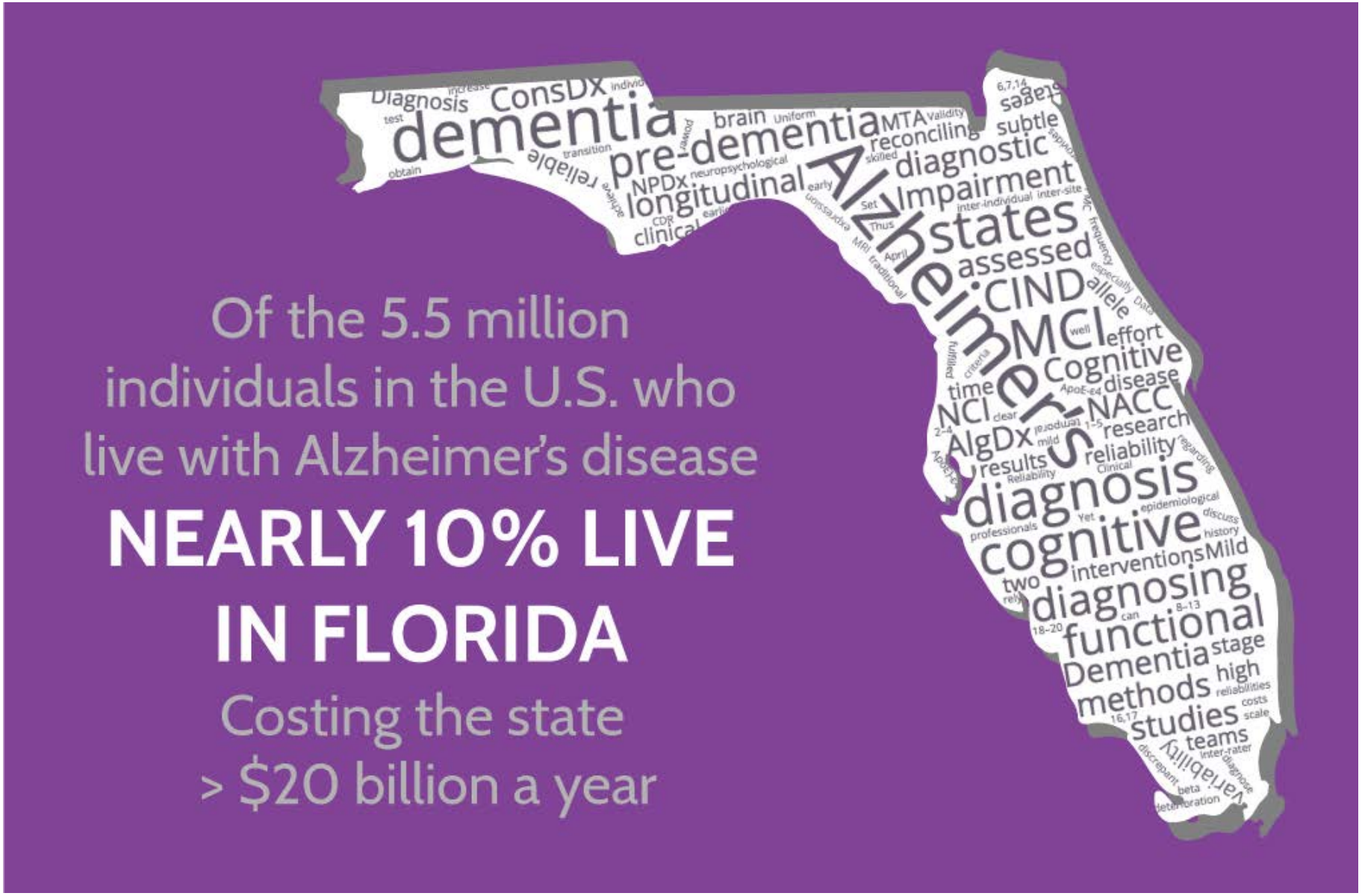


1Florida ADRC

NACCulator:
A translator for CSV files to NACC

1Florida ADRC

Talking Points:

- Introduction
- Business Case
- The Software
- Collaboration

KEVIN HANSON

Co-Investigator

kshanson@ufl.edu

http://1floridaadrc.org

1Florida ADRC

BUSINESS CASE

Business 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 (450+)
  - 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?
  - Transfer data consistently to NACC
  - Pre-check data issues prior to upload
  - Scheduled data export and upload
  - Repeatable programmable way that can be modified as changes happen

