



CIVITAS
LEARNING

Guest Lecture

Daniel Dao & Chad Cotton

OVERVIEW

What is Civitas Learning

What We Do

Mission Statement

Demo

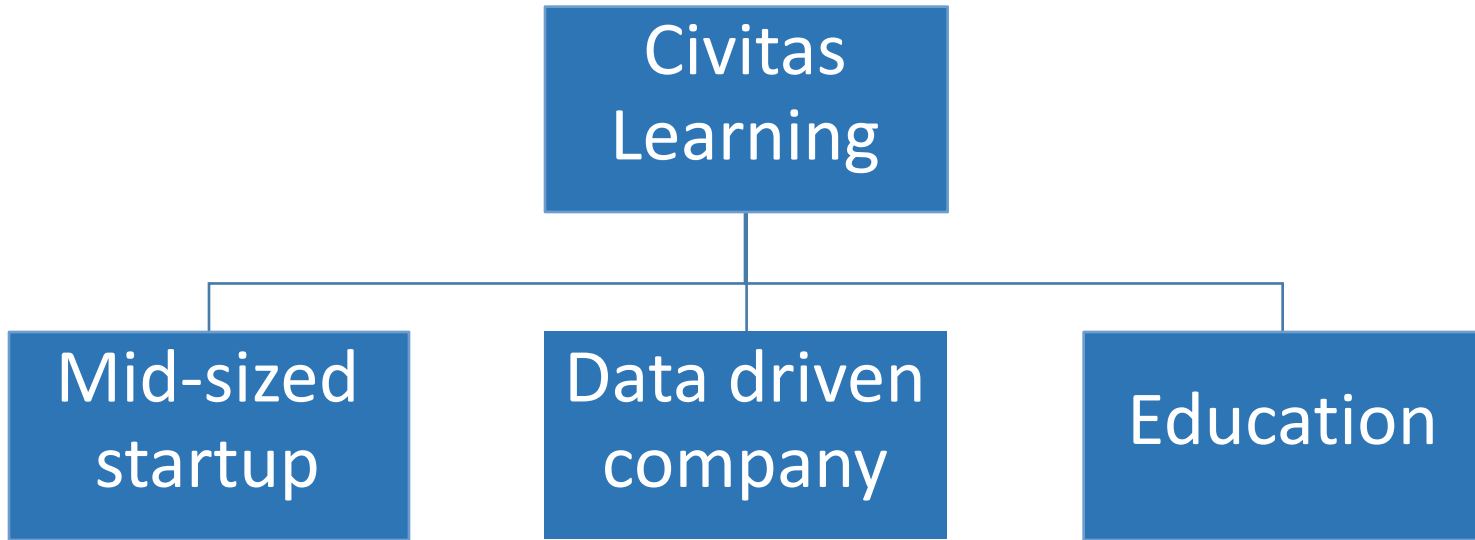
What I Do

How I Use Databases

Chad Cotton



WHAT IS CIVITAS LEARNING



“We partner with forward-thinking colleges and universities, harnessing the power of insight and action analytics to help a million more students learn well and finish strong.” – The Million More Mission



WHAT WE DO

- Work with institutions to provide insights through various applications
 - Inspire



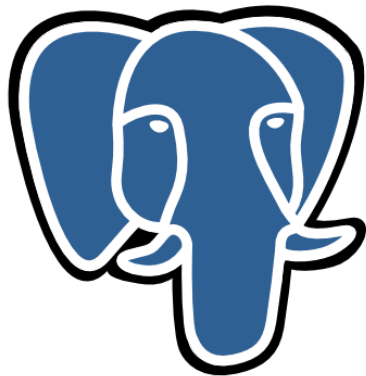
Inspire for Faculty Demo



HOW I USE DATABASES



elasticsearch



PostgreSQL



AMAZON
REDSHIFT



WHAT I DO

- My role in the company
- How my work is broken down
 - Product
 - Dev managers, PSMs, engineers
 - Frontend
 - Work with HTML/CSS/ReactJS
 - Backend
 - Writing APIs
 - Working with models
 - Writing SQL
 - Optimizing performance
 - Writing tests



Chad Cotton - DE

- Graduated from THE University of Texas
 - Bachelors in Physic
 - Bachelors in Computer Science
- Data Engineering
 - I've been working in Data Engineering for about 7 ½ years.
 - I've been at Civitas for a year and a half
 - Before I was a Data Engineer I taught Math and Computer Science in High School



Goals

At the end of this lecture, you will be able to:

- Distinguish jobs in the data industry.
- Describe the process Civitas uses to manipulate data.
- Evaluate methods of loading data into a database.



