

# Hydrating a Data Lake to Power Innovation

Enabling Pharmaceutical Companies to Automate  
Data De-Identification for Secondary Use

d-wise™

An Instem Company



Kristin McDougall, Director of Transparency

# The Cost of Transparency

## RESOURCES NEEDED



### Time

Reacting to new regulations and developing new processes takes time.



### Money

Both outsourced and internal methods require additional costs.



### Knowledge

Keeping up-to-date with current regulations can feel like a full-time job.

## ROI



### Public Education

Transparency enables the public to make informed decisions regarding their health care.



### Trust

Being transparent with trial data builds trust in the clinical trial process.



### Secondary Research

Empower medical breakthroughs by utilizing existing data to minimize duplicative efforts and reduce timelines.

# Thinking Beyond Regulations



**Completed Study Data**



**Compliance Standards  
Roadblocks**



**Exploratory analysis to  
support future studies**

**Minimize duplicate efforts**

**Medical Breakthroughs**

**Potential**

# Automating De-Identification at a Large Scale | Case Study

A global sponsor enlisted d-wise to provide a proof-of-concept capability to **accelerate the flow of data to and from an internal data lake** through d-wise's Blur anonymization solution, providing sufficiently anonymized content to achieve re-identification risk levels appropriate for internal secondary use.

## Challenges

What to Anonymize

Internal Process Alignment

Scalability

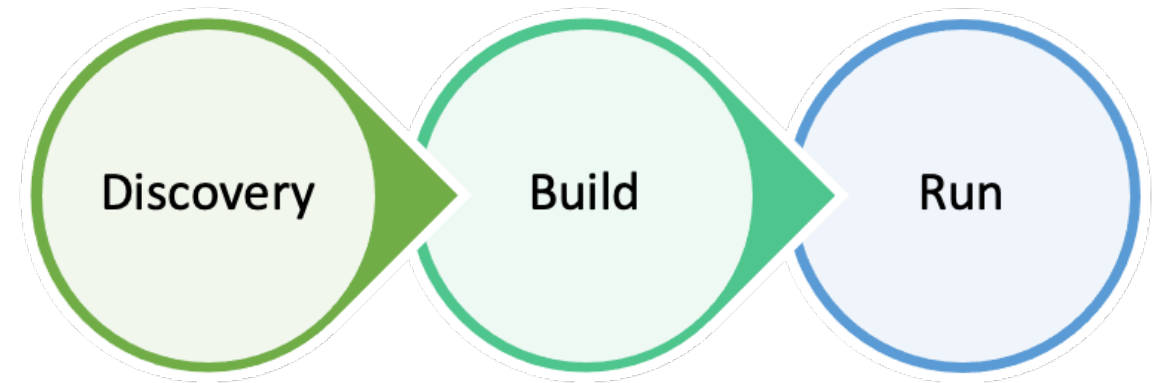
Risk Threshold Strategy

## Solution

Automated de-identification  
of datasets through Blur

# d-wise Approach

- **Discovery, Build, Run Process**
- **Master Template Format**
- **Audit Trail**
- **Variable Analysis**
- **Verbatim Text**
- **Inbound Folder Structure for System Input**
- **Outbound Folder to Receive Final, Resorted Data**



# Run Phase

## Manual

- Place study folders into Inbound folder

## Automated

- Template applied
- Calculating Match Score
- Anonymization Report
- Resorting datasets
- Exporting into study folders in Outbound folder

## Results

**84**

Studies  
Anonymized

**30**

Minutes

**1**

Manual  
Task

# Contact

**Kristin McDougall**

**Director of Transparency**

[kristin.mcdougall@d-wise.com](mailto:kristin.mcdougall@d-wise.com)

[www.d-wise.com](http://www.d-wise.com)