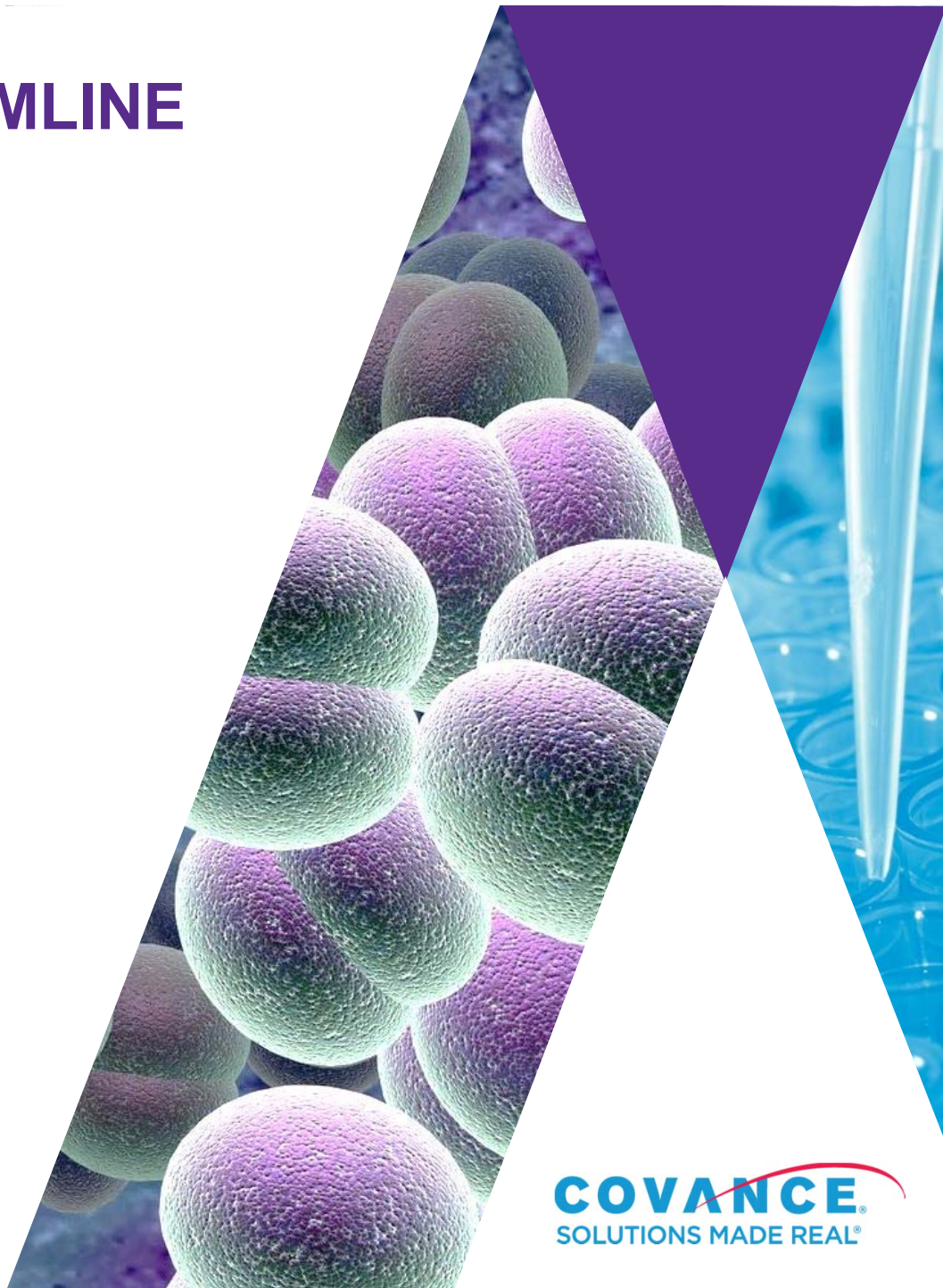


SIMPLIFY AND STREAMLINE USING PYTHON

October 25th, 2018



NORTH CAROLINA • OCT. 25, 2018

Copyright © 2018 Covance. All Rights Reserved

Public

COVANCE
SOLUTIONS MADE REAL®

Michael Stackhouse

BIOGRAPHY

Michael Stackhouse holds a bachelor's degree from Arcadia University where he studied Business Administration, Economics, and Statistics. He is currently an MIDS student at UC Berkeley School of Information, studying Data Science.

Mr. Stackhouse has extensive CDISC experience, working with both SDTM and ADaM standards as well as serving as a subject matter expert for Define.xml. Mr. Stackhouse's favorite career projects include macro and utility development, as well as molding the minds of young programmer Padawans into the programming Jedis of the future.

Mr. Stackhouse is currently a Manager of Statistical Programming at Covance and father to three young, furry, four-legged children.



