


Configure Intelligence



XURMO
Configure Intelligence



Do you deal with large quantities of data that are messy or difficult to read?

PDF, JPG, large file systems, etc.

Is there a need to bring different data sets together to get a business view?

Mails, data from RDBMS systems, Text files, etc.

Would you rather having an easy to use analytics tool that requires no coding

Replace cumbersome scripting with something easier

Are you extracting data manually or using multiple tools to read data sources?

Using scripts that need change as data changes

Do you use multiple tools to get insights from raw data?

Stitching a data lake, data integration & analytics tools

About Xurmo

Xurmo provides *SOLUTIONS* to data problems using its *SINGLE TECHNOLOGY STACK* that automates

- **Data Extraction**
- **Data Preparation**
- **Analysis**

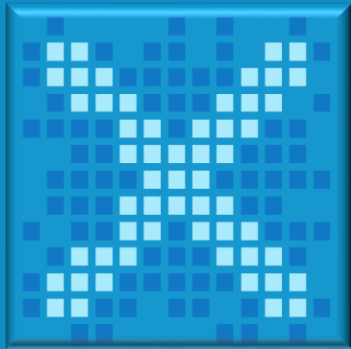
Proprietary Analytics Platform + Experienced Team

=

Accelerated Data Solutions

The Xurmo Proposition

To
Simplify &
Automate
Data Analysis

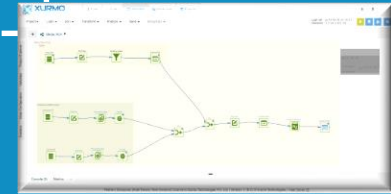


Xurmo Suite

Xurmo Bridge
For Admin of Xurmo



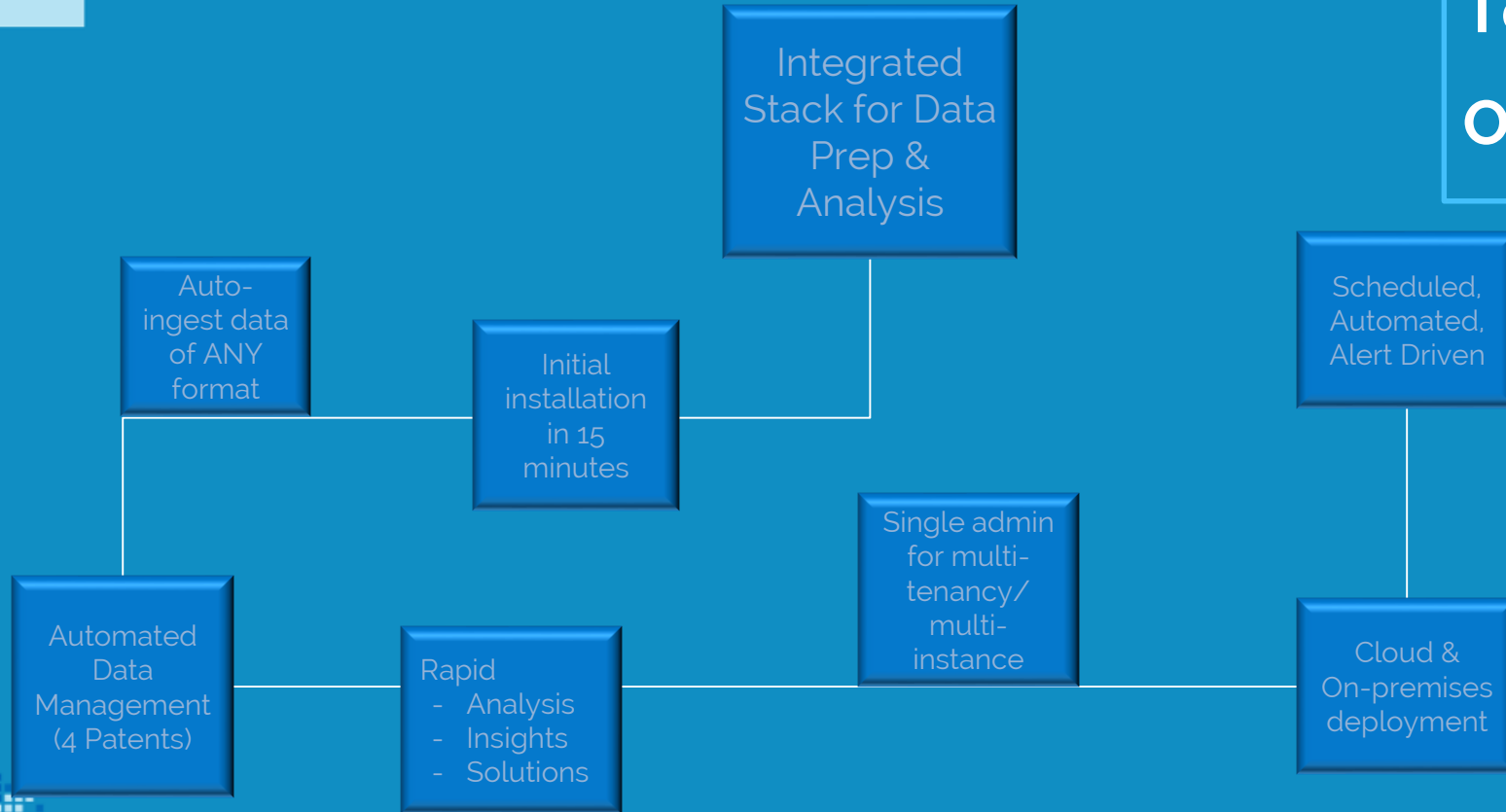
Xurmo Canvas
Data Prep & Analysis workbench



