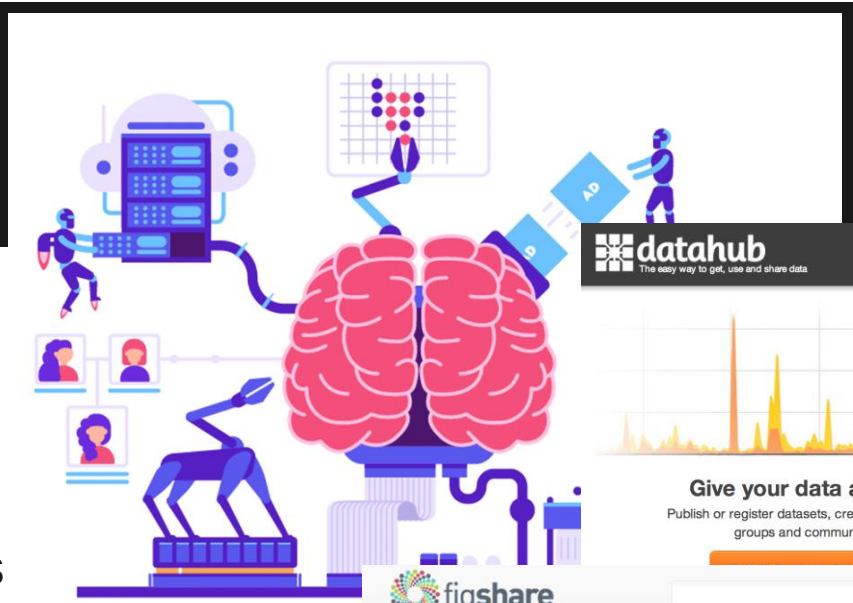
Help researchers REPRODUCIBLY publish data and software

- Open Data - Legal right to use
- DOI - make it citable

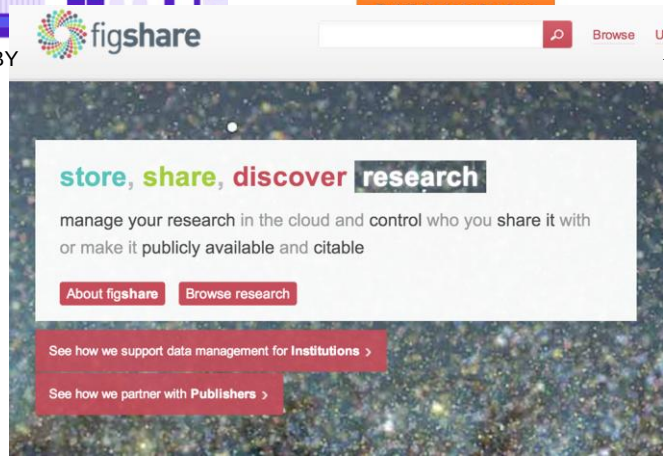


Future

- Help more researchers!
- MACHINE LEARNING & ARTIFICIAL INTELLIGENCE (AI)
- Encourage standards (ontologies and cont.voc. : CDISC, SNOMED, LOINC, UMLS)
- Help people access and use open datasets (use reference data before you have to disambiguate)
- High Quality Data Visualization



SOURCE: DRIBBBLE.COM BY ANDRII BEZVERSHENKO



Project Stories

Solving informatics problems for researchers with REDCap, R, custom coding and bureaucratic wrangling

Philip Chase
Assistant Director, CTS-IT

Solving informatics problems for researchers with REDCap, R, custom coding and bureaucratic wrangling

