

# The People behind AI

ConnectAI 2022 Digital Masterclass

“Working with a black box: Reproducibility and quality assurance in AI”

**Natasa Milic-Frayling, Intact Digital**

**26 January 2022, 13:00–13:45 UTC+1**

# Working with a black box: Reproducibility and quality assurance in AI

*“AI systems are complex. Their performance depends on intricate interactions of parameter settings and data characteristics that are practically intractable. They are black boxes for all practical purposes.”*

Dr Natasa Milic-Frayling  
CEO & Founder, Intact Digital

# Topics

## COMPLEXITY OF AI SYSTEMS

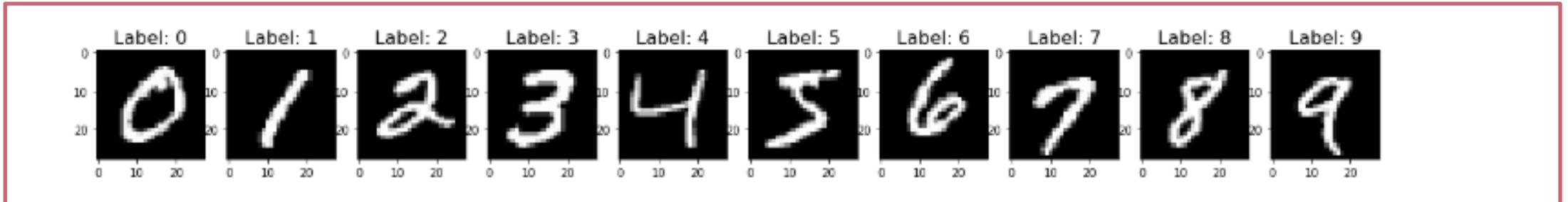
- AI Project Example: Single digit recognition
- Modularity and reusability: Use of libraries and data resources

## REPRODUCIBILITY & AI ASSURANCE

- Stability of systems configurations and performance
- Reproducibility: Quality evidence and assurance

# AI programming

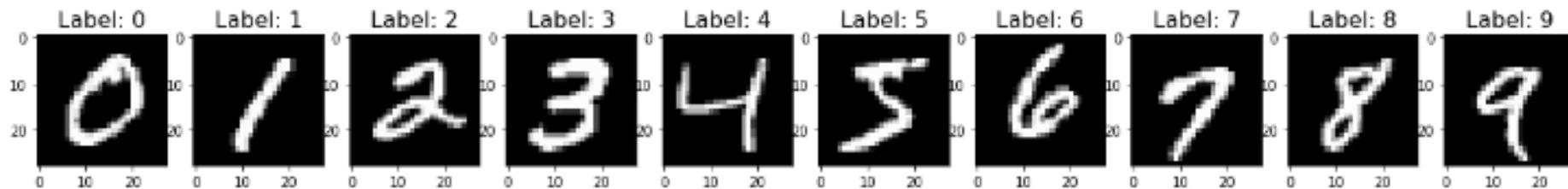
- Problem: Train a neural network model to recognize single digits





# AI programming: Modular and reusable

- Problem: Train a neural network model to recognize single digits



```
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline

import keras
from keras.models import Sequential
from keras.layers import Dense, Dropout
```

```
from sklearn.metrics import confusion_matrix
import seaborn as sns
import tensorflow as tf

np.random.seed(0)

from keras.datasets import mnist
```

Mnist\_Digits\_Classification.ipynb - x

colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=2hLSCxKRk...

Mnist\_Digits\_Classification.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text Reconnect Editing

### Create Model - fully Connected Neural Network

```

model = Sequential()

model.add(Dense(units=128, input_shape=(784,), activation='relu'))
model.add(Dense(units=128, activation='relu'))
model.add(Dropout(0.25))
model.add(Dense(units=10, activation='softmax'))

model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
model.summary()

```

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 128)	100480
dense_1 (Dense)	(None, 128)	16512
dropout (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 10)	1290

=====  
Total params: 118,282  
Trainable params: 118,282  
Non-trainable params: 0  
=====

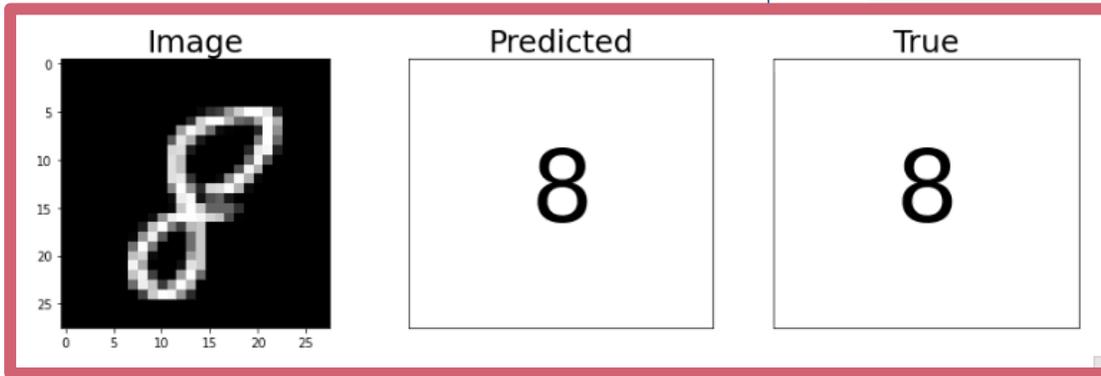
0s completed at 1:24 AM

Feed Forward Neural Network

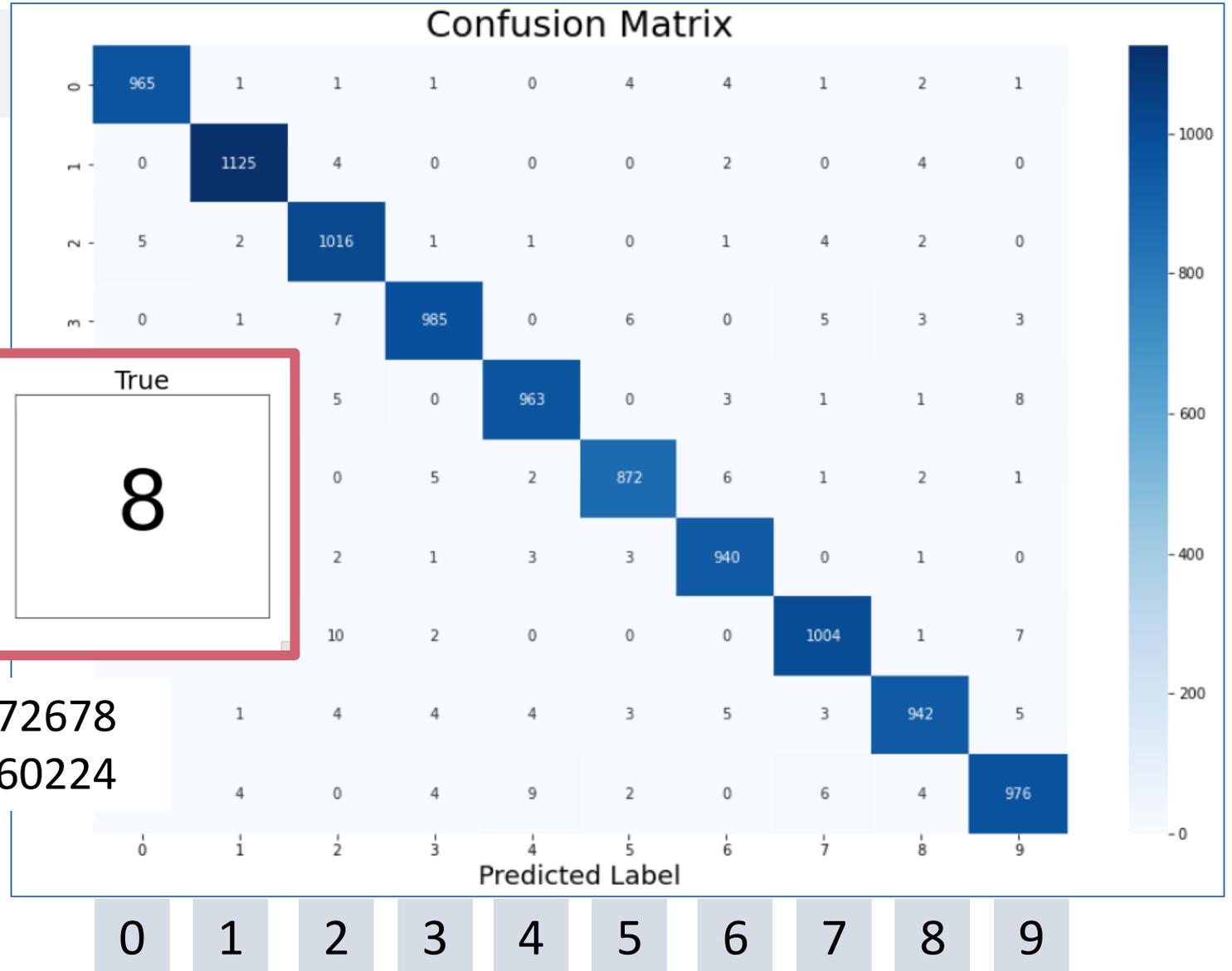
4 layers, including the Dropout

Cross Entropy Loss function with Adam optimizer

# Performance



Test Loss: **0.068174876272678**  
 Test Accuracy: **0.978799998760224**



# Quality: Consistent performance

- Application: Pre-trained model used to recognize US postcodes
- Issue: Hand-written digits are significantly different and system underperforms.

A

MODEL TRAINED AND TESTED ON DATASET A



B

NEW DATASET B



# Quality assurance: Remediation of the problem

## Practical approaches

- Use the same model structure and retrain on additional labelled data from the new data set
- Change the structure of the model, e.g., number of layers
- Change the representation of the images (i.e., features) used to represent the content.