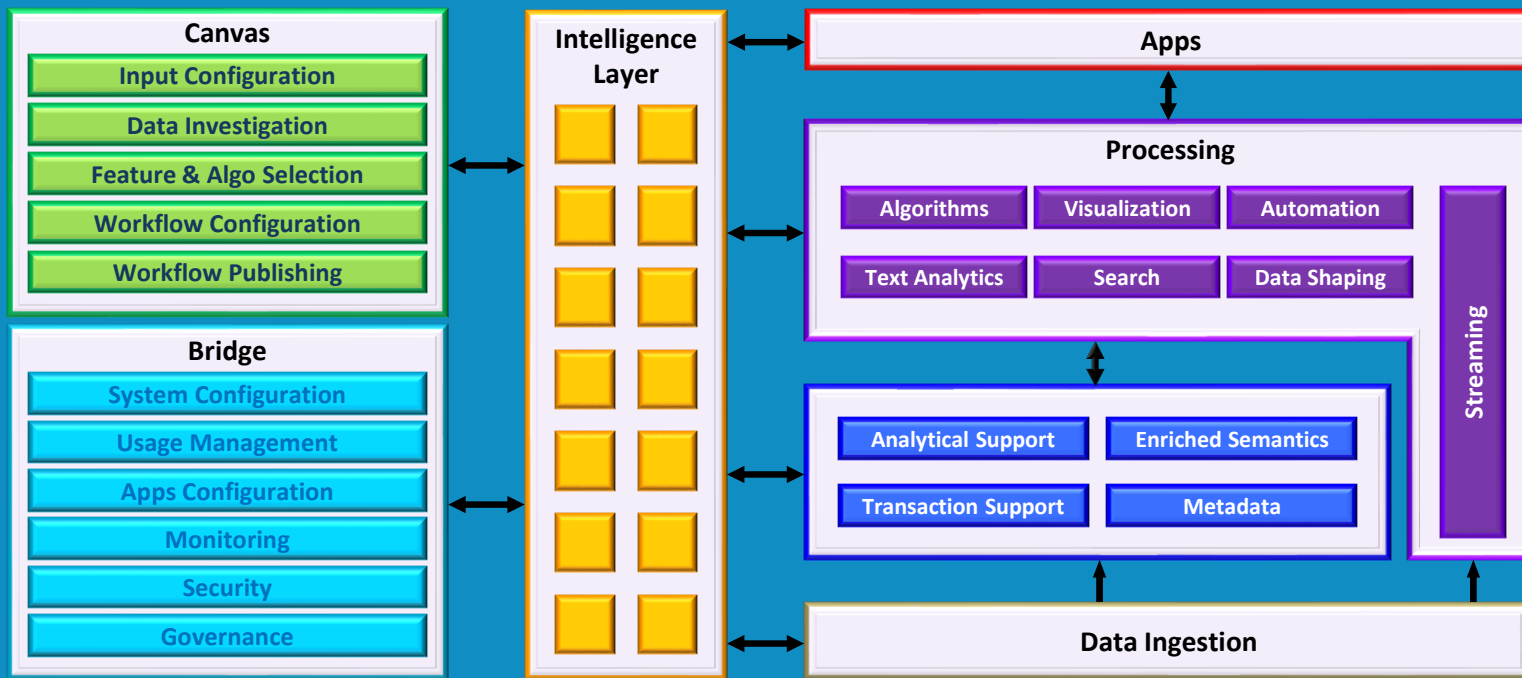
Xurmo Platform
The core for storage and processing



