

Workday + Greenhouse Data Integration

How to overcome the data challenges with
evergreens, historical data, and data
inconsistencies

PREPARED BY

Jen Ceresia & Gabe Horwitz

eqtble Inc.

Version 1.0

August 12, 2021

Introduction:

Integrating multiple systems into an analytics pipeline is always a challenge. Data is scattered and incomplete - each system has its own interpretation of the same underlying data.

The first challenge is with Evergreen or sourcing job requisitions that are perpetually open. The second challenge is dealing with incomplete & historical data when integrating Greenhouse and Workday.

Challenge 1: Evergreen Requisitions

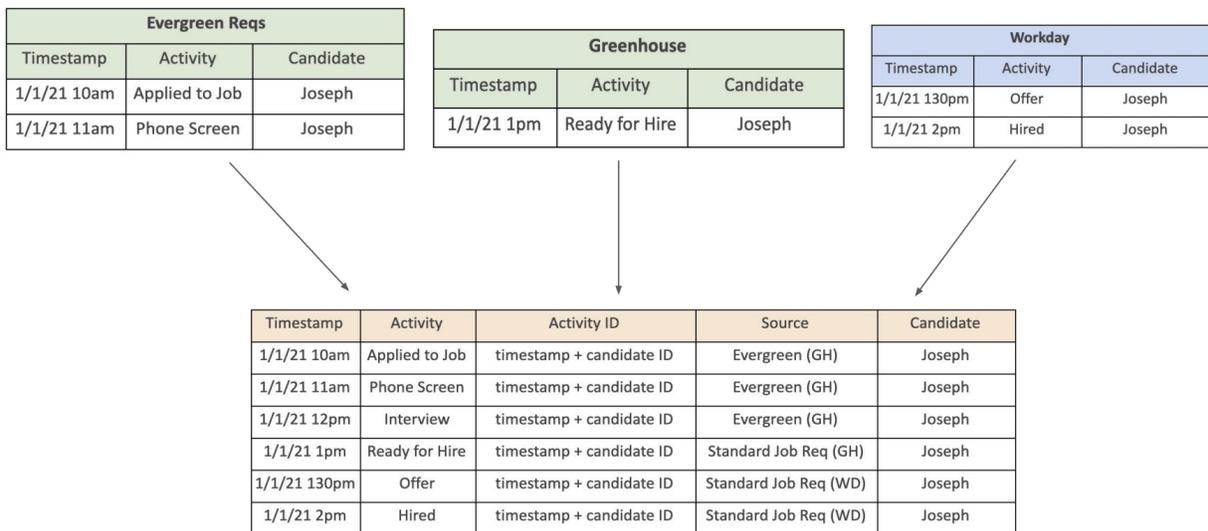
Evergreen or Sourcing requisitions are perpetually open. When a candidate applies to these roles and successfully moves through the process, their application is transferred to another non-evergreen job requisition. This action causes the deletion of candidate process data. Without this data, companies are unable to analyze bottlenecks, potential biases, and identify areas for improvement.

eqtble Solution:

Eqtble proposes using a novel activity-based data model instead of the traditional snowflake or star-based schemas. Doing so enables the tracking of candidate activity regardless of which job requisition they transferred from.

The outcome is a single record for each candidate and application, aggregated from both Evergreens and actual requisitions.

At its core, the activity-based schema turns the data model into a time-series table. Each row in the table represents a single activity taken by a candidate at a point in time. This activity would be a phone screen that happened in the Evergreen requisition. By defining the action as a time series, we can identify the succession of events tied to the candidate, regardless of which job requisition the activity came from.



We then join this data to core base tables that provide us with additional data features to complete the analysis.

This approach enables us to deliver clean candidate process data which allows companies to identify bottlenecks, potential biases and ultimately improve their talent acquisition efficiency and equity.

Challenge 2: Greenhouse & Workday Integration Leads to Incomplete and Unusable Data

Workday and Greenhouse do not have an analytics integration. Most companies try to solve this by just joining the data on a unique candidate ID. However, this method leaves gaps in analyses such as candidate process, offer acceptance, and time KPIs (time to hire/fill/start).

eqtble Solution:

Eqtble proposes an innovative way of joining GH to WD to capture all of the relevant data for the critical people analyses. Instead of joining on candidate ID, eqtble unpacks these large entities so that they can be rebuilt with accurate relationship data. We start by creating two job tables: one in Workday and one in Greenhouse. When creating these tables it is important to ensure the relevant primary key is used. The primary keys in both tables may be different and you would need to introduce a foreign key for accurate joining.

The value of creating these two tables, one from WD and one from GH, is because most job requisition information comes from Workday since that is where companies typically do their headcount planning. These tables will be used to calculate time to fill, open jobs, average days a job is open, recruiter capacity, and more.

Additionally, via this integration, we can immediately identify when a candidate is moved to the offer stage and mark the job status as filled. The integration we built enables the user to have access to correct data faster than the Workday + Greenhouse integration.

1. Next, we isolate two candidate tables - the nuances in each depend on the use case.
 - a. Workday candidate table
 - i. Historical candidate information (if Workday was used as an ATS before Greenhouse)
 - ii. Hires
 - b. Greenhouse candidate Table
 - i. Current candidates
 - ii. Historical candidates

At this point, our algorithm combines the Greenhouse and Workday candidate tables to create a single record for each time a candidate applies to a job.

Application Date	Application ID	Candidate ID	Candidate Name	Status
1/1/21	123	987	Joseph	Rejected
3/1/21	345	987	Joseph	On-Site

Depending on which stage the candidate is passed from Greenhouse to Workday, our model identifies which fields should come from which source. The three possibilities of where a candidate comes from are:

1. Candidates that only exist in Workday before Greenhouse was implemented or candidates that exist in Greenhouse before Workday

2. Candidates that exist in Workday and Greenhouse because they applied to Greenhouse and made it to the offer stage in Workday
3. Candidates that only live in Greenhouse because haven't made it to the offer stage yet

An example would be if Greenhouse candidates who have reached Offer Stage or later are pushed to Workday - our code identifies this and pulls all the fields from before the Offer Stage from Greenhouse and anything after from Workday.

Conclusion:

The result is the following clean and robust data model.

1. Candidates table
 - a. Single row per candidate per application, combining all 3 datasets and field identifies where this candidate source is
 - b. Transformation depends on what source they come from and pulls the relevant fields according to the source of the candidate and stages each system is responsible for.
2. Jobs table
 - a. Single row per job that combines two datasets from Workday and Greenhouse — consolidating to make sure that every JR is captured and its status is updated. For example, if a candidate is pushed to the offer stage in Workday, Greenhouse job status shows 'filled'.
3. Valuable Data
 - a. Candidate Process:

- i. For a single application, see the complete candidate funnel from application review to hired - no matter the system each stage was captured in. For example, a candidate applied through Greenhouse and moved through a phone screen and the interview stages. Once given an offer, the remaining recruitment process (offer signed, background check, marked as hired) was captured in Workday.
- b. Track hires in real-time
- c. Job-status updates regardless of where the hire is made
- d. No duplication of candidates or jobs