About Covance

Covance Inc., the drug development business of Laboratory Corporation of America Holdings (LabCorp), is the world's most comprehensive drug development company.

We are dedicated to advancing healthcare and delivering Solutions Made Real[®] by providing high quality nonclinical, clinical, commercialization and informatics services to pharmaceutical and biotechnology companies to help increase the speed, precision and effectiveness of drug development.

Covance is currently working on about half of all active clinical trials globally and has been involved in all top 50 best-selling drugs on the market.



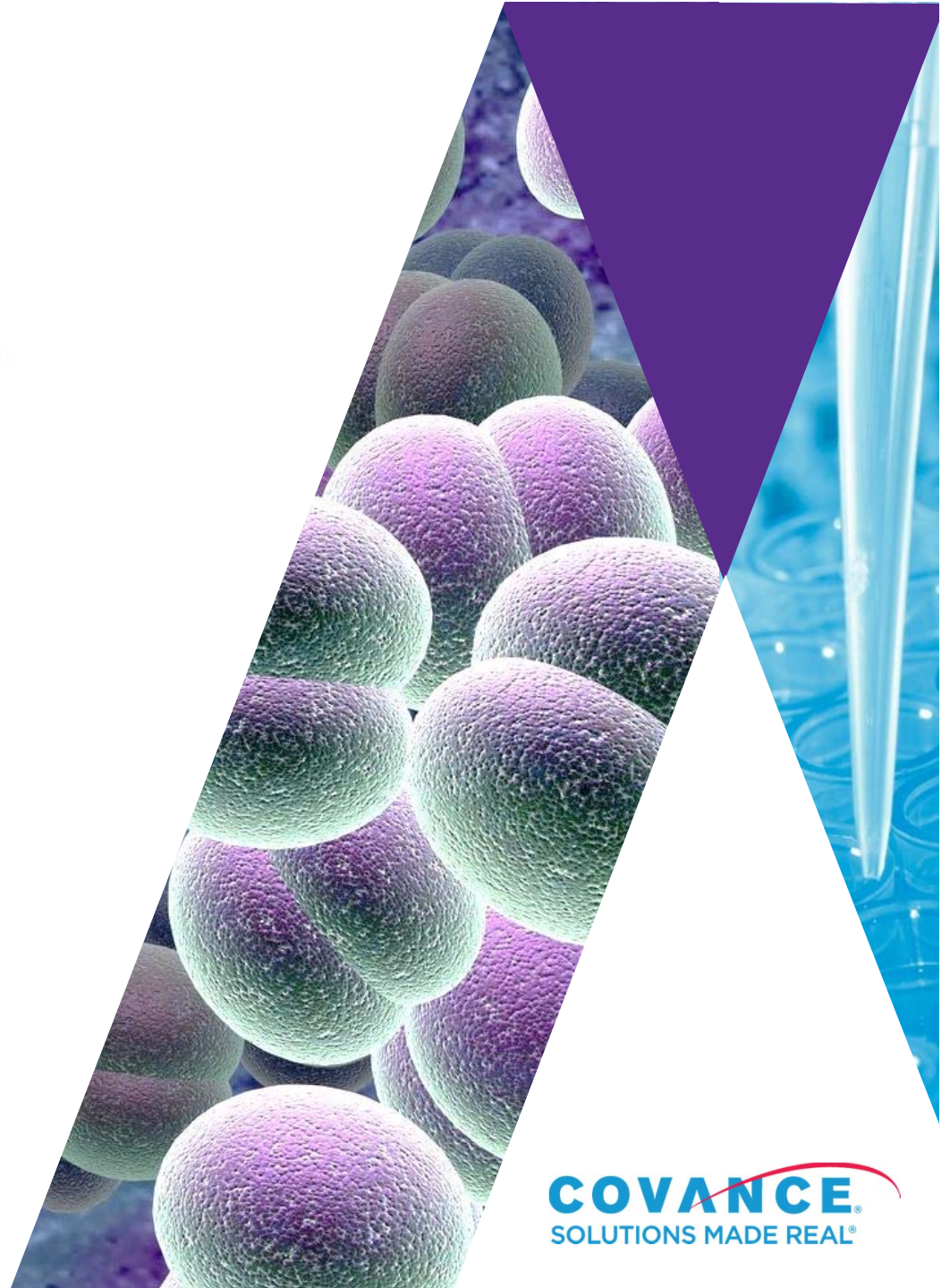
Why Python?



NORTH CAROLINA • OCT. 25, 2018

4 | Simplify and Streamline Using Python, October 25th, 2018

Public



Why Python?

- ▶ Simple, Clean, Efficient
- ▶ System Operations are very straightforward
- ▶ Quick development time for otherwise complex operations
- ▶ OS Independent
- ▶ Distribute programs to team members without access to SAS
- ▶ Free!