Tech Overview



Xurmo dramatically reduces the cost of Data Analytics





Case Studies



Increase ROI
from Media
Spend

Customer
Tier-1 Marketing
Analytics
Service Provider

Objective
To automate generation of
a single view of sales,
discounts and commissions
across regions and brands
every year



Challenges

Input files have different formats

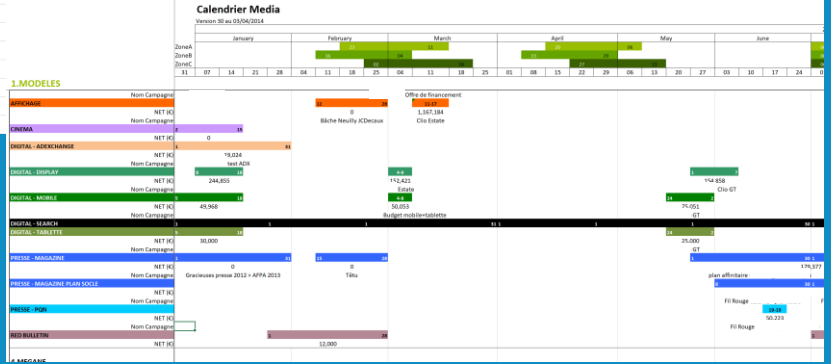
Like these

Model Market/Region Campaign Tier 1, 2, or 3 Format Any Comments?	APRIL			MAY			AUGUST		
	Traffic / Facings / Impressions	Actual Spend		Traffic / Facings / Impressions	Actual Spend		Traffic / Facings / Impressions	Actual Spend	
TV TAM AIRLINES (INDOOR)	2,852,000 impacts by month	RS 150,778		2,852,000 impacts by month	RS 150,778		26 screens / 200 insertions by day / 156,000 impressions by month	RS 33,564	
TV AIR - AIR_SP - SCREENS	26 screens / 200 insertions by day / 156,000 impressions by month	RS 33,564		26 screens / 200 insertions by day / 156,000 impressions by month	RS 33,564		26 screens / 200 insertions by day / 156,000 impressions by month	RS 33,564	
SCREENS - BARRISQU, CURITIBA	26 screens / 200 insertions by day / 156,000 impressions by month	RS 16,800		26 screens / 200 insertions by day / 156,000 impressions by month	RS 16,800		260 facings / 27,000,000 impacts by month		
STRECALIN (STREET CLOCKS)									

April 2013 - Sept 2015									
Week	1/4/2013	8/4/2013	15/4/2013	22/4/2013	29/4/2013	6/5/2013	13/5/2013	20/5/2013	27/5/2013
Brand									
B-1									
Budget	880916	458764	378068	253015	740687	868960	445988	741326	123626
GRP 25-49	45.6	23.7	19.6	13.1	38.3	45	23.1	38.4	6.4
B-Max									
Budget	979903	735822	711188	499166	962859	651015	883464	705272	706489
GRP 25-49	50.7	38.1	36.8	25.8	49.8	33.7	45.7	36.5	36.6
D-1									
Budget	290011	430303	862291	458625	588278	476354	949581	264864	395564
GRP 25-49	15	22.3	44.6						
FE-23									
Budget	861008	243779	584395						
GRP 25-49	44.6	12.6	30.2						
FF-18									
Budget	760378	987656	734252						
GRP 25-49	39.4	51.1	38						

Significant manual effort expended to re-create data flow to generate same reports