## Knowledge acquisition

- Has anyone solved the problem on a different data set?
- Has anyone solved the problem in a different domain?
- Is the problem known as being intractable at this point?
- What would make it possible to solve the problem?

At which point is it necessary to dig deeper into the building blocks and start building new ones?

# AI Assurance: Quality evidence

---

Government interventions

# EU Commission: Rules and actions for excellence & trust in AI

Press release | 21 April 2021 | Brussels

## Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence

Page contents

[Top](#)

[Print friendly pdf](#)

[Related media](#)

[Press !\[\]\(24ebf582a58af7318d9e75a2b147597b\_img.jpg\) contact](#)

The Commission proposes today new rules and actions aiming Europe into the global hub for trustworthy Artificial Intelligence (combination of the first-ever [legal framework on AI](#) and a new [Coordinated Plan with Member States](#) will guarantee the safety fundamental rights of people and businesses, while strengthening uptake, investment and innovation across the EU. New rules or [Machinery](#) will complement this approach by adapting safety ru increase users' trust in the new, versatile generation of product:

*"The combination of the first-ever legal framework on AI and a new Coordinated Plan with Member States will guarantee the safety and fundamental rights of people and businesses"*

[Europe fit for the Digital Age: Artificial Intelligence \(europa.eu\)](#)

21 APRIL 2021

# United Kingdom: AI Assurance through AI Governance

Independent report

## The roadmap to an effective AI assurance ecosystem

Published 8 December 2021

*"The UK intends to establish the most trusted and pro-innovation system for AI governance in the world"*

[The roadmap to an effective AI assurance ecosystem - GOV.UK \(www.gov.uk\)](https://www.gov.uk)  
8 DECEMBER 2021

Contents

Ministerial foreword

The roadmap to an effective AI assurance ecosystem

 [Print this page](#)

### Ministerial foreword

Artificial intelligence (AI) is one of the game-changing developments of our time. From powering innovative businesses of all sizes across the length and breadth of the UK, to enabling trailblazing research into some of the greatest scientific challenges of our time, AI offers a world of transformative potential. As AI becomes an increasingly important driver of economic and social progress, we need our government to keep up. Getting governance right will create the trust that will drive AI adoption and unlock its full potential. As we announced in the National AI Strategy, the UK intends to establish [the most trusted and pro-innovation system for AI governance in the world](#).

# UK AI Assurance Focus

## The roadmap to an effective AI assurance ecosystem

Artificial intelligence (AI) offers transformative opportunities for the economy and society, but these benefits will only be realised if **organisations, users and citizens can trust AI systems and how they are used.**

The roadmap to an effective AI assurance ecosystem - GOV.UK ([www.gov.uk](http://www.gov.uk))

## AI assurance provides the tools to build trust and ensure trustworthy adoption

Assurance is about building confidence or trust in something, for example a system or process, documentation, a product or an organisation.

**Assurance services help people to gain confidence in AI systems by evaluating and communicating reliable evidence about their trustworthiness.**

# UK AI Assurance Audits and Regulations

To provide meaningful and reliable assurance for AI, organisations need to overcome:

- An **information problem**: **reliably evaluate evidence** to assess whether an AI system is trustworthy.
- A **communication problem**: communicate the evidence at the right level, to inform assurance users' views on whether to trust an AI system.

[The roadmap to an effective AI assurance ecosystem - GOV.UK \(www.gov.uk\)](https://www.gov.uk)

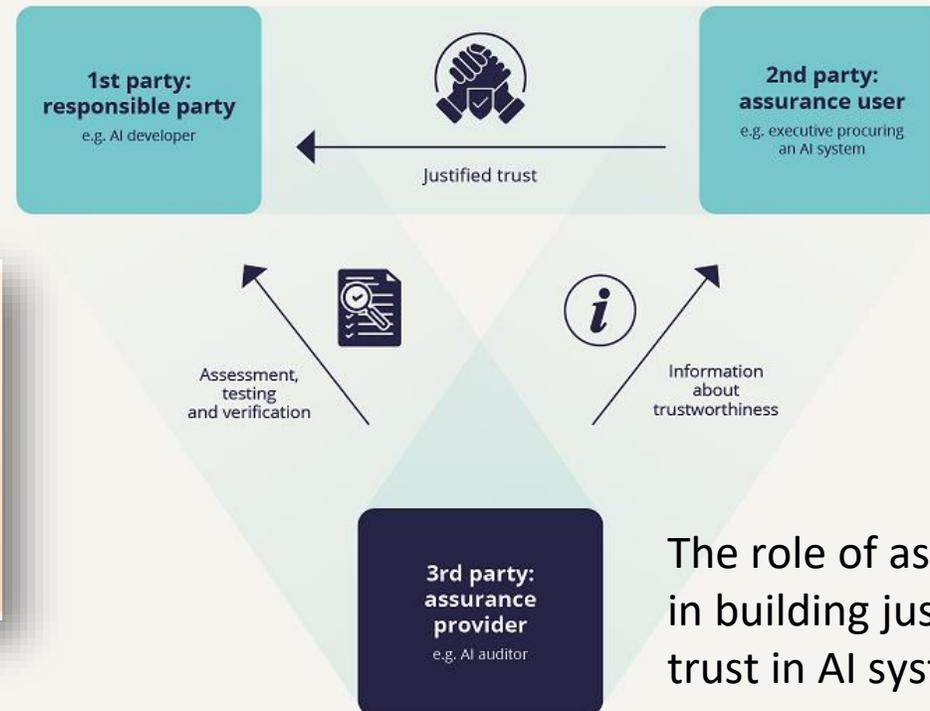
# United Kingdom: AI Assurance services



Mechanisms for assuring AI systems



The roadmap to an effective AI assurance ecosystem - GOV.UK ([www.gov.uk](http://www.gov.uk))



The role of assurance in building justified trust in AI systems

# Life Sciences: Reproducibility as a regulatory requirement

---

Pre-requisite: computer system validation

## Regulated sectors: Computing System Validation

- **Computer System Validation** (CSV) is a process used to ensure (and document) that a computer-based system will produce information or data that meets a set of predefined requirements.
  - Organizations are required to keep the system in the validated states at all times.
- CSV includes a **technology retirement** strategy, i.e., the plans for system decommissioning and data management.
- CSV and system decommissioning are documented and documented steps are strictly followed.

# Regulated sectors: Long-term data retention

## A

### Attributable

Document must clearly identify who has created and contributed to them, and be protected against falsification or forgery of those details

## L

### Legible

Stored documentation must be legible and easy to read

## C

### Contemporaneous

Documentation should demonstrate and support contemporaneous record-keeping

## O

### Original

Storing original copies of documentation guarantees accuracy and confidentiality

## A

### Accurate

The processes and procedures by which companies record and keep their documentation up to date must ensure accuracy and reliability

## +Complete

All documentation must have an audit trail to show no data has been deleted or lost

## +Consistent

Documentation must be date and time stamped and stored in such a way to prove it has been assembled in the expected sequence

## +Enduring

Data must be available for as long as the regulation requires

## +Available

## ALCOA+

### Data integrity requirements

- Raw data used in pre-clinical and clinical research must be retained for decades after the studies are completed
- Organizations must comply with Good Laboratory Practices (GLP) and able to **re-construct the studies from raw data**
- Organizations are regularly audited for compliance with GLP regulations

# Software integrity: Long-term reliable use of software

- Issue: Software is subject to rapid obsolescence if not regularly updated
- Using alternative software cannot guarantee the same output as the original software version.

## DATA INTEGRITY

Extent to which data is complete, consistent and accurate throughout the data lifecycle from data collection to analysis and archiving