Philip Chase
Assistant Director

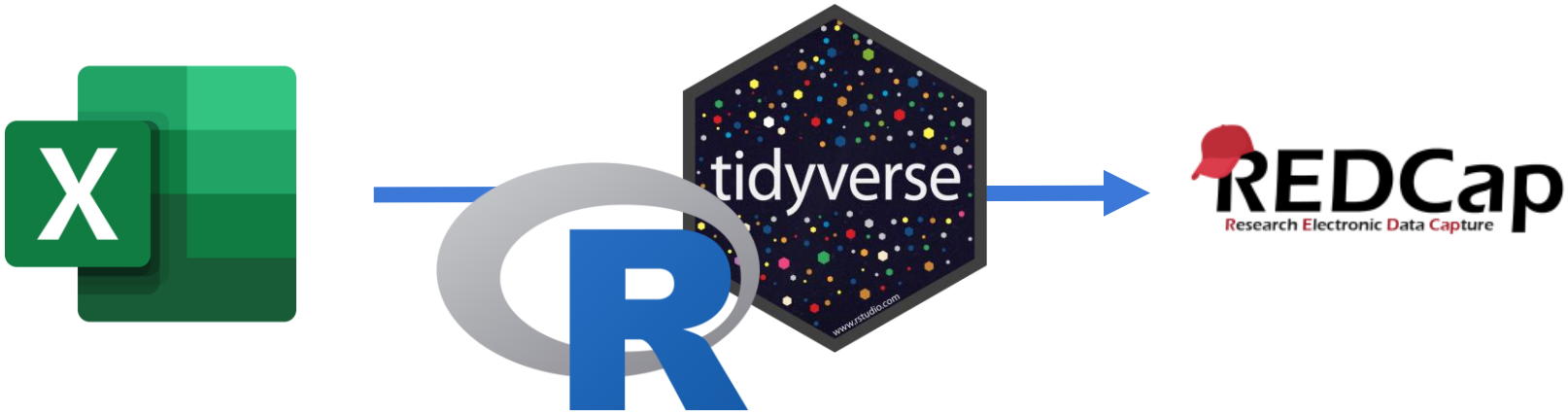
PRISM Registry

Emily Weber, M.D.
UF Department of Obstetrics and Gynecology

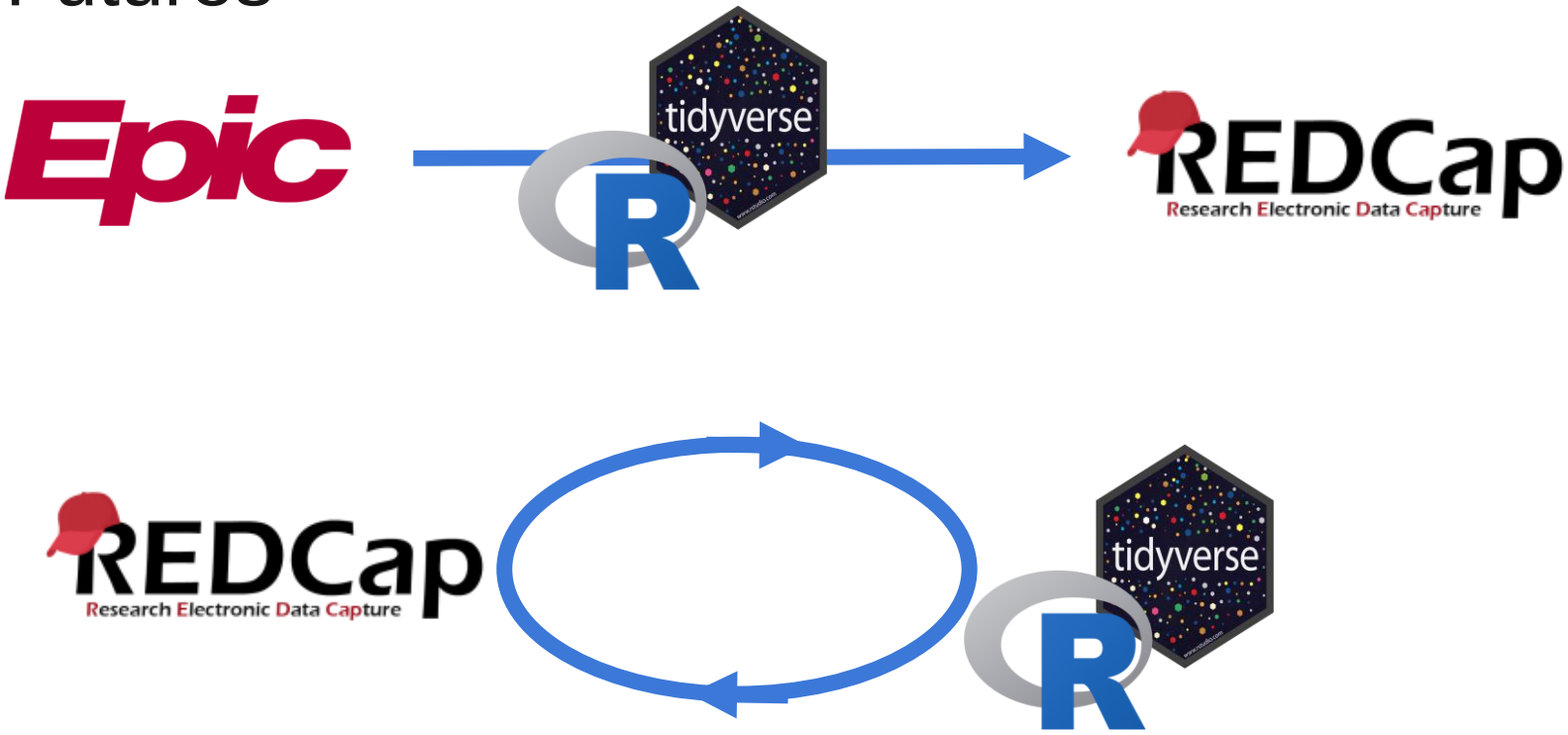
Existing Data

1331 surgeries in an Excel file
Data entry is unwieldy, delayed
Data quality is not enforced
Work is hard to distribute

Solution


tidyverse
www.rstudio.com
REDCap
Research Electronic Data Capture

Futures


Epic
tidyverse
REDCap
Research Electronic Data Capture
REDCap
Research Electronic Data Capture
tidyverse

Risky Behaviors Projects

Robert Leeman, PhD
Associate Professor
Dept of Health Education & Behavior
College of Health & Human Performance

Project design - Easy stuff

Survey research
Longitudinal studies
Daily follow-up
Young population

Why REDCap?

Designed for Clinical Research

Great data entry, survey, and data mgmt

3700 Institutions, 1m users, 7100 citations

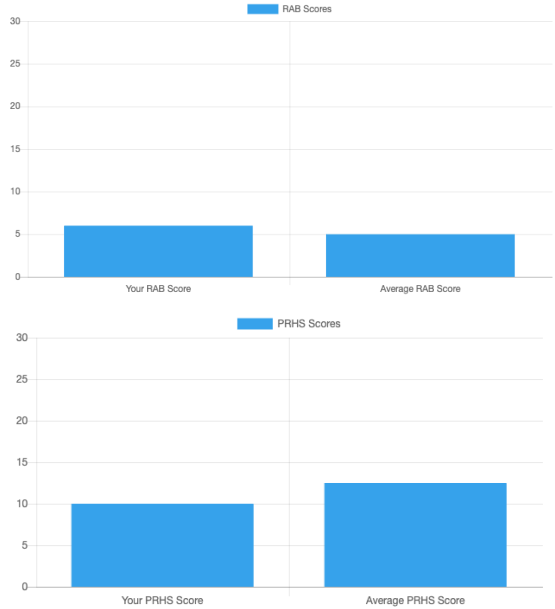
REDCap

Research Electronic Data Capture

Project design - Challenge #1

Normative Behavior with graphical feedback

Module title and description	Date Added	Downloads	
REDCap Chart Field *(redcap_chart_field_v1.1.1)* View on GitHub *Description:* Provides a chart drawing feature for data entries and surveys. Integrates REDCap with third-party chart libraries - currently Charts.js and Chartist are supported. Piping can be used on field configuration, so charts may display facts to survey participants based on their previous answers. *Authors:* Tiago Bember (University of Florida - CTSI), Philip Chase (University of Florida - CTSI), Taryn Stoffs (University of Florida - CTSI), Marly Cormar (University of Florida - CTSI)	2018-08-24	110	You must be a REDCap administrator logged in to REDCap to download modules.



Project design - Challenge #2

Project Stories

- KEVIN HANSON

Matthew Gurka, Ph.D.