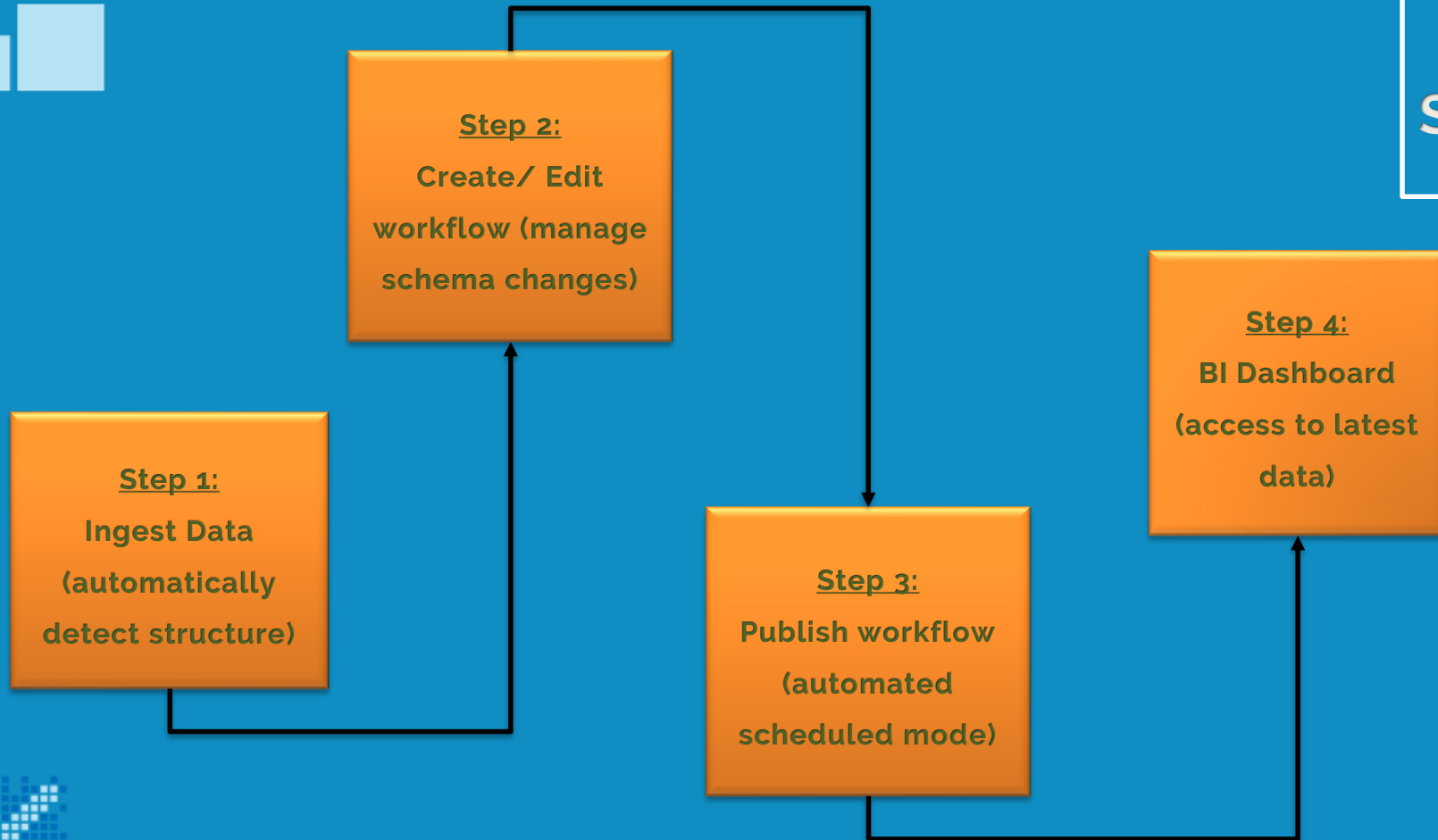
Time taken to prepare data and analysis = 3 months




Like this
In some cases color has meaning

Schema changes across time

The Solution



Time taken to develop the solution on Xurmo = 1 day
Time taken to subsequently get the output for new input = almost real-time



2-Wheeler Dealer Pricing Analysis

Customer
Tier-1
Motorcycle
Manufacturer

Objective
Identify Dealers who are
over charging customers,
thereby negating
Manufacturer's pricing
strategy

Challenges

Transactions across
18 States, by more
than 800 dealers on
more than 70
motorcycle variants
were to be analysed