Jobs in the Data Industry

- When you are out looking for work, you may wonder what types of things to look for:
 - Integration Specialist – getting data out of, or into a database (**ETL: Extract Transform Load**)
 - Database developer – getting data out of a database and using it in software, or creating routines to transform the data within the database.
 - Database Administrator – make the database perform as efficiently as possible.
 - Database architect – Designing schema to organize the data and or model a business process in a database
 - Data Scientist – Much more about math than databases.
 - Data Analyst – Investigate data to find insights.

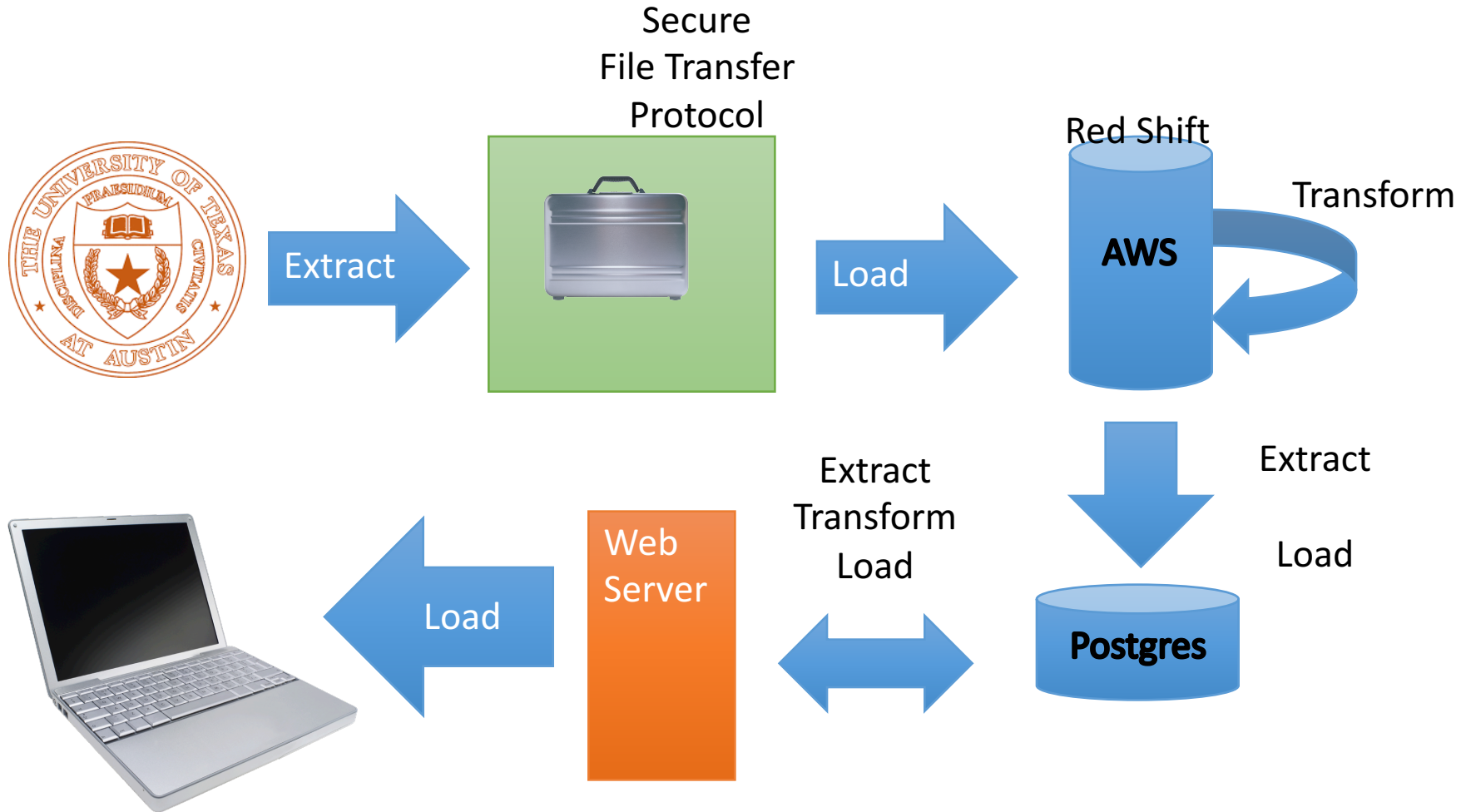


What I Do

- Data Engineer
 - Made up title to describe someone that has to do both the Integration specialist role and database developer
