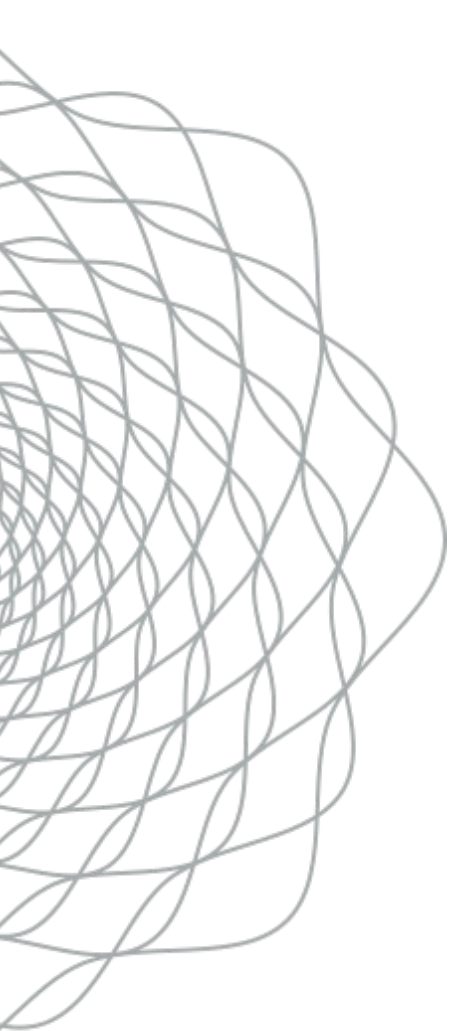




TARPON
HEALTH

a healthcare automation community

Getting Data Out Of Your EMR

Agenda

1. Why it's important
2. Process and common tools
3. Pros and cons of different approaches
4. Discussion on hierarchy

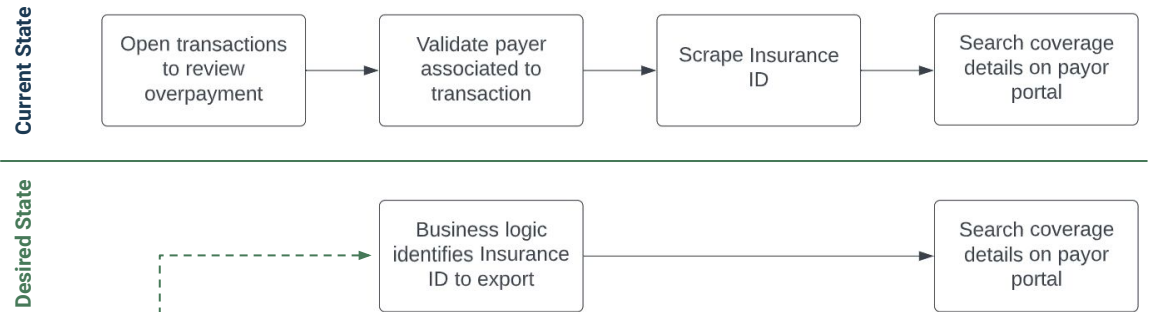
Why this is important?

Stability and speed

Example - Overpayment Identification

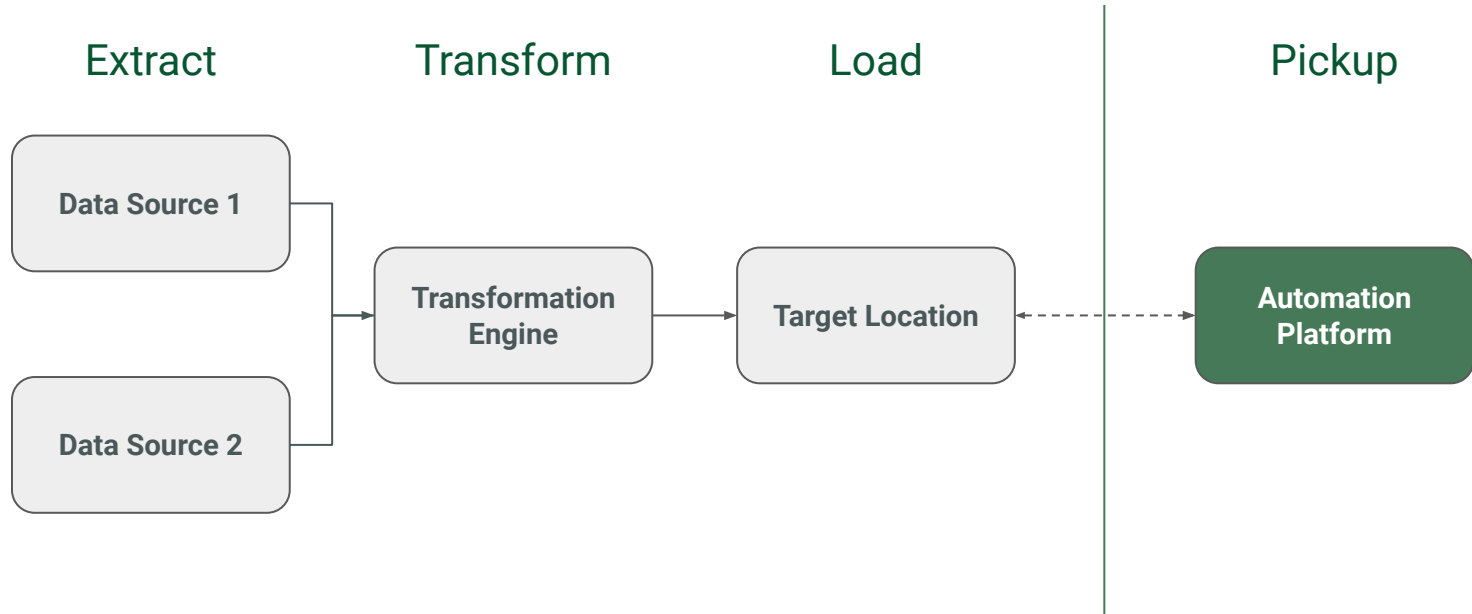
How do you simplify
the number of
actions?