1Florida ADRC

Challenges

- Tightly controlled variable definitions
- Data validation rules that may change
- Form changes

1Florida ADRC

THE SOFTWARE

1Florida ADRC

Authors and Designers

- Tarun Akirala
- Christopher Barnes
- Naomi Braun
- Kevin Hanson
- Matthew McConnell
- Ajantha Ramineni
- Taeber Rapczak

1Florida ADRC

NACCulator

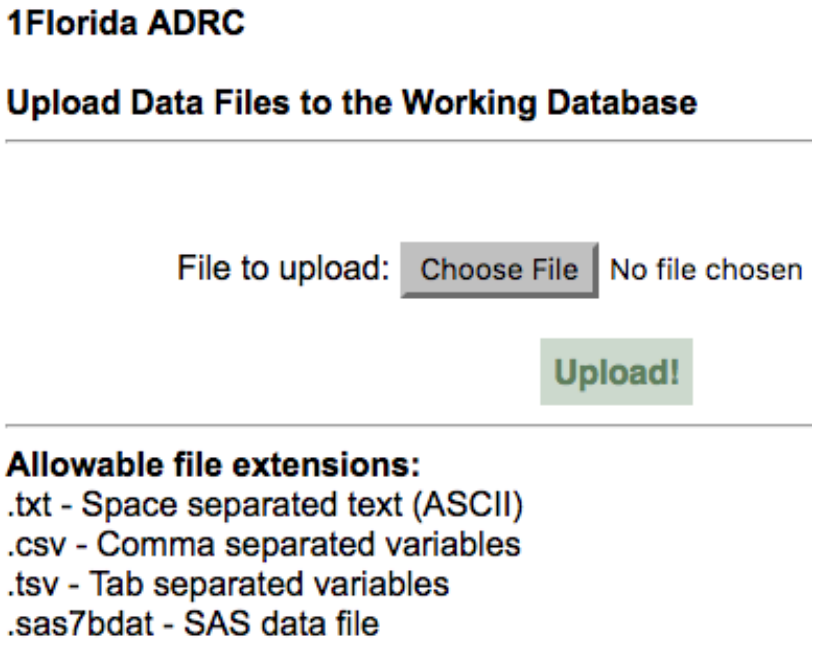
- NACCulator: a translator for CSV to NACC fixed width format
- Written in Python
- https://github.com/ctsit/nacculator

```
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), ler
    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,
```

1Florida ADRC

How it Works

- Install NACCulator: pip install nacculator
- 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



1Florida ADRC

```
$ pip install nacculator
$ redcap2nacc < data.csv > data.txt
```

Or, if you're using the source code:

```
$ PYTHONPATH=. ./nacc/redcap2nacc.py < data.csv > data.txt
```