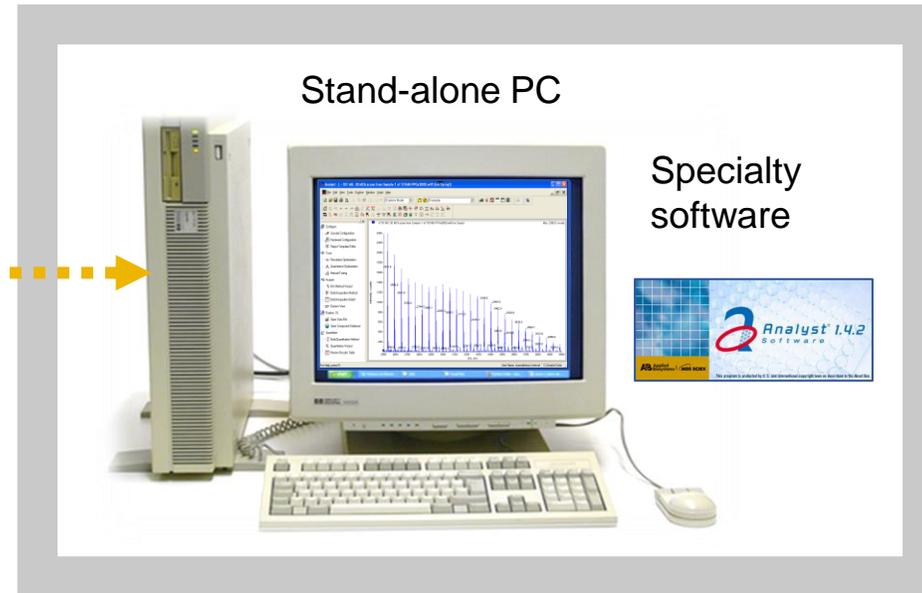
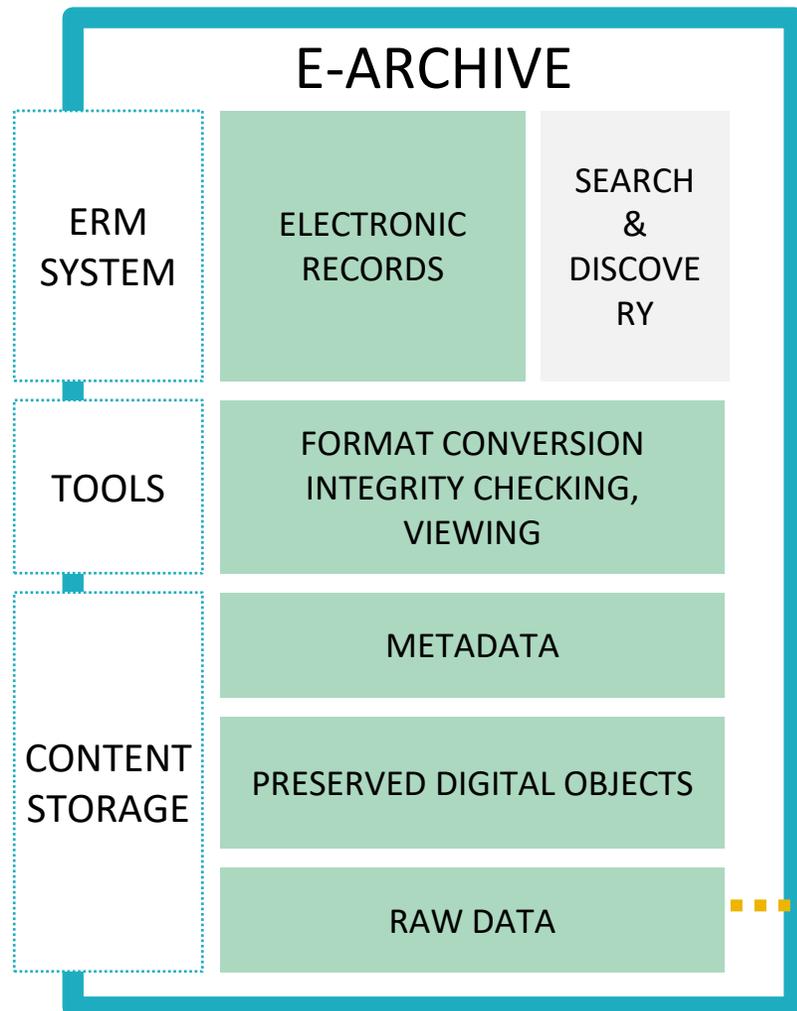
- ✓ **Complete**  
All the required data needs to be collected and stored
- ✓ **Unaltered**  
Data must not be changed
- ✓ **Secure**  
Data must not be destroyed
- ✓ **Confidential**  
Data must not be disclosed to unauthorised individuals
- ✓ **Usable**  
Study metadata needs to provide context for experts reproducing study results

## SOFTWARE INTEGRITY

Extent to which a software installation is functional, reliable, and usable throughout its lifecycle from operational use to archived study reproducibility.

- ✓ **Functional**  
Software installation must stay operational
- ✓ **Unaltered**  
Software must not be changed
- ✓ **Secure**  
Software installation must not present risks from cyber attacks and confidentiality breaches
- ✓ **Accessible**  
Software installation needs to be easily accessed for repeated use

# Technology obsolescence impact

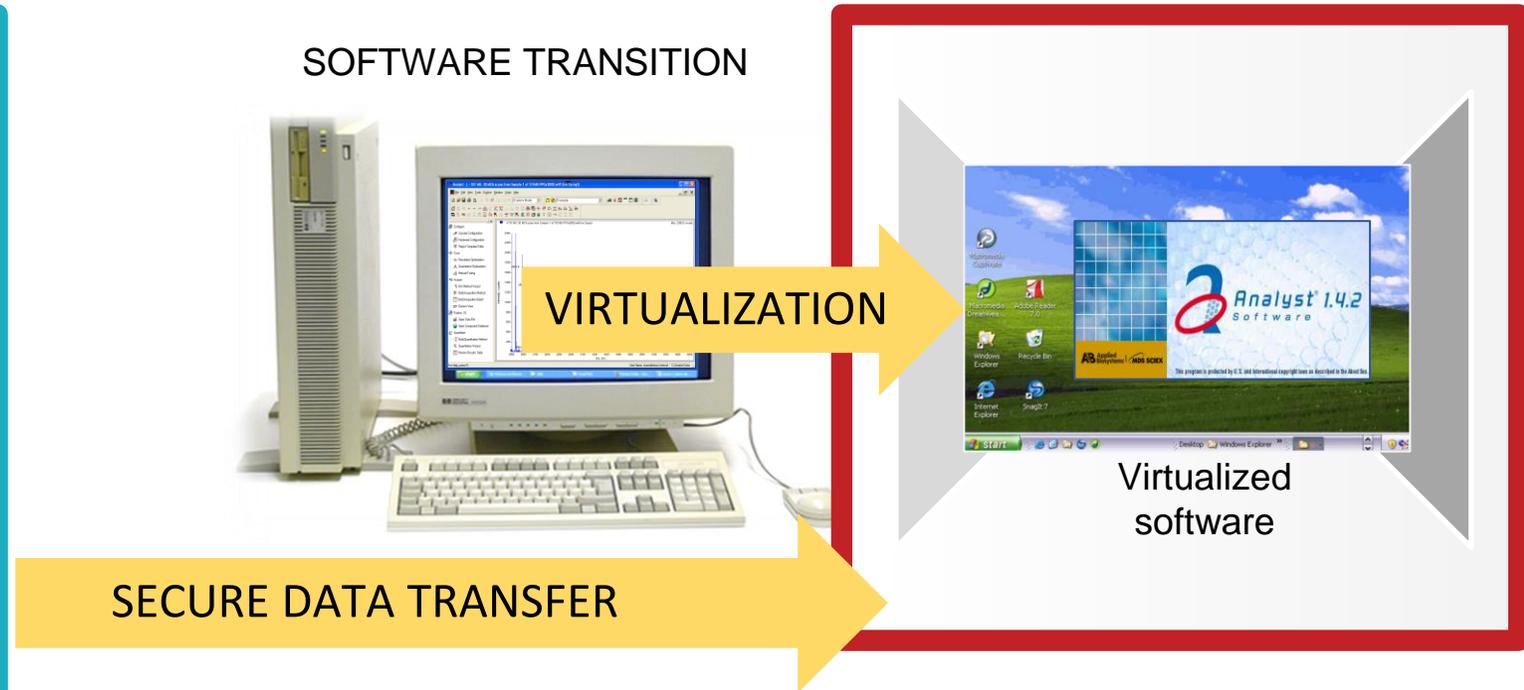
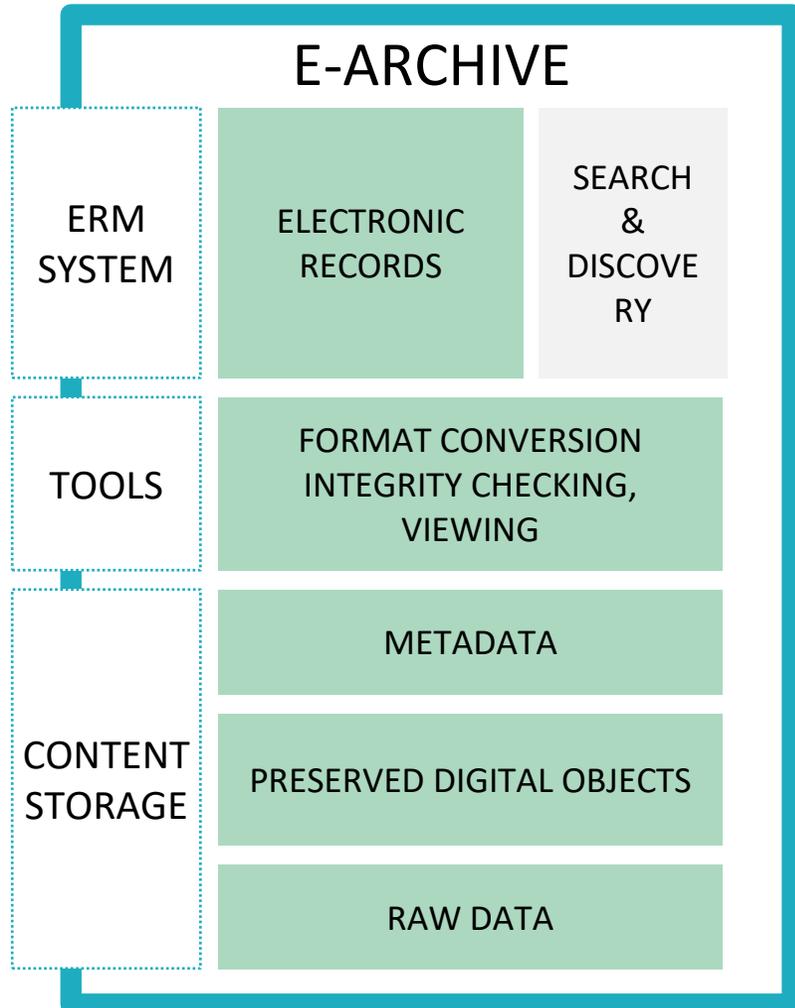


PHYSICAL  
DATA  
TRANSFER

## Technology obsolescence

- Unsupported PC hardware
- Unsupported operating system (Windows XP, Windows 7)
- Unsupported application
- Compatible software may not be available.