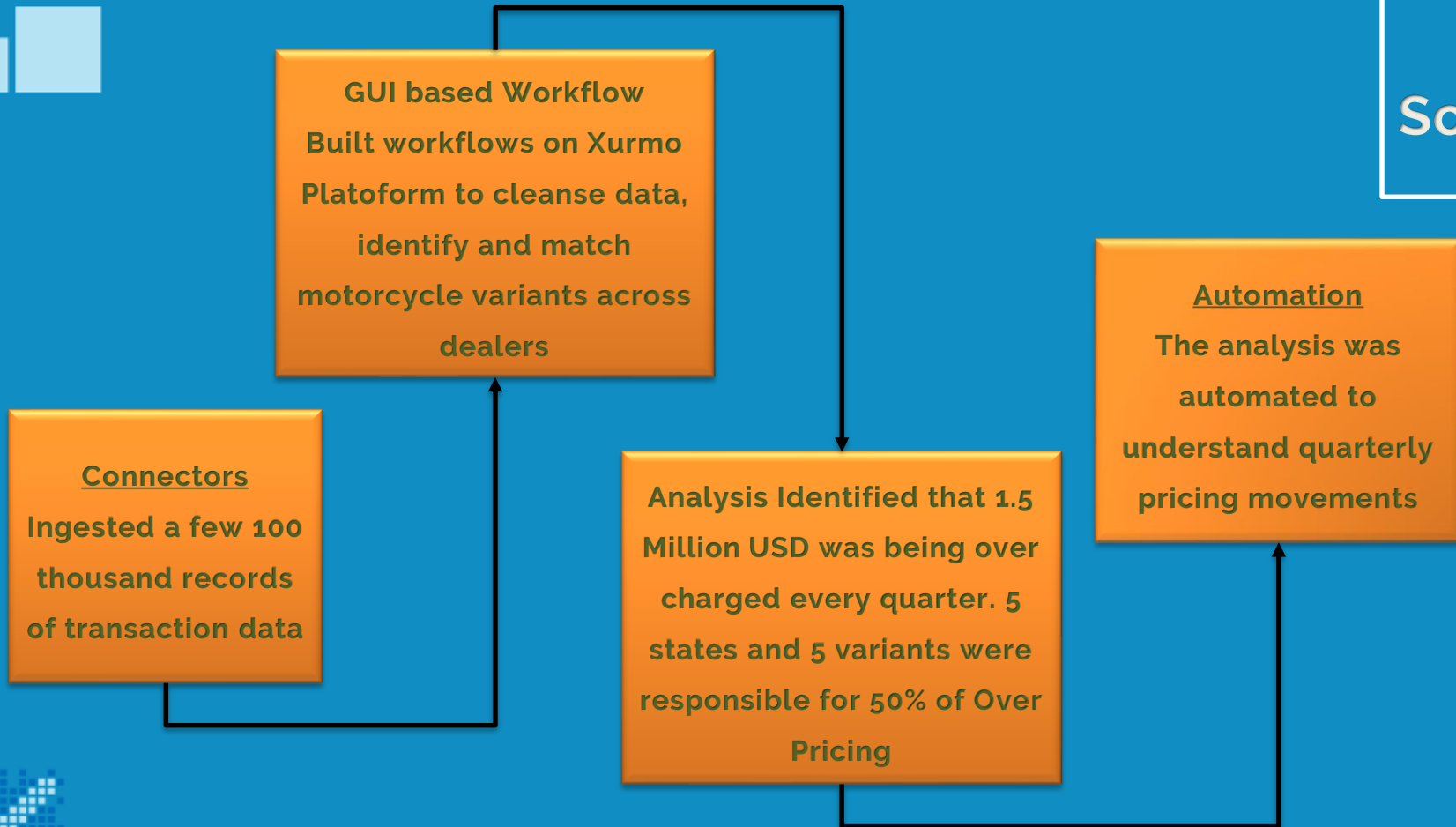
Nomenclatures
across Dealers
and States varied

Data was spread across
multiple systems
including Databases,
SAP systems, Emails
and Excel files

Vehicle On Road
Pricing formats
and pricing
breakups varied
across Dealers

Repeated analysis
on the latest
transactions were
time consuming
and never readily
available

The Solution



Dealer Pricing in-line with Pricing Strategy



Prospect Identification from Web Data

Customer
Tier-1 Bank

Objective
Identify potential
customers for forex
products from their
data on public web



Challenges

Information
about companies
was found across
numerous web
sources

The only
available
information was a
vast list of SME
Customer names

Manual data
extraction and
analysis of each
company takes an
inordinate amount
of time

Unstructured
web data
was noisy!