The program accepts two arguments -file and -(ivp|fvp|np). Both the arguments are optional. See the python help as:

```
$ PYTHONPATH=. ./nacc/redcap2nacc.py -h
usage: redcap2nacc.py [-h]
                      [-fvp | -ivp | -np | -f {cleanPtid,updateField,replaceDrugId,fillDefault,fixC1S}]
                      [-file FILE] [-meta FILTER_META]

Process redcap form output to nacculator.

optional arguments:
-h, --help            show this help message and exit
-fvp                  Set this flag to process as fvp data
-ivp                  Set this flag to process as ivp data
-np                   Set this flag to process as np data
-f or --filter        Accepts one of {cleanPtid,updateField,replaceDrugId,fillDefault,fixC1S}
                      Set this flag to process the filter
-file FILE            Path of the csv file to be processed.
-meta FILTER_META     Input file for the filter metadata (in case cleanPtid is used)
```

```
class FormA1(nacc.uds3.FieldBag):
    def __init__(self):
        self.fields = header_fields()
        self.fields['REASON'] = nacc.uds3.Field(name='REASON', typename='Num', position=(45, 45), length=1, inclusive_range=(1, 4), allowable_values=['4', '2', '1', '9'], blanks=[
        self.fields['REFERSC'] = nacc.uds3.Field(name='REFERSC', typename='Num', position=(47, 47), length=1, inclusive_range=(1, 6), allowable_values=['8', '9', '3', '2', '1', '6
        self.fields['LEARNED'] = nacc.uds3.Field(name='LEARNED', typename='Num', position=(49, 49), length=1, inclusive_range=(1, 4), allowable_values=['3', '2', '1', '9', '8', '4
        self.fields['PRESTAT'] = nacc.uds3.Field(name='PRESTAT', typename='Num', position=(51, 51), length=1, inclusive_range=(1, 3), allowable_values=['3', '2', '1'], blanks=[])
        self.fields['PRESPART'] = nacc.uds3.Field(name='PRESPART', typename='Num', position=(53, 53), length=1, inclusive_range=(1, 2), allowable_values=['2', '1'], blanks=[])
        self.fields['SOURCENW'] = nacc.uds3.Field(name='SOURCENW', typename='Num', position=(55, 55), length=1, inclusive_range=(1, 2), allowable_values=['2', '1'], blanks=[])
        self.fields['BIRTHMO'] = nacc.uds3.Field(name='BIRTHMO', typename='Num', position=(57, 58), length=2, inclusive_range=(1, 12), allowable_values=[], blanks=[])
        self.fields['BIRTHYR'] = nacc.uds3.Field(name='BIRTHYR', typename='Num', position=(60, 63), length=4, inclusive_range=(1875, CURRENT_YEAR-15), allowable_values=[], blanks=
        self.fields['SEX'] = nacc.uds3.Field(name='SEX', typename='Num', position=(65, 65), length=1, inclusive_range=(1, 2), allowable_values=['2', '1'], blanks=[])
        self.fields['HISPANIC'] = nacc.uds3.Field(name='HISPANIC', typename='Num', position=(67, 67), length=1, inclusive_range=(0, 1), allowable_values=['9', '1', '0'], blanks=[]
        self.fields['HISPOR'] = nacc.uds3.Field(name='HISPOR', typename='Num', position=(69, 70), length=2, inclusive_range=(1, 6), allowable_values=['50', '99', '3', '2', '1', '6
        self.fields['HISPORX'] = nacc.uds3.Field(name='HISPORX', typename='Char', position=(72, 131), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Quest
        self.fields['RACE'] = nacc.uds3.Field(name='RACE', typename='Num', position=(133, 134), length=2, inclusive_range=(1, 5), allowable_values=['99', '3', '2', '1', '50', '5',
        self.fields['RACEX'] = nacc.uds3.Field(name='RACEX', typename='Char', position=(136, 195), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Question
        self.fields['RACESEC'] = nacc.uds3.Field(name='RACESEC', typename='Num', position=(197, 198), length=2, inclusive_range=(1, 5), allowable_values=['88', '0', '99', '3', '2'
        self.fields['RACESECX'] = nacc.uds3.Field(name='RACESECX', typename='Char', position=(200, 259), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Qu
        self.fields['RACETER'] = nacc.uds3.Field(name='RACETER', typename='Num', position=(261, 262), length=2, inclusive_range=(1, 5), allowable_values=['88', '99', '3', '2', '1'
        self.fields['RACETERX'] = nacc.uds3.Field(name='RACETERX', typename='Char', position=(264, 323), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Qu
        self.fields['PRIMLANG'] = nacc.uds3.Field(name='PRIMLANG', typename='Num', position=(325, 325), length=1, inclusive_range=(1, 6), allowable_values=['8', '9', '3', '2', '1'
        self.fields['PRIMLANX'] = nacc.uds3.Field(name='PRIMLANX', typename='Char', position=(327, 386), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Qu