Professor, Health Outcomes & Biomedical Informatics

Associate Director, Institute for Child Health Policy

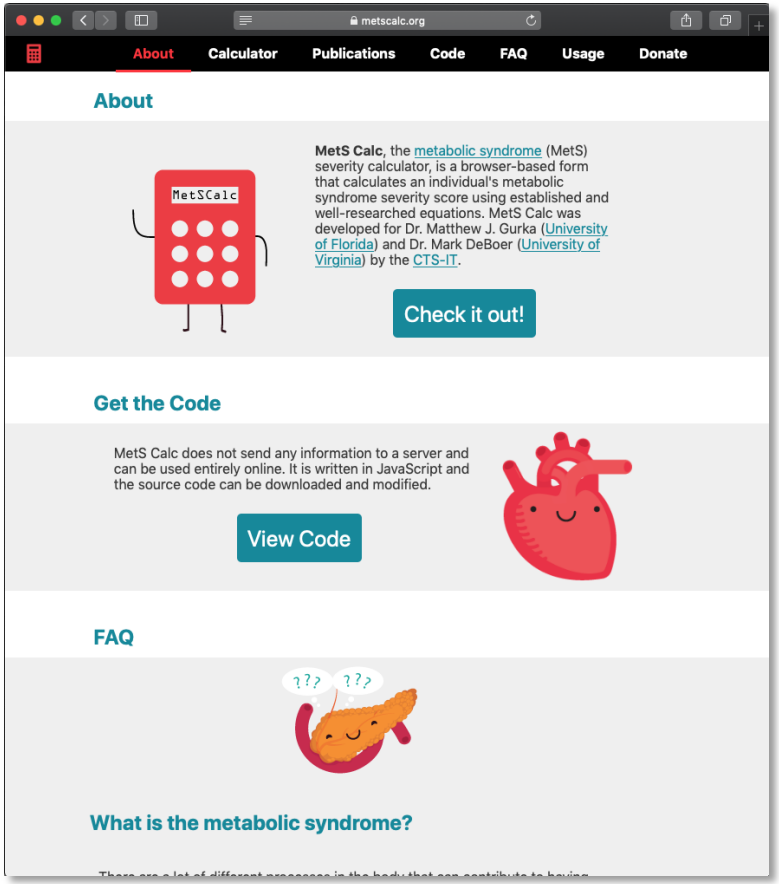
Business Case

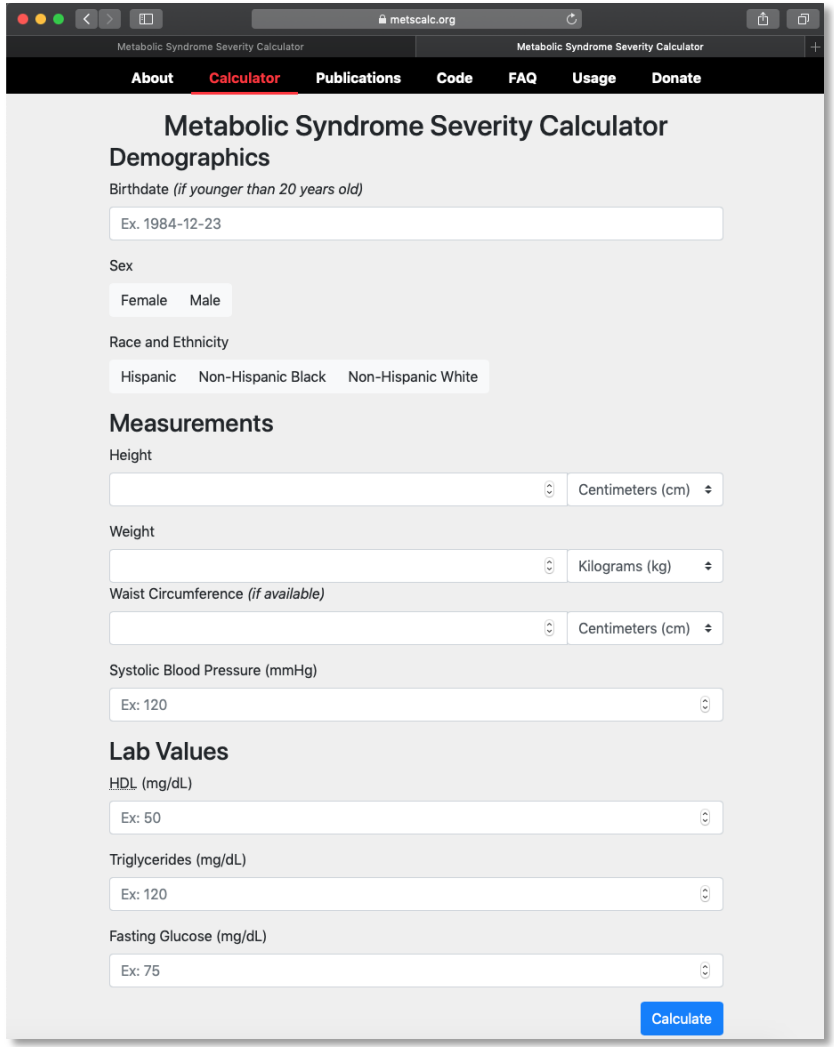
- Existing code for calculator became outmoded through advancing research.
- Changes needed to calculate Metabolic Syndrome Severity Score.
- Previous builds not licensed properly for open-source sharing
- Multi-institutional branding