Step 1:
The web was crawled to search and extract text dump from numerous web sources

Step 2:
The drag and drop interface was used to build data transformation workflows

Step 3:
Analytics models were built and assembled on the same Platform

Step 4:
Workflows to ingest, and transform data and identify prospects was automated

The Solution

Quick access to immediate revenue opportunities





Extracting data from vast repository

Customer

Top 5 Global
Law Firm

Objective

Making the vast repository of Trademark and other legal documents more useful and accessible by extracting required attributes





Challenges

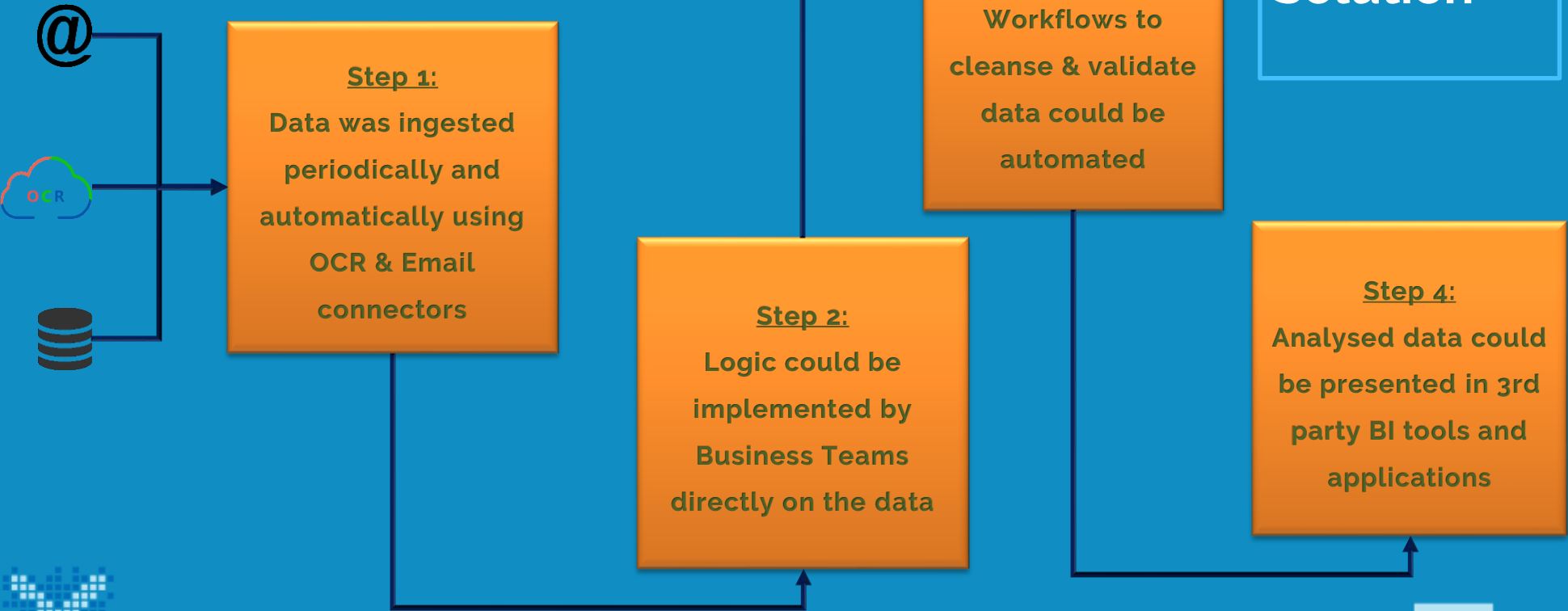
Documents are stored as scanned images and require OCR

Significant manual effort and domain expertise required

Many documents are present as email attachments

Effort to extract attributes linearly increases with number of publications

The Solution



Time taken to develop solution on Xurmo < 1 day



Document Tagging

Customer
Top 5 Global
Pharmaceutical
Company

Objective
Making the Publication
repository more useful
and accessible by
tagging documents with
medical terminology



Challenges

Required
going to
multiple web
sources to tag
documents

Significant
manual effort
and domain
expertise was
required

Manual
tagging was
inconsistent

Tagging effort
linearly
increased with
number of
publications

The Solution



XURMO
Configure Intelligence

Step 1:
Raw Data
including Clinical
Trials Web data
was ingested

Step 2:
Pattern matching,
dictionary based tagging
and Machine Learning
based classification
were implemented
through GUI based
workflows

Step 3:
Workflows to cleanse
and transform data
were automated to
provide tags on new
data

Step 4:
An Ensemble of
multiple ML
classification
approaches were used
for high accuracy
predictions

Step 5:
Analysed data was
accessible by a
custom application
built for the
scientists

Tagging time reduced from months to minutes



Thank you