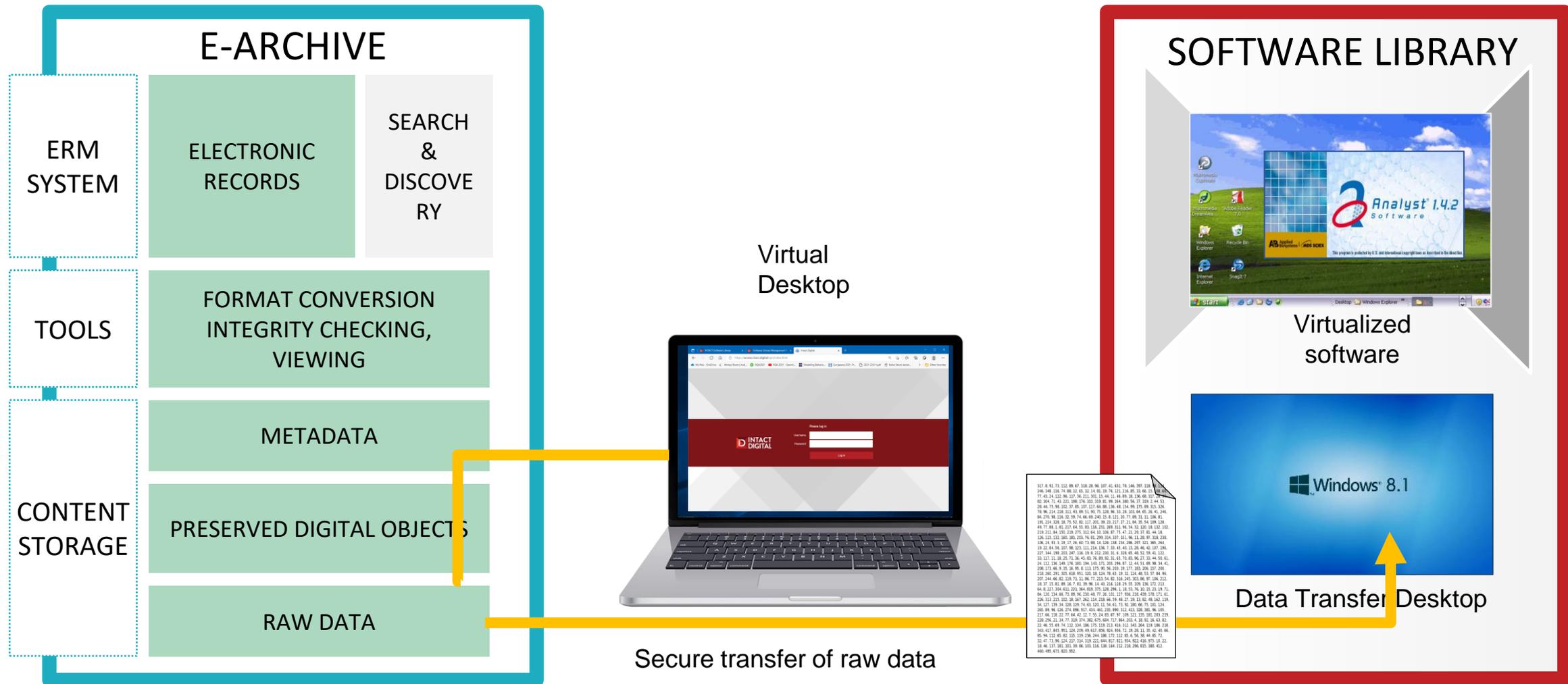
# Approach: Re-install software in virtual machines



- Data is transferred secure for study reconstruction

- Validated software installations
- Remote access through Virtual Desktops (in standard browsers).

# Safe data transfer for use with validated software



# Software Library

---

Long-term maintenance of computing environments



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Intact Digital

https://access.intact.digital/vpn/index.html



Please log in

Username

Password



Type here to search



7°C Mostly cloudy

11:02 04/11/2021

- Recycle Bin
- PDFsam Basic
- Adobe Acrobat DC
- Droplet
- 3D Objects - Shortcut
- Zoom
- Slack

Citrix Receiver

https://access.intact.digital/Citrix/IntactWeb/

**INTACT DIGITAL** DESKTOPS IDUser-007

Search Desktops

 Analyst 1-4-2 <a href="#">Details</a>	 Cell-IQ Analyser <a href="#">Details</a>	 CellActivision VM <a href="#">Details</a>	 DC Analyst <a href="#">Details</a>	 Digital Vault <a href="#">Details</a>
 Transfer Desktop <a href="#">Details</a>				



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

Please wait...

Microsoft Windows xp Professional

Copyright © 1985-2001 Microsoft Corporation

Applying your personal settings...



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

My Computer

Analyst Software

**IDLuser-007**

- Internet Explorer
- Outlook Express
- Windows Media Player
- Windows Messenger
- Tour Windows XP
- Windows Movie Maker
- Files and Settings Transfer Wizard
- My Documents
- My Recent Documents
- My Pictures
- My Music
- My Computer
- Control Panel
- Set Program Access and Defaults
- Printers and Faxes
- Help and Support
- Search
- Run...
- Windows Security

All Programs

Log Off Shut Down Disconnect

start

EN 11:02



Type here to search



7°C Mostly cloudy

11:05 04/11/2021



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184



My Computer  
Analyst Software

start

EN 11:04

Recycle Bin

Type here to search



7°C Mostly cloudy ENG

11:06 04/11/2021



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

My Computer Analyst Software

**Analyst**

File Edit View Tools Explore Window Script Help

Explore Mode Default

- Configure
  - Security Configuration
  - Hardware Configuration
  - Report Template Editor
- Tune
  - Resolution Optimization
  - Quantitative Optimization
  - Manual Tuning
- Acquire
  - IDA Method Wizard
  - Build Acquisition Method
  - Build Acquisition Batch
  - Express View
- Explore
  - Open Data File
  - Open Compound Database
- Quantitate
  - Build Quantitation Method
  - Quantitation Wizard
  - Review Results Table

For Help, press F1 User Name: iduser-007@intact.internal C:\Analyst Data

start Analyst EN 11:05



Type here to search



7°C Mostly cloudy

11:07 04/11/2021



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

Analyst

File Edit View Tools Explore Window Script Help

Default

U:\Data\_LIT&TOF\LIT

File and Folder Tasks

- Make a new folder
- Publish this folder to the Web

Other Places

- Data\_LIT&TOF
- My Documents
- My Computer
- My Network Places

Details

Bromocriptine Analyst Document 1,837 KB

Reserpine Analyst Document 158 KB

IDA BSA Digest Analyst Document 4,077 KB

For Help, press F1

User Name: iduser-007@intact.internal C:\Analyst Data

start Analyst LIT EN 11:05



Type here to search



7°C Mostly cloudy

11:08 04/11/2021



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver window with tabs: Analyst 1-4-2, Transfer Desktop

Browser address bar: <https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184>

Analyst - [ +EPI (609.20): 50 MCA scans from Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray)]

File Edit View Tools Explore Window Script Help

Default

EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray) Max. 3.0e7 cps.

U:\Data\_LIT&TOF\LIT

File and Folder Tasks

- Rename this file
- Move this file
- Copy this file
- Publish this file to the Web
- E-mail this file
- Delete this file

Other Places

- Data\_LIT&TOF
- My Documents
- My Computer
- My Network Places

Details

Bromocriptine Analyst Document 1,837 KB

IDA BSA Digest Analyst Document 4,077 KB

Reserpine Analyst Document 158 KB

365.2 397.3 448.3 609.4

m/z, amu

For Help, press F1

User Name: iduser-007@intact.internal C:\Analyst Data

start Analyst - [ +EPI (609... ) LIT EN 11:06



Type here to search



7°C Mostly cloudy

11:08 04/11/2021



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver window with tabs: Analyst 1-4-2, Transfer Desktop

Browser address bar: <https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184>

Analyst - [+EPI (609.20): 50 MCA scans from Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray)]

File Edit View Tools Explore Window Script Help

Explore Mode Default

Configure

- Security Configuration
- Hardware Configuration
- Report Template Editor
- Tune
  - Resolution Optimization
  - Quantitative Optimization
  - Manual Tuning
- Acquire
  - IDA Method Wizard
  - Build Acquisition Method
  - Build Acquisition Batch
  - Express View
- Explore (1)
  - Open Data File
  - Open Compound Database
- Quantitate
  - Build Quantitation Method
  - Quantitation Wizard
  - Review Results Table

Intensity, cps

m/z, amu

Max. 3.0e7 cps

159.2, 174.2, 182.2, 192.2, 195.2, 214.3, 224.3, 236.2, 238.2, 276.2, 288.3, 304.3, 322.1, 336.2, 340.3, 365.2, 368.2, 395.3, 397.3, 404.3, 413.3, 436.3, 448.3, 460.2, 489.2, 493.3, 533.1, 543.3, 547.3, 577.3, 593.2, 608.4, 609.4

For Help, press F1

User Name: iduser-007@intact.internal C:\Analyst Data

start Analyst - [+EPI (609... LIT EN 11:06

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

### Analyst - [Data List for "+EPI (609.20): 50 MCA scans from Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray)"]

File Edit View Tools Explore Window Script Help

Explore Mode Default

Configure

- Security Configuration
- Hardware Configuration
- Report Template Editor

Tune

- Resolution Optimization
- Quantitative Optimization
- Manual Tuning

Acquire

- IDA Method Wizard
- Build Acquisition Method
- Build Acquisition Batch
- Express View

Explore (1)