Solution

- Create GitHub.com organization
- Use GitHub.io free website hosting service
- Register metscalc.org since the project is multi-institutional
- Place all code on GitHub.com with open-source Apache2 license
- Develop JavaScript to reflect latest research to calculate severity score
- Style website to showcase the software, how to collaborate, how to cite the work, and access the calculator

MetSCalc.org





Todd Golde, M.D.
Director, McKnight Brain Institute
Director, 1Florida Alzheimer's Disease Research Center

1Florida ADRC

- New center in 2015
- Clinical site at Mt. Sinai Medical Center in Miami Beach, FL
- Administrative, Data and Neuropath cores in Gainesville, FL

Needs

- Workflows
- Website
- Data management
- Analysis
- Reporting

NACCulator Use Case

- What is the problem we are trying to solve?
  - NACC input format is fixed length (position matters)
  - Need a platform to reliably apply the rule set (500+)
  - Upload data to NACC's data system
- How can we address the problem?
  - Use NACC's web entry system
  - Use SAS
  - **Write software**
- What value is generated from this option?
  - Automated data transfer to NACC
  - Pre-check data issues prior to upload
  - Scheduled data export and upload
  - Modifiable as changes happen

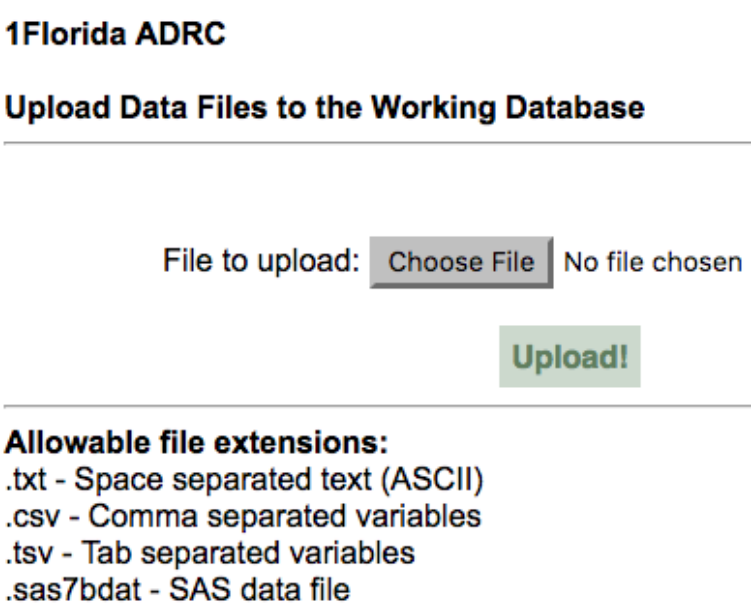
NACCulator

- NACCulator: a translator from CSV to fixed width format
- Written in Python
- https://github.com/ctsit/nacculator

```
def header_fields():
    fields = {}
    fields['PACKET'] = nacc.uds3.Field(name='PACKET', typename='Char', position=(1, 2), l
    fields['FORMID'] = nacc.uds3.Field(name='FORMID', typename='Char', position=(4, 6), l
    fields['FORMVER'] = nacc.uds3.Field(name='FORMVER', typename='Num', position=(8, 10),
    fields['ADCID'] = nacc.uds3.Field(name='ADCID', typename='Num', position=(12, 13), le
    fields['PTID'] = nacc.uds3.Field(name='PTID', typename='Char', position=(15, 24), len
    fields['VISITMO'] = nacc.uds3.Field(name='VISITMO', typename='Num', position=(26, 27)
    fields['VISITDAY'] = nacc.uds3.Field(name='VISITDAY', typename='Num', position=(29, 3
    fields['VISITYR'] = nacc.uds3.Field(name='VISITYR', typename='Num', position=(32, 35)
    fields['VISITNUM'] = nacc.uds3.Field(name='VISITNUM', typename='Char', position=(37,
    fields['INITIALS'] = nacc.uds3.Field(name='INITIALS', typename='Char', position=(41,
```

How it works

- Install NACCulator:
  - pip install nacculator
  - git clone git@github.com:ctsit/nacculator.git
- curl -v -d token=123456 -d content=record -d format=csv -d type=flat https://redcap.ctsi.ufl.edu/redcap/api/ > data.csv
- Run nacculator: redcap2nacc –ivp < data.csv > data.txt
- Upload to NACC website



NACCulator Demonstration

Thank you!

- Thank you for listening!
- Questions?