What steps do you
need to take to get
the right data?



Concept review: Tarpon Health Solution
Architecture, Part 1

ETL - High Level Process



ETL and SSIS defined

Extract, Transform, Load (ETL)

*ETL is a **general term** that refers to the process of extracting data from various sources, transforming that data to meet specific requirements, and then loading the transformed data into a target destination, such as a database.*

*ETL is a concept in data processing and is **not specific** to any particular technology or platform.*

SQL Server Integration Services (SSIS)

SSIS is a **specific data integration tool** developed by Microsoft as part of the SQL Server suite of products. SSIS provides a graphical user interface and a comprehensive set of tools for building, deploying, **and managing ETL processes.**

It allows users to create complex data integration workflows, perform data transformations, and automate data loading tasks.

ETL and SSIS differences

Platform Specificity: SSIS is a Microsoft-specific tool, while ETL is a general concept that can be implemented using various tools and technologies, including SSIS, but also other platforms like IBM's Datastage, Airflow ETL from Apache, or Informatica.

Scope: ETL refers to the overall process of data integration, while SSIS is a specific tool that provides a comprehensive set of features and capabilities for implementing ETL workflows.

Flexibility: SSIS is tightly integrated with the Microsoft SQL Server ecosystem, making it well-suited for Microsoft-centric environments. ETL, on the other hand, is more platform-agnostic and can be used across various data sources and target systems.

Features: SSIS offers a wide range of features, including data transformation, data cleansing, data validation, error handling, and scheduling, among others. ETL, as a general concept, can incorporate these features, but the specific implementation and capabilities may vary depending on the tool or technology used.

sFTP defined

sFTP is a common way to “Load” (of Extract, Transfer, “Load”) data.

Purpose: sFTP is a protocol used to securely transfer files between systems over a network. This is the “Load” part of the ETL process.

Functionality:

- File Transfer: Enables the transfer of files between computers or servers.
- Security: Encrypts data during transfer to protect it from interception.
- Authentication: Uses secure methods like SSH (Secure Shell) keys or passwords for user authentication.

Data Integrity: Ensures files are not altered during transfer.

Scope: Focused **solely on the secure transfer of files**, not on data transformation or loading into databases.

Platform: Can be used across various systems and platforms, **not limited to any specific database or technology**.

Overview of approaches

Below is a brief description of common approaches taken to build a queue of accounts in your automation platform.

Approach	Description
RPA pulls data line by line	Have an automation work a list of accounts built into the EMR via login and search functionality.
RPA to generate an export	Have an automation login into a system get an export
ETL Exports from a database or business intelligence tool	Have a defined extract sent to the automation platform on a defined schedule directly from a database (Clarity) or business intelligence tool (Strata).
Direct query	Have an automation run a query directly to obtain a queue
APIs	APIs can bring in additional data to supplement extracts

How do I know what approach is best for me?



Tarpon suggests talking with your security and compliance team about your projects and the importance of reducing clicks and actions, stability, autonomy, and speed.

Focus on:

- Reducing clicks and actions
- Stability
- Autonomy
- Speed and iteration

AND

- ***Whatever Security and Compliance allows!***

RPA pulls data line by line

Using RPA to access an existing worklist can work, but it adds clicks and actions potentially resulting in more issues arising.

Pros

- Like RPA in general - it's extremely flexible
- You know the data is up-to-date
- Existing worklists can be tapped-into which means less report-building work
- Potentially faster timelines because you're not reliant on other team members

Cons

- Creates more steps and complexity in the process than is needed
- More prone to breaks / issues
- Much harder to iterate and isolate the right population; the exception handling will have to happen within the automation itself
- Longer run times = more licence costs

Tarpon recommends using this option when all other options are extinguished

RPA pulls an extract

Using RPA to access a predefined report is a better approach, but it still has unnecessary clicks and actions.