- Open Data File
- Open Compound Database

Quantitate

- Build Quantitation Method
- Quantitation Wizard
- Review Results Table

Intensity, cps

m/z, amu

Max. 3.0e7 cps

+EPI (609.20): 50 MCA scans from Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray)

	m/z (amu)	Intensity (cps)	Centroid mass (amu)	Peak area (counts)	Charges	Peak start (amu)	Peak end (amu)	Width (amu)	Resolution
1	151.2000	1.8333e5	151.1336	5.6667e5	1	150.9600	151.3200	0.2310	654.2847
2	152.1000	2.5000e5	152.1267	9.3333e5	Undefined	151.8600	152.3400	0.2400	633.8466
3	153.3600	1.6667e4	153.3360	3.3333e4	1	153.1800	153.4200	Undefined	Undefined
4	154.1400	1.0000e5	154.1660	2.0000e5	Undefined	153.9600	154.2600	0.1160	1329.0713
5	154.9200	3.3333e4	154.9200	3.3333e4	Undefined	154.8600	154.9800	Undefined	Undefined
6	155.2200	5.0000e4	155.2320	1.5000e5	1	154.9800	155.4000	0.3600	Undefined
7	156.1800	1.0000e5	156.1978	3.6667e5	Undefined	156.0000	156.4200	0.2300	679.1033
8	157.0800	5.0000e4	157.1109	2.1667e5	2	156.8400	157.3200	0.1160	805.7298
9	157.6800	3.3333e4	157.6900	5.0000e4	2	157.6200	157.8000	0.2400	Undefined
10	158.1000	2.6667e5	158.1274	7.1667e5	Undefined	157.9200	158.2800	0.1440	1098.0583
11	158.5200	6.6667e4	158.5255	6.6667e4	Undefined	158.4600	158.7000	Undefined	Undefined

For Help, press F1

User Name: iduser-007@intact.internal C:\Analyst Data

- Recycle Bin
- PDFsam Basic
- Adobe Acrobat DC
- Droplet
- 3D Objects - Shortcut
- Zoom
- Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

### Analyst - [File Information for Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff]

File Edit View Tools Explore Window Script Help

Explore Mode Default

Configure

- Security Configuration
- Hardware Configuration
- Report Template Editor

Tune

- Resolution Optimization
- Quantitative Optimization
- Manual Tuning

Acquire

- IDA Method Wizard
- Build Acquisition Method
- Build Acquisition Batch
- Express View

Explore (1)

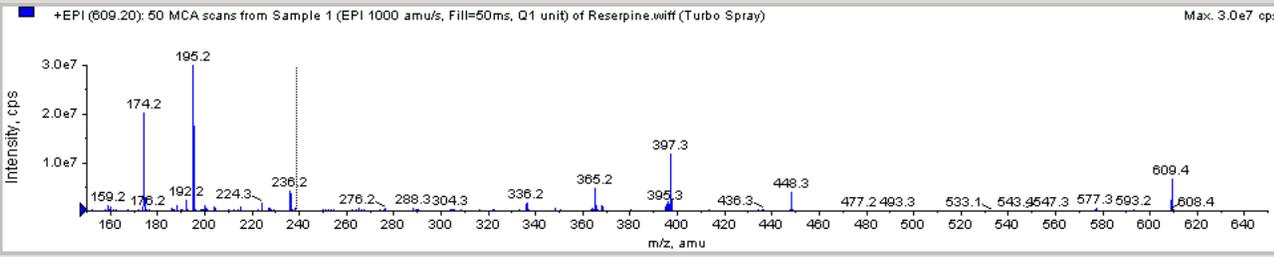
- Open Data File
- Open Compound Database

Quantitate

- Build Quantitation Method
- Quantitation Wizard
- Review Results Table

For Help, press F1

User Name: iduser-007@intact.internal C:\Analyst Data



+EPI (609.20): 50 MCA scans from Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray)

	m/z (amu)	Intensity (cps)	Centroid mass (amu)	Peak area (counts)	Charges	Peak start (amu)	Peak end (amu)	Width (amu)	Resolution
1	151.2000	1.8333e5	151.1336	5.6667e5	1	150.9600	151.3200	0.2310	654.2847
2	152.1000	2.5000e5	152.1267	9.3333e5	Undefined	151.8600	152.3400	0.2400	633.8466
3	153.3600	1.6667e4	153.3360	3.3333e4	1	153.1800	153.4200	Undefined	Undefined
4	154.1400	1.0000e5	154.1660	2.0000e5	Undefined	153.9600	154.2600	0.1160	1329.0713
5	154.9200	3.3333e4	154.9200	3.3333e4	Undefined	154.8600	154.9800	Undefined	Undefined

File Info

Log Info

Acquisition Info

Quant. Info

Period 1:

Resolution tables

Calibration tables

Instrument Parameters:

Keyed Text:

**Acquisition Info**

Acquisition Method: testTune.dam  
 Acquisition Path: D:\Analyst Data\Projects\Tune April 15\Example\Acquisition Methods\  
 First Sample Started: Not Available  
 Last Sample Finished: Not Available  
 Sample Acq Time: 26 April 2002 08:56:21  
 Sample Acq Duration: 28.410sec  
 Number of Scans: 51  
 Periods in File: 1

- Recycle Bin
- PDFsam Basic
- Adobe Acrobat DC
- Droplet
- 3D Objects - Shortcut
- Zoom
- Slack

Citrix Receiver | Analyst 1-4-2 | Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

Analyst - [File Information for Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff]

File Edit View Tools Explore Window Script Help

Default

EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray) Max. 3.0e7 cps.

Sample 1 (EPI 1000 amu/s, Fill=50ms, Q1 unit) of Reserpine.wiff (Turbo Spray)

Peak mass (amu)	Peak area (counts)	Charges	Peak start (amu)	Peak end (amu)	Width (amu)	Resolution
336	5.6667e5	1	150.9600	151.3200	0.2310	654.2847
367	9.3333e5	Undefined	151.8600	152.3400	0.2400	633.8466
360	3.3333e4	1	153.1800	153.4200	Undefined	Undefined
360	2.0000e5	Undefined	153.9600	154.2600	0.1160	1329.0713
300	3.3333e4	Undefined	154.8600	154.9800	Undefined	Undefined

Information Info

ion Method: testTune.dam  
 ion Path: D:\Analyst Data\Projects\Tune April 15\Example\Acquisition Methods\  
 Sample Started: Not Available  
 Sample Finished: Not Available  
 Acq Time: 26 April 2002 08:56:21  
 Acq Duration: 28.410sec  
 Number of Scans: 51  
 Periods in File: 1

User Name: iduser-007@intact.internal | C:\Analyst Data

LIT

File Edit View Favorites Tools Help

Address U:\Data\_LIT&TOF\LIT

File and Folder Tasks

- Rename this file
- Move this file
- Copy this file
- Publish this file to the Web
- E-mail this file
- Delete this file

Other Places

- Data\_LIT&TOF
- My Documents
- My Computer
- My Network Places

Details

Bromocriptine Analyst Document 1,837 KB

Reserpine Analyst Document 158 KB

IDA BSA Digest Analyst Document 4,077 KB

Type: Analyst Document Size: 3.98 MB

- Recycle Bin
- PDFsam Basic
- Adobe Acrobat DC
- Droplet
- 3D Objects - Shortcut
- Zoom
- Slack



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver | Analyst 1-4-2 | Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

### Analyst - [IDA Explorer - , Sample # 1 ()]

File Edit View Tools Explore Window Script Help

Explore Mode | Default

Mass Tolerance: 0.100 amu | Mass Window (amu): 5.000

m/z(amu)	Time(min)	Scan	CE	Z
<input checked="" type="checkbox"/> 385.370	0.017	ER		
<input type="checkbox"/> 367.995	0.083	ER		
<input type="checkbox"/> 391	0.083	ER		
<input type="checkbox"/> 394.4	0.328	ER		
<input type="checkbox"/> 367.534	0.369	ER		
<input type="checkbox"/> 395.468	0.410	ER		
<input type="checkbox"/> 445.152	0.410	ER		
<input type="checkbox"/> 416.952	1.140	ER		
<input type="checkbox"/> 429.050	1.687	ER		
<input type="checkbox"/> 373.268	2.167	ER		
<input type="checkbox"/> 475.368	2.562	EPI	16	U...
<input type="checkbox"/> 357.308	2.736	EPI	11	U...
<input type="checkbox"/> 475.380	2.537	ER		
<input type="checkbox"/> 412.409	2.628	ER		
<input type="checkbox"/> 440.340	2.628	ER		
<input type="checkbox"/> 445.043	2.669	ER		
<input type="checkbox"/> 357.397	2.710	ER		
<input type="checkbox"/> 566.086	2.710	ER		
<input type="checkbox"/> 359.236	2.777	ER		
<input type="checkbox"/> 566.210	2.777	ER		
<input type="checkbox"/> 566.213	2.802	EPI	20	U...
<input type="checkbox"/> 367.269	2.843	ER		
<input type="checkbox"/> 500.274	2.843	ER		
<input type="checkbox"/> 500.277	2.868	EPI	38	1
<input type="checkbox"/> 351.439	2.909	ER		
<input type="checkbox"/> 363.180	2.950	ER		
<input type="checkbox"/> 383.307	3.017	EPI	12	U...
<input type="checkbox"/> 369.267	3.149	EPI	30	1

Configure

- Security Configuration
- Hardware Configuration
- Report Template Editor
- Tune
  - Resolution Optimization
  - Quantitative Optimization
  - Manual Tuning
- Acquire
  - IDA Method Wizard
  - Build Acquisition Method
  - Build Acquisition Batch
  - Express View
- Explore (2)
  - Open Data File
  - Open Compound Database
- Quantitate
  - Build Quantitation Method
  - Quantitation Wizard
  - Review Results Table