        self.fields['EDUC'] = nacc.uds3.Field(name='EDUC', typename='Num', position=(388, 389), length=2, inclusive_range=(0, 36), allowable_values=['99'], blanks=[])
        self.fields['MARISTAT'] = nacc.uds3.Field(name='MARISTAT', typename='Num', position=(391, 391), length=1, inclusive_range=(1, 6), allowable_values=['9', '3', '2', '1', '6'
        self.fields['LIVSITUA'] = nacc.uds3.Field(name='LIVSITUA', typename='Num', position=(393, 393), length=1, inclusive_range=(1, 6), allowable_values=['9', '3', '2', '1', '6'
        self.fields['INDEPEND'] = nacc.uds3.Field(name='INDEPEND', typename='Num', position=(395, 395), length=1, inclusive_range=(1, 4), allowable_values=['3', '2', '1', '9', '4'
        self.fields['RESIDENC'] = nacc.uds3.Field(name='RESIDENC', typename='Num', position=(397, 397), length=1, inclusive_range=(1, 4), allowable_values=['3', '2', '1', '9', '4'
        self.fields['ZIP'] = nacc.uds3.Field(name='ZIP', typename='Char', position=(399, 401), length=3, inclusive_range=(6, 999), allowable_values=[], blanks=['Blank if Question
        self.fields['HANDED'] = nacc.uds3.Field(name='HANDED', typename='Num', position=(403, 403), length=1, inclusive_range=(1, 3), allowable_values=['3', '2', '1', '9'], blanks


class FormA2(nacc.uds3.FieldBag):
    def __init__(self):
        self.fields = header_fields()
        self.fields['INBIRMO'] = nacc.uds3.Field(name='INBIRMO', typename='Num', position=(45, 46), length=2, inclusive_range=(1, 12), allowable_values=['99'], blanks=[])
        self.fields['INBIRYR'] = nacc.uds3.Field(name='INBIRYR', typename='Num', position=(48, 51), length=4, inclusive_range=(1875, CURRENT_YEAR-15), allowable_values=['9999'], b
        self.fields['INSEX'] = nacc.uds3.Field(name='INSEX', typename='Num', position=(53, 53), length=1, inclusive_range=(1, 2), allowable_values=['2', '1'], blanks=[])
        self.fields['INHISP'] = nacc.uds3.Field(name='INHISP', typename='Num', position=(55, 55), length=1, inclusive_range=(0, 1), allowable_values=['9', '1', '0'], blanks=[])
        self.fields['INHISPOR'] = nacc.uds3.Field(name='INHISPOR', typename='Num', position=(57, 58), length=2, inclusive_range=(1, 6), allowable_values=['50', '99', '3', '2', '1'
        self.fields['INHISPOX'] = nacc.uds3.Field(name='INHISPOX', typename='Char', position=(60, 119), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Ques
        self.fields['INRACE'] = nacc.uds3.Field(name='INRACE', typename='Num', position=(121, 122), length=2, inclusive_range=(1, 5), allowable_values=['99', '3', '2', '1', '50',
        self.fields['INRACEX'] = nacc.uds3.Field(name='INRACEX', typename='Char', position=(124, 183), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Ques
        self.fields['INRASEC'] = nacc.uds3.Field(name='INRASEC', typename='Num', position=(185, 186), length=2, inclusive_range=(1, 5), allowable_values=['88', '99', '3', '2', '1'
        self.fields['INRASECX'] = nacc.uds3.Field(name='INRASECX', typename='Char', position=(188, 247), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Qu
        self.fields['INRATER'] = nacc.uds3.Field(name='INRATER', typename='Num', position=(249, 250), length=2, inclusive_range=(1, 5), allowable_values=['88', '99', '3', '2', '1'
        self.fields['INRATERX'] = nacc.uds3.Field(name='INRATERX', typename='Char', position=(252, 311), length=60, inclusive_range=None, allowable_values=[], blanks=['Blank if Qu
        self.fields['INEDUC'] = nacc.uds3.Field(name='INEDUC', typename='Num', position=(313, 314), length=2, inclusive_range=(0, 36), allowable_values=['99'], blanks=[])
        self.fields['INRELTO'] = nacc.uds3.Field(name='INRELTO', typename='Num', position=(316, 316), length=1, inclusive_range=(1, 6), allowable_values=['3', '2', '1', '6', '5',
        self.fields['INKNOWN'] = nacc.uds3.Field(name='INKNOWN', typename='Num', position=(318, 320), length=3, inclusive_range=(0, 120), allowable_values=['999'], blanks=[])
        self.fields['INLIVWTH'] = nacc.uds3.Field(name='INLIVWTH', typename='Num', position=(322, 322), length=1, inclusive_range=(0, 1), allowable_values=['1', '0'], blanks=[])
        self.fields['INVISITS'] = nacc.uds3.Field(name='INVISITS', typename='Num', position=(324, 324), length=1, inclusive_range=(1, 6), allowable_values=['3', '2', '1', '6', '5'
        self.fields['INCALLS'] = nacc.uds3.Field(name='INCALLS', typename='Num', position=(326, 326), length=1, inclusive_range=(1, 6), allowable_values=['3', '2', '1', '6', '5',
```