NORTH CAROLINA • OCT. 25, 2018

5 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 1: Rename SAS Programs to Text Files (sas2text.py)

- ▶ eTCG Section 4.1.2.10 Software Programs
 - Sponsors should submit software programs in ASCII text format; however, executable file extensions should not be used.
- ▶ Change the extension on all `.sas` files to `.txt`
- ▶ Benefits of Python
 - Simple to iterate over files
 - Easily replace files in all subfolders
 - Very concise (9 lines of code)



NORTH CAROLINA • OCT. 25, 2018

6 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 2: Build a hyperlinked TOC of programs (program_toc.py)

- ▶ Assists reviewers in finding the programs of interest
- ▶ Easily open programs right from the spreadsheet
- ▶ Automatically list all macros called within a program to aid review
- ▶ Leverage data available data sources
 - Uses ADaM metadata files available to find variable labels
 - Uses TLF metadata sources to link titles to program names



NORTH CAROLINA • OCT. 25, 2018

7 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 2: Build a hyperlinked TOC of programs (program_toc.py)

► Why automate?

- Offers a sense of validation (are all programs present? All outputs?)
- Recreate easily with new or updated information
- Avoid manual mistakes



NORTH CAROLINA • OCT. 25, 2018

8 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 3: Comparing Spec Files (spec_compare.py)

- ▶ Useful to find needles in haystack
 - Evaluate minor changes in two versions of spec
 - Compare two sister studies to look for inconsistencies
- ▶ Easily summarize all findings into one clear spreadsheet with only relevant information
- ▶ Useful close to submission
 - Aids creation of define.xml by catching minor changes that may otherwise be missed



NORTH CAROLINA • OCT. 25, 2018

9 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 4: Extract SAS code (without macros) from SAS Logs (ConvertLogDir.py)

► Goal:

- Read in SAS log and output an executable SAS program
- Resulting SAS program resembles the original program but with all macros expanded into Base SAS syntax

► First question – Why???

- Reviewer may not know SAS
- Macros can be complicated, nested, and hard to digest – even for experienced programmers



NORTH CAROLINA • OCT. 25, 2018

10 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 4: Extract SAS code (without macros) from SAS Logs (ConvertLogDir.py)

- ▶ From the eTCG section 4.1.2.10:
 - The main purpose of requesting the submission of these programs is to understand the process by which the variables for the respective analyses were created and to confirm the analysis algorithms.
- ▶ (You can probably avoid these questions by submitting Analysis Results Metadata)



NORTH CAROLINA • OCT. 25, 2018

11 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 4: Extract SAS code (without macros) from SAS Logs (ConvertLogDir.py)

► Process:

- Run all SAS programs and ensure that MPRINT is turned on everywhere
- Step 1: Toss out any autoexec process prior to the start of the program
- Step 2: Identify all program lines
 - Pull out any base SAS or MPRINT lines
- Step 3: Determine which lines to write
 - Exclude any in program written macros (even macro-caption)
- Step 4: Clean up the lines and write out to the SAS file
 - For base SAS, trim the log line numbers
 - For MPRINT lines, take the text after the colon



NORTH CAROLINA • OCT. 25, 2018

12 | Simplify and Streamline Using Python, October 25th, 2018

Public



Example 4: Extract SAS code (without macros) from SAS Logs (ConvertLogDir.py)

- ▶ Secondary benefit – recover original SAS code
 - Ever lose a program, but have a log? It happens
 - Program can also ignore MPRINT syntax and restore the original SAS code
- ▶ Why Python?
 - Can it be done in SAS? Yes, but with Python...
 - Text processing is more straightforward
 - Easily done with a double click or a command
 - Processes a program in a single pass



NORTH CAROLINA • OCT. 25, 2018

13 | Simplify and Streamline Using Python, October 25th, 2018

Public



Conclusion

▶ Python can help you:

- Automate otherwise tedious activities
- Find changes within documents that could easily be overlooked
- Dynamically create well formatted excel sheets
- Save you hours of program preparation from unexpected requests

▶ See other benefits of Python!

- <https://www.pharmasug.org/proceedings/2018/AD/PharmaSUG-2018-AD12.pdf>



Contact Information

Name: Michael Stackhouse

Organization: Covance, Inc.

Address: 4000 Centergreen Way # 300

City, State ZIP: Cary, NC 27513

E-mail: Michael.Stackhouse@Chiltern.com

Web: <https://www.linkedin.com/in/michael-s-stackhouse/>
<https://github.com/mstackhouse>



Covance Inc., headquartered in Princeton, NJ, USA, is the drug development business of Laboratory Corporation of America Holdings (LabCorp). COVANCE is a registered trademark and the marketing name for Covance Inc. and its subsidiaries around the world.



15 | Simplify and Streamline Using Python, October 25th, 2018

Public