Mass List - List View

TIC of +EMS: Exp 1, from Sampl... Max. 2.2e8 cps

+ER: Exp 2, 0.013 min from Sampl... Max. 0.0 cps

XIC of +EMS: Exp 1, 385.1 to 385... Max. 3.3e5 cps

Intensity, cps vs Time, min

Intensity, cps vs m/z, amu

Intensity, cps vs Time, min

Intensity, cps vs Mass, amu

For Help, press F1

User Name: iduser-007@intact.internal | C:\Analyst Data



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

### Analyst - [IDA Explorer - U:\Data\_LIT&TOF\LIT\IDA BSA Digest.wiff, Sample # 1 (4pmol all charges)]

File Edit View Tools Explore Window Script Help

Explore Mode Default

Configure

- Security Configuration
- Hardware Configuration
- Report Template Editor
- Tune
  - Resolution Optimization
  - Quantitative Optimization
  - Manual Tuning
- Acquire
  - IDA Method Wizard
  - Build Acquisition Method
  - Build Acquisition Batch
  - Express View
- Explore (2)
  - Open Data File
  - Open Compound Database
- Quantitate
  - Build Quantitation Method
  - Quantitation Wizard
  - Review Results Table

Mass Tolerance: 0.100 amu Mass Window (amu): 5.000

m/z(amu)	Time(min)	Scan	CE	Z
<input type="checkbox"/> 385.370	0.017	ER		
<input type="checkbox"/> 367.395	0.083	ER		
<input type="checkbox"/> 391.519	0.083	ER		
<input type="checkbox"/> 394.380	0.328	ER		
<input type="checkbox"/> 367.534	0.369	ER		
<input type="checkbox"/> 395.468	0.410	ER		
<input checked="" type="checkbox"/> 445.152	0.410	ER		
<input type="checkbox"/> 416.952	1.140	ER		
<input type="checkbox"/> 429.050	1.687	ER		
<input type="checkbox"/> 373.268	2.167	ER		
<input type="checkbox"/> 475.368	2.562	EPI	16	U...
<input type="checkbox"/> 357.308	2.736	EPI	11	U...
<input type="checkbox"/> 475.380	2.537	ER		
<input type="checkbox"/> 412.409	2.628	ER		
<input type="checkbox"/> 440.340	2.628	ER		
<input type="checkbox"/> 445.043	2.669	ER		
<input type="checkbox"/> 357.397	2.710	ER		
<input type="checkbox"/> 566.086	2.710	ER		
<input type="checkbox"/> 359.236	2.777	ER		
<input type="checkbox"/> 566.210	2.777	ER		
<input type="checkbox"/> 566.213	2.802	EPI	20	U...
<input type="checkbox"/> 367.269	2.843	ER		
<input type="checkbox"/> 500.274	2.843	ER		
<input type="checkbox"/> 500.277	2.868	EPI	38	1
<input type="checkbox"/> 351.439	2.909	ER		
<input type="checkbox"/> 363.180	2.950	ER		
<input type="checkbox"/> 383.307	3.017	EPI	12	U...
<input type="checkbox"/> 369.267	3.149	EPI	30	1

TIC of +EMS: Exp 1, from Sampl... Max. 2.2e8 cps

+ER (367.53): Exp 2, 0.406 min fr... Max. 5.8e4 cps

XIC of +EMS: Exp 1, 444.9 to 445... Max. 8.0e5 cps

Intensity, cps vs Time, min

Intensity, cps vs m/z, amu

Intensity, cps vs Time, min

Intensity, cps vs Mass, amu

For Help, press F1

User Name: iduser-007@intact.internal C:\Analyst Data



Recycle Bin



PDFsam Basic



Adobe Acrobat DC



Droplet



3D Objects - Shortcut



Zoom



Slack

Citrix Receiver Analyst 1-4-2 Transfer Desktop

https://access.intact.digital/Citrix/IntactWeb/clients/HTML5Client/src/SessionWindow.html?launchid=1636023883184

My Computer Analyst Software

- Preferences
- Logging
- About
- Display Resolution
- Ctrl+Alt+Del
- Disconnect
- Log Off

start LIT EN 11:09



Type here to search



9°C Mostly cloudy

11:12 04/11/2021

# Reproducibility of AI

---

Performance optimization and benchmarking

# AI system audits, validation and reproducibility

What would it take to ensure long-term use and impact assessment of an AI system?

- Code and code documentation
- Software components (building blocks)
- Tools (e.g., for writing and executing code)
- Configuration setting
- Training and test data, or relevant representations of data
- Documentation

# Framework for technology life-cycle management

## Technology

- Infrastructure for storing software assets, back-up and disaster recovery
- Infrastructure and services for hosting validated software installations
- Services for remote access control and security.

## Legal

- Licenses for operating systems, hypervisors and remote access
- Provisions for managing DRM protection of software to enable re-installations.

## Quality assurance

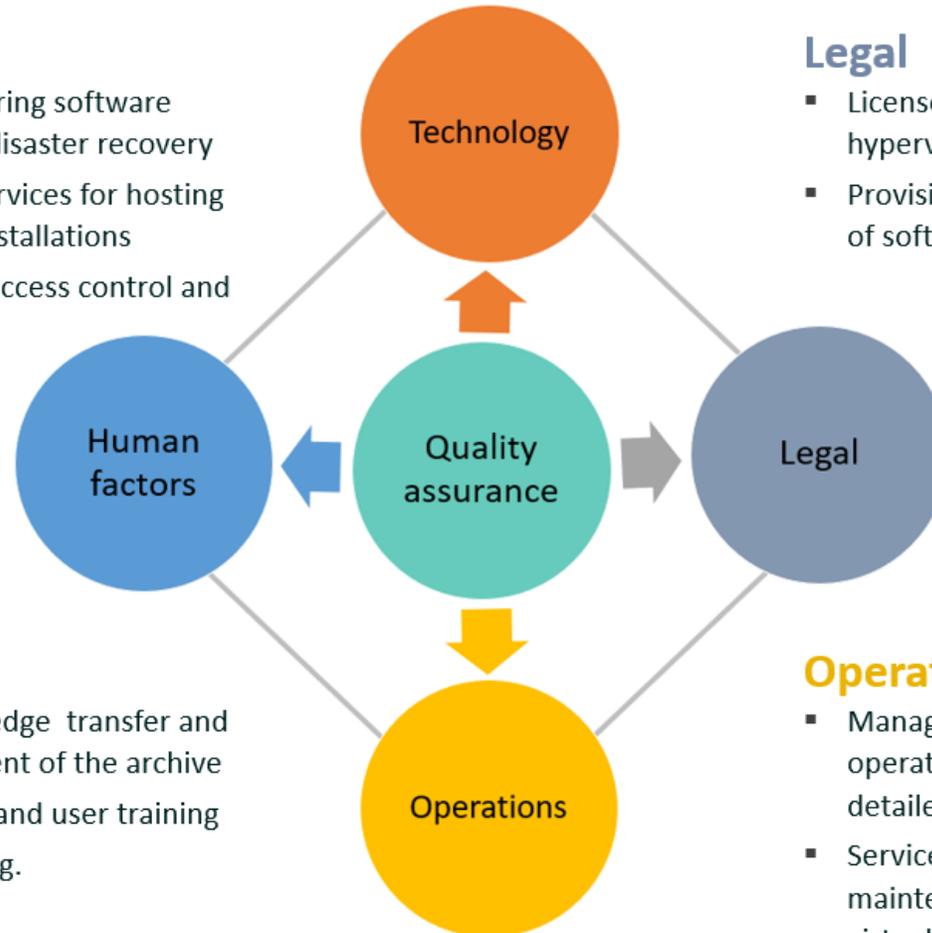
- Monitoring and periodic testing of platform components
- Review and update of the platform services and procedures in response to ecosystem changes.

## Human factors

- Provisions for knowledge transfer and long term management of the archive
- Community building and user training
- Technical staff training.

## Operations

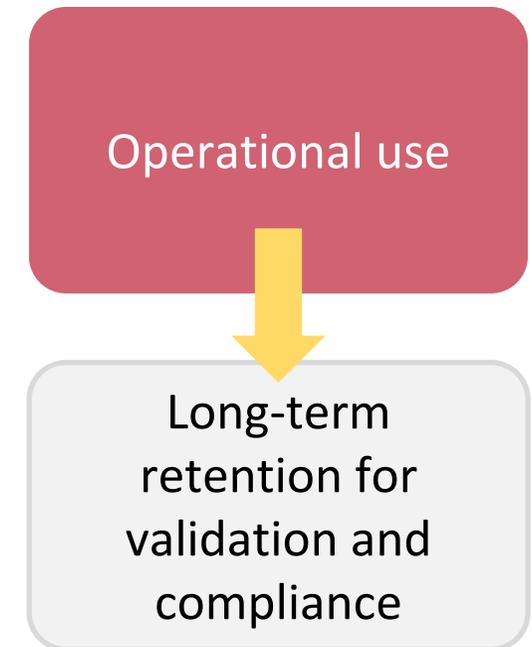
- Management of software transition from operational to archival use, including detailed documentation
- Services for installation, validation, delivery, maintenance and re-installation of virtualized software.



## Complexities due to use of AI cloud platforms

Use of AI environments with **pre-installed machine learning modules and tools**:

- Components constantly change
- There is complex relationship between versions of various software and tools
  - For example, some versions of Python are not compatible with computational packages (TensorFlow) and tools.
- Cloud platforms cannot promise **long-term maintenance of specific environments**.



# Complete and self-contained installations

---

## Replicating configurations of AI cloud installations



DESKTOPS

IDLuser-007

Search Desktops



Analyst 1-4-2

Details



Cell-IQ Analyser

Details



CellActivision VM

Details



DC Analyst

Details



Digital Vault

Details



Scholarly Desktop

Details



Scholarly GPU

Details



Transfer Desktop

Details



Mnist\_Digits\_Classification.ipynb

colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=CH1t\_kSCSfP0

Mnist\_Digits\_Classification.ipynb

File Edit View Insert Runtime Tools Help All cha...