1 Florida ADRC

COLLABORATION

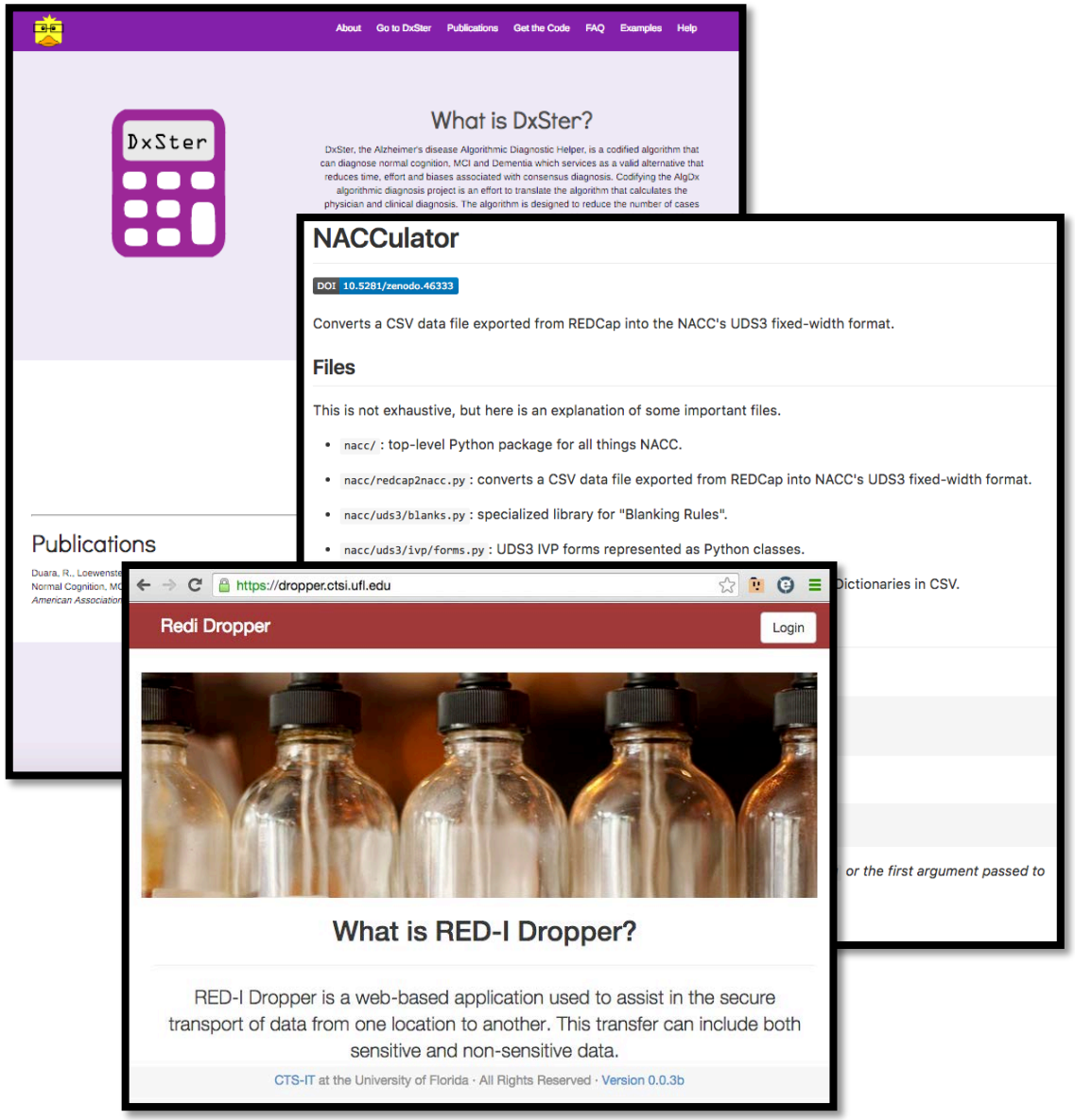
1 Florida ADRC

Open Source Software

- DxSter for algorithmic diagnosis of Alzheimer's disease
  https://ctsit.github.io/dxster/

- NACCulator for REDCap to NACC data transfer
  https://github.com/ctsit/nacculator

- Dropper for multi-site large file transfer
  https://github.com/ctsit/redi-dropper-client



1Florida ADRC

Acknowledgements

- Kansas ADC
  - Jonathan Mahnken

  - Suzanne Hunt
- Michigan ADC
  - Hiroko Dodge

  - Sherry Teboe
  - Sean Ma
  - Ari Bhaumik

1Florida ADRC

If we don’t share our work we are doomed to reinvent it

Thank You