Pros

- Like RPA in general - it's extremely flexible
- Data is near up-to-date
- Less prone to breaks and issues than working directly out of a worklist
- Exception handling can happen within the export

Cons

- Creates more steps and complexity in the process than is needed
- Much harder to iterate and isolate the right population as you're reliant on others
- Longer run times = more licence costs

This option uses RPA unnecessarily and still requires resources outside the automation team

ETL extracts from the EMR, database or BI tool

This is a preferred method due to its flexibility, reliability, and accessibility (often)

Pros

- Very secure and stable
- Good option for iteration
- Can be learned, so you are not reliant on others
- Access to tools are common

Cons

- Access and set-up is required
- The query / report writing language of the tools must be learned which takes time to master
 - This may need to come from an outside resource
- Data delays need to be well understood

One of Tarpon's preferred approaches
Data delay requires attention - **especially for BI tools**

Direct query from a database

This is our preferred approach, because it reduces the likelihood of data delays

Pros

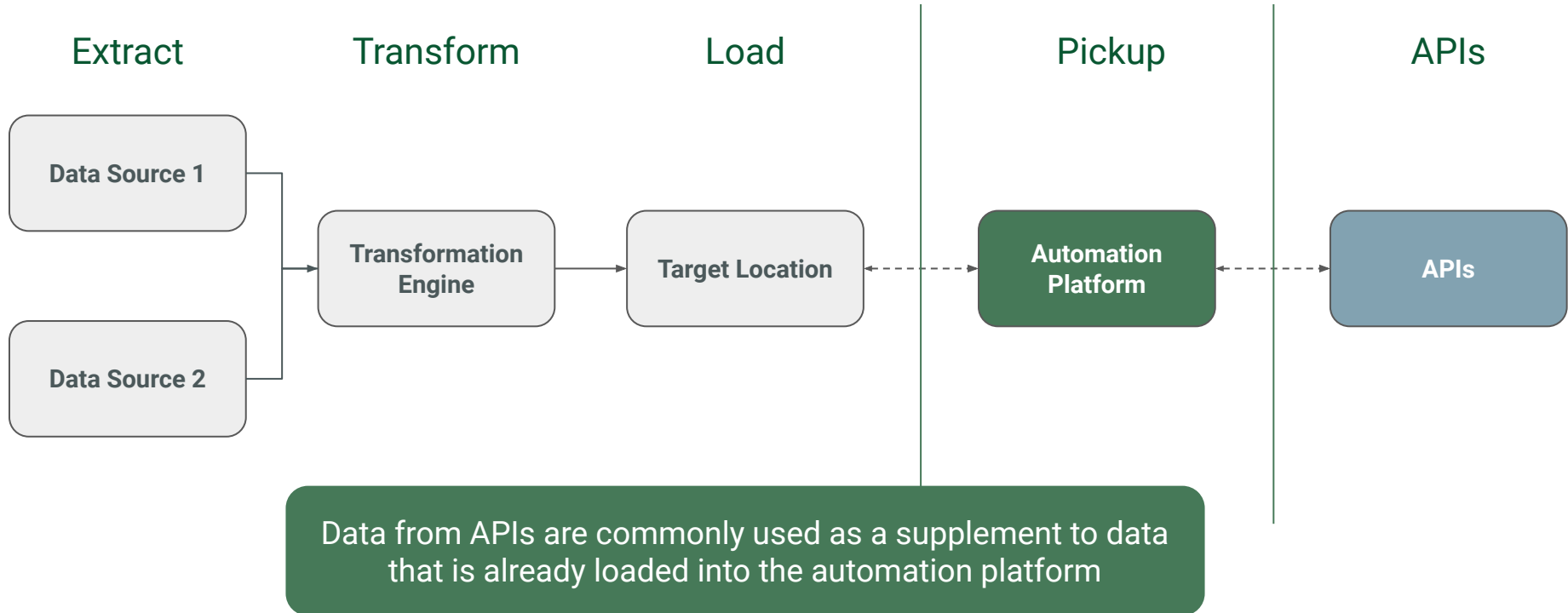
- Very secure and stable
- Best option for iteration
- The data nearly up-to-date
 - Delay is minimal because you are querying the underlying data
- Can be learned, so you are not reliant on others

Cons

- Need to obtain access and set-up if its not done already
- Need to learn a query writing language which takes time to master
- Commonly challenged by IT and security

While this is our preferred approach, it does commonly take additional training to learn the EMR database tables, IT approval, and query-writing skills

High Level Process with APIs



Hierarchy

Tarpon's hierarchy based on:

- Reducing clicks and actions
- Stability
- Autonomy
- Speed and iteration