+ Code + Text

RAM Disk Editing

imports

```
[1] import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import keras
from keras.models import Sequential
from keras.layers import Dense, Dropout
from sklearn.metrics import confusion_matrix
import seaborn as sns
import tensorflow as tf

np.random.seed(0)
```

Data

0s completed at 6:46 AM





Mnist\_Digits\_Classification.ipynb

colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=ewD7\_R6WSQOh

Mnist\_Digits\_Classification.ipynb

File Edit View Insert Runtime Tools Help All cha...

+ Code + Text

RAM  Disk  Editing

### Visualize Examples

```
[4] num_classes = 10
f, ax = plt.subplots(1, num_classes, figsize=(20,20))

for i in range(0, num_classes):
    sample = x_train[y_train == i][0]
    ax[i].imshow(sample, cmap='gray')
    ax[i].set_title("Label: {}".format(i), fontsize=16)
```

```
[5] for i in range(10):
    print(y_train[i])
```

0s completed at 6:46 AM





### Mnist\_Digits\_Classification.ipynb

colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=ewD7\_R6WSQOh

File Edit View Insert Runtime Tools Help All cha...

+ Code + Text RAM Disk Editing

## ▼ Create Model - fully Connected Neural Network

```
[10] model = Sequential()

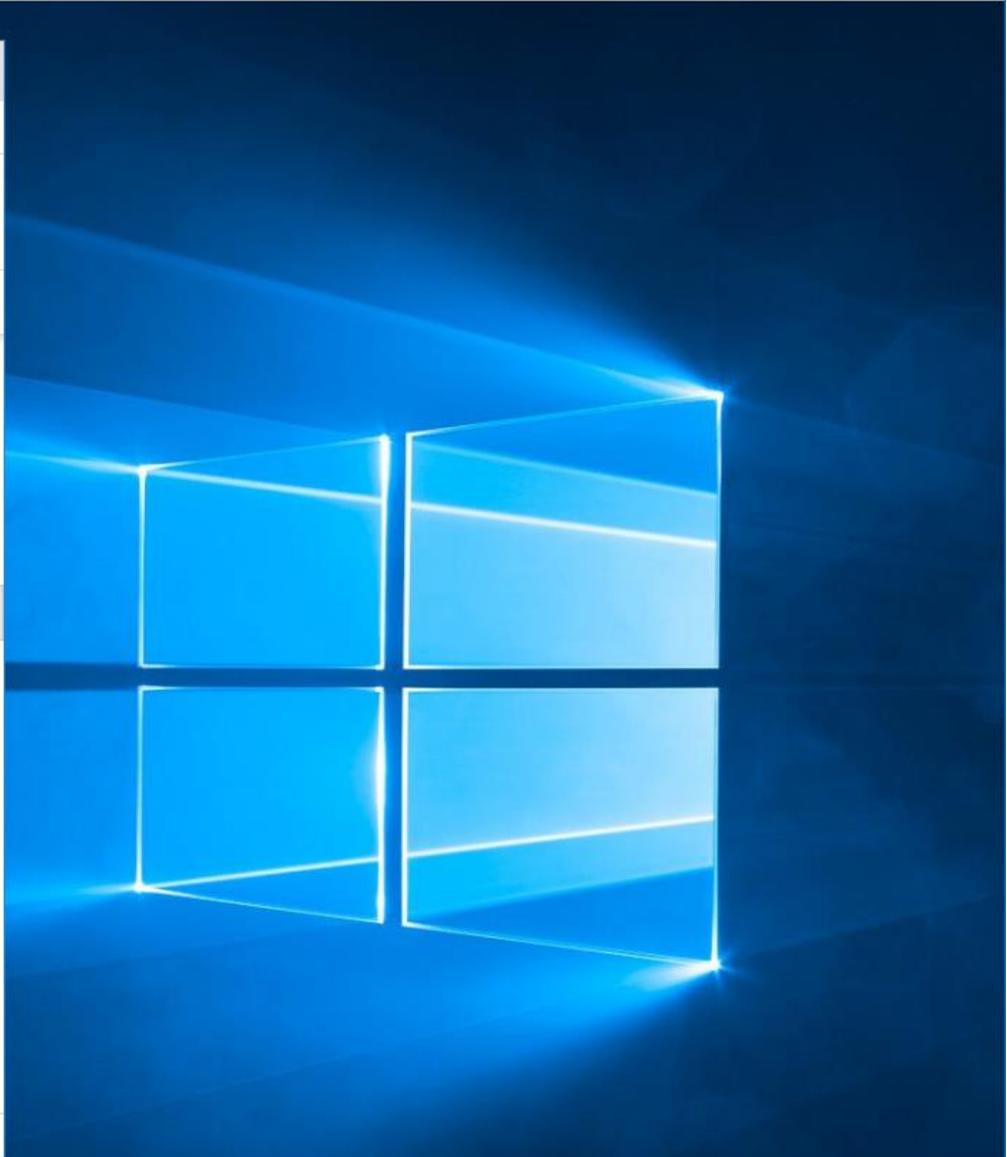
model.add(Dense(units=128, input_shape=(784,), activation='relu'))
model.add(Dense(units=128, activation='relu'))
model.add(Dropout(0.25))
model.add(Dense(units=10, activation='softmax'))

model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 128)	100480
dense_1 (Dense)	(None, 128)	16512

0s completed at 6:46 AM





```
Mnist_Digits_Classification.ipynb · x +  
colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=2hLSCxKRkwBJ  
Mnist_Digits_Classification.ipynb ☆  
File Edit View Insert Runtime Tools Help All cha...  
+ Code + Text RAM Disk Editing  
[11] batch_size = 512  
epochs=10  
model.fit(x=x_train, y=y_train, batch_size=batch_size, epochs=epochs)  
  
Epoch 1/10  
118/118 [=====] - 3s 17ms/step - loss: 0.5988 - accuracy: 0.82  
Epoch 2/10  
118/118 [=====] - 2s 20ms/step - loss: 0.2255 - accuracy: 0.93  
Epoch 3/10  
118/118 [=====] - 2s 17ms/step - loss: 0.1656 - accuracy: 0.95  
Epoch 4/10  
118/118 [=====] - 2s 19ms/step - loss: 0.1320 - accuracy: 0.96  
Epoch 5/10  
118/118 [=====] - 2s 16ms/step - loss: 0.1072 - accuracy: 0.96  
Epoch 6/10  
118/118 [=====] - 1s 11ms/step - loss: 0.0916 - accuracy: 0.97  
Epoch 7/10  
118/118 [=====] - 1s 10ms/step - loss: 0.0788 - accuracy: 0.97  
Epoch 8/10  
118/118 [=====] - 1s 11ms/step - loss: 0.0693 - accuracy: 0.97  
Epoch 9/10  
118/118 [=====] - 1s 11ms/step - loss: 0.0605 - accuracy: 0.98  
0s completed at 6:46 AM
```





Mnist\_Digits\_Classification.ipynb x +  
colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=3WxgbXuyMpU0  
Mnist\_Digits\_Classification.ipynb ☆  
File Edit View Insert Runtime Tools Help All cha...  
+ Code + Text RAM Disk Editing ^  
Evaluate  
[12] test\_loss, test\_acc = model.evaluate(x\_test, y\_test)  
print("Test Loss: {}, Test Accuracy: {}".format(test\_loss, test\_acc))  
313/313 [=====] - 1s 2ms/step - loss: 0.0697 - accuracy: 0.977  
Test Loss: 0.06969588994979858, Test Accuracy: 0.9775000214576721  
y\_pred = model.predict(x\_test)  
y\_pred\_classes = np.argmax(y\_pred, axis=1)  
print(y\_pred)  
print(y\_pred\_classes)  
[[8.4542842e-07 9.4627342e-07 1.7323036e-05 ... 9.9983215e-01  
1.9796314e-06 8.2834624e-05]  
[6.6998376e-07 3.4376166e-03 9.9346453e-01 ... 1.3521279e-07  
2.1691810e-04 2.7049094e-07]  
[8.5400588e-06 9.9746203e-01 3.2820937e-04 ... 7.7688455e-04  
6.7511125e-04 1.9763247e-05]  
...  
0s completed at 6:46 AM



- Recycle Bin
- Acrobat Reader DC
- Colab
- Google Chrome
- Google Earth Pro
- Jupyter Notebook
- Map Warper

Mnist\_Digits\_Classification.ipynb

colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=namQmMhqmnadb

Mnist\_Digits\_Classification.ipynb

File Edit View Insert Runtime Tools Help All cha...

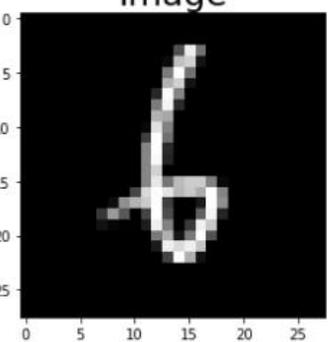
+ Code + Text

RAM  Disk  Editing

```
[14] ax[1].set_xticks([])
ax[1].set_yticks([])
ax[2].set_xticks([])
ax[2].set_yticks([])