 - What I am not is a Data Architect/Scientist/Admin
- ETL (Extract – Transform – Load)
 - This is an idiom for many combinations of the three.
 - As you will see, it is rarely as simple as just those three.

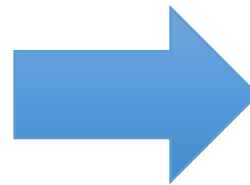


Civitas Data Flow

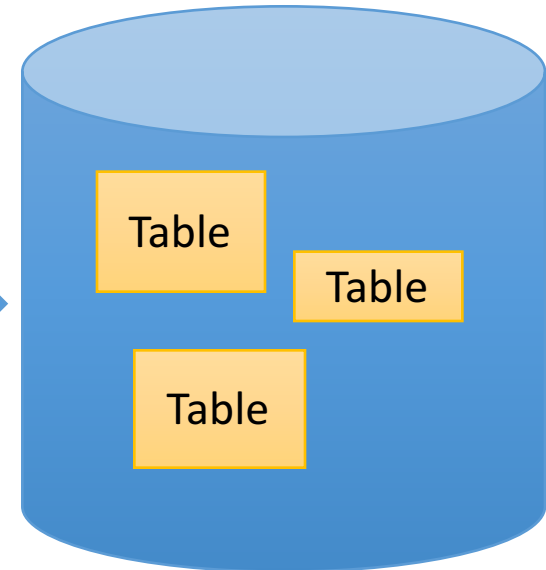


Load

SFTP



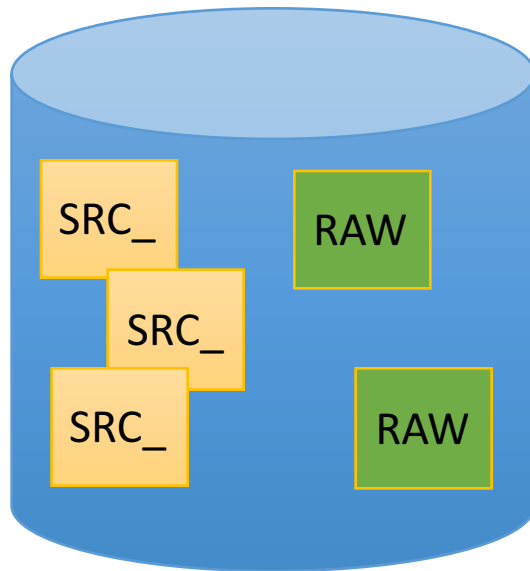
Red Shift



Flat file: Plain Text file that is non-hierarchical, usually in the form of CSV, or TSV. Each row represents one row in the database.



Transform



```
SELECT
  SPBPERS.SPBPERS_PIDM AS raw_person_id
  , SPBPERS.SPBPERS_BIRTH_DATE AS raw_birth_dt
  , SPBPERS.SPBPERS_DEAD_DATE AS raw_death_dt
  , SPBPERS.SPBPERS_SEX AS raw_gender
  , null AS raw_primary_language
  , null AS raw_country_of_origin
FROM
  src_banner_saturn.spbpers
```