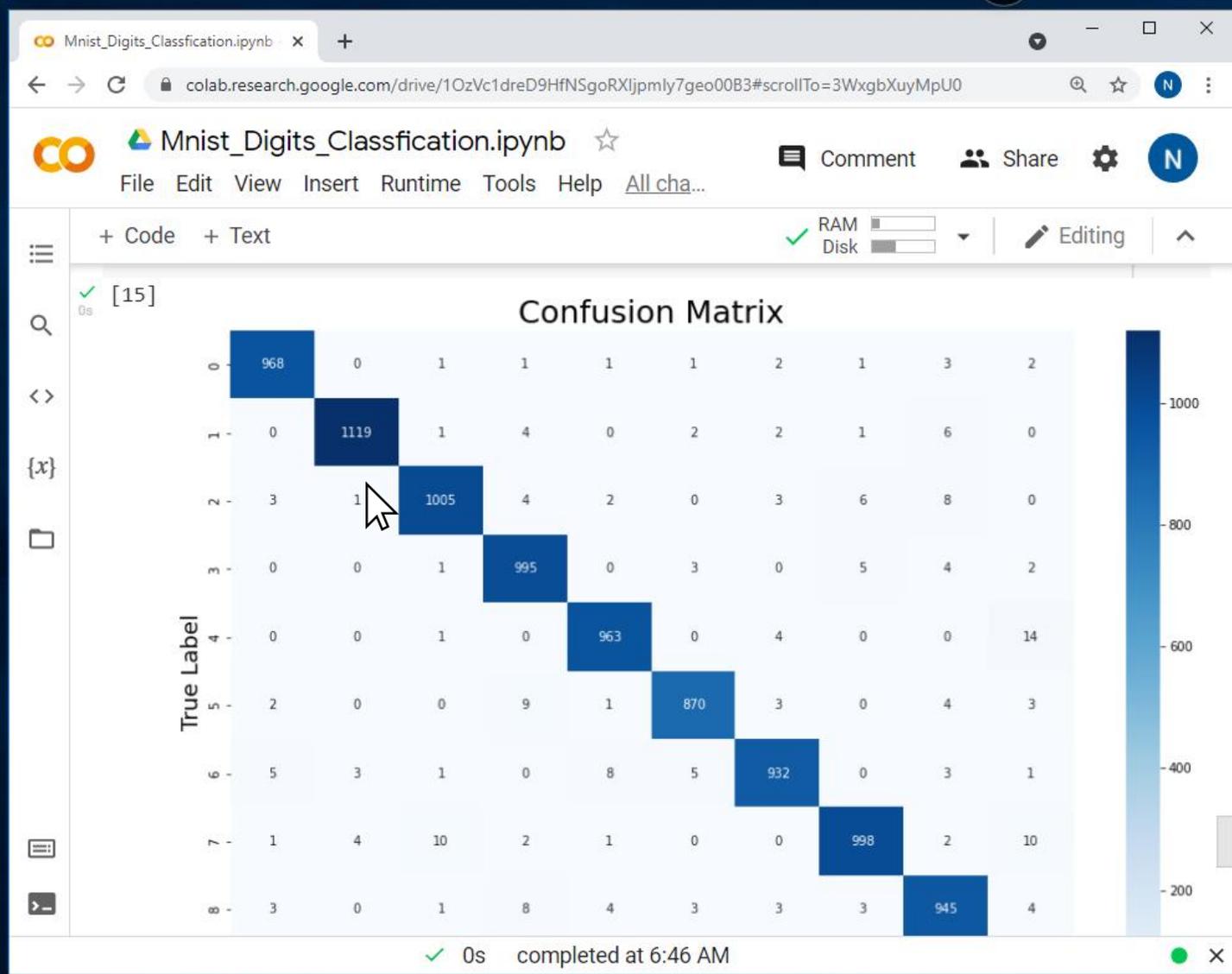
plt.title("Predicted: {}, True: {}".format(y_sample_pred_class,y_sample_true), fontsi
plt.imshow(x_sample.reshape(28,28), cmap='gray')

plt.show()
```

Image	Predicted	True
		

0s completed at 6:46 AM







Mnist\_Digits\_Classification.ipynb · x

colab.research.google.com/drive/1OzVc1dreD9HfNSgoRXIjpmly7geo00B3#scrollTo=CH1t\_kSCSfP0

Mnist\_Digits\_Classification.ipynb ☆

File Edit View Insert Runtime Tools Help All cha...

+ Code + Text

RAM  Disk  Editing

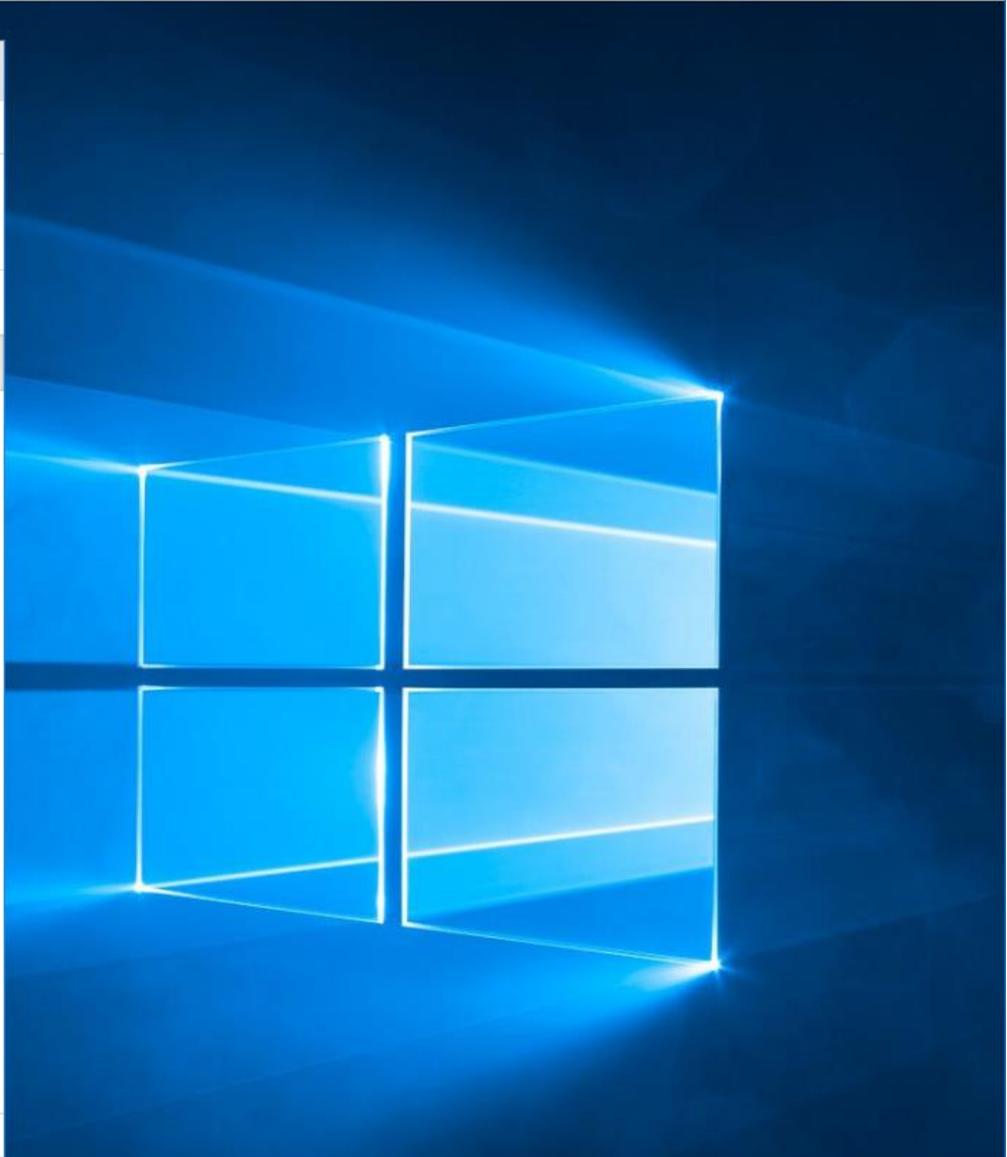
### imports

```
[1] import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import keras
from keras.models import Sequential
from keras.layers import Dense, Dropout
from sklearn.metrics import confusion_matrix
import seaborn as sns
import tensorflow as tf

np.random.seed(0)
```

Data

0s completed at 6:46 AM



The image shows a Jupyter Notebook interface with a code cell containing the following Python code:

```
[1] import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import keras
from keras.models import Sequential
from keras.layers import Dense, Dropout
from sklearn.metrics import confusion_matrix
import seaborn as sns
import tensorflow as tf

np.random.seed(0)
```

Below the code cell, there is a 'Data' section. A file browser overlay is visible on the right side of the notebook, showing the file structure:

- 0 /
- Data\_LIT&TOF (3 months ago)
- Projects (a year ago)
- Windows (a year ago)
- Mnist\_Digits\_Classfication.ipynb (13 hours ago, 10.2 KB)
- eventvwr.lnk (a year ago, 1.26 KB)
- Steps Recorder.lnk (a year ago, 1.22 KB)

At the bottom of the notebook, a status bar indicates: ✓ 0s completed at 6:46 AM

The image shows a Windows desktop environment with a Jupyter Notebook open in a browser. The notebook is titled "Mnist\_Digits\_Classification" and is running on a Python 3 (ipykernel) environment. The code in the notebook is as follows:

```
[1] import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import keras
from keras.models import Sequential
from keras.layers import Dense, Dropout
from sklearn.metrics import confusion_matrix
import seaborn as sns
import tensorflow as tf

np.random.seed(0)
```

The notebook interface includes a menu bar (File, Edit, View, Insert, Runtime, Tools, Help), a toolbar with icons for file operations and execution, and a status bar at the bottom indicating the execution time (0s) and completion time (6:46 AM).

# Reproducibility and quality of assurance

---

Looking ahead

# Key considerations for AI validation and assurance

## AI VALIDATION

- **Reproducibility of AI** processes is important for benchmarking and validation
- Systems are complex, with many **building blocks and resources** that are not fully accessible – black boxes
- System properties can be understood only through **repeated and systematic observations** of a stable system.

## AI ASSURANCE

- AI Assurance will involve **system audits** and evaluation based on **evidence of system performance**
- AI impact will lead to requirements for **long-term system retention** and use and require management of technology obsolescence
- Cloud based installations do not support long-term use of AI systems and separate **archiving installations** of the system are required.

# Thank you!

Natasa Milic-Frayling

Intact Digital Ltd  
[natasamf@intact.digital](mailto:natasamf@intact.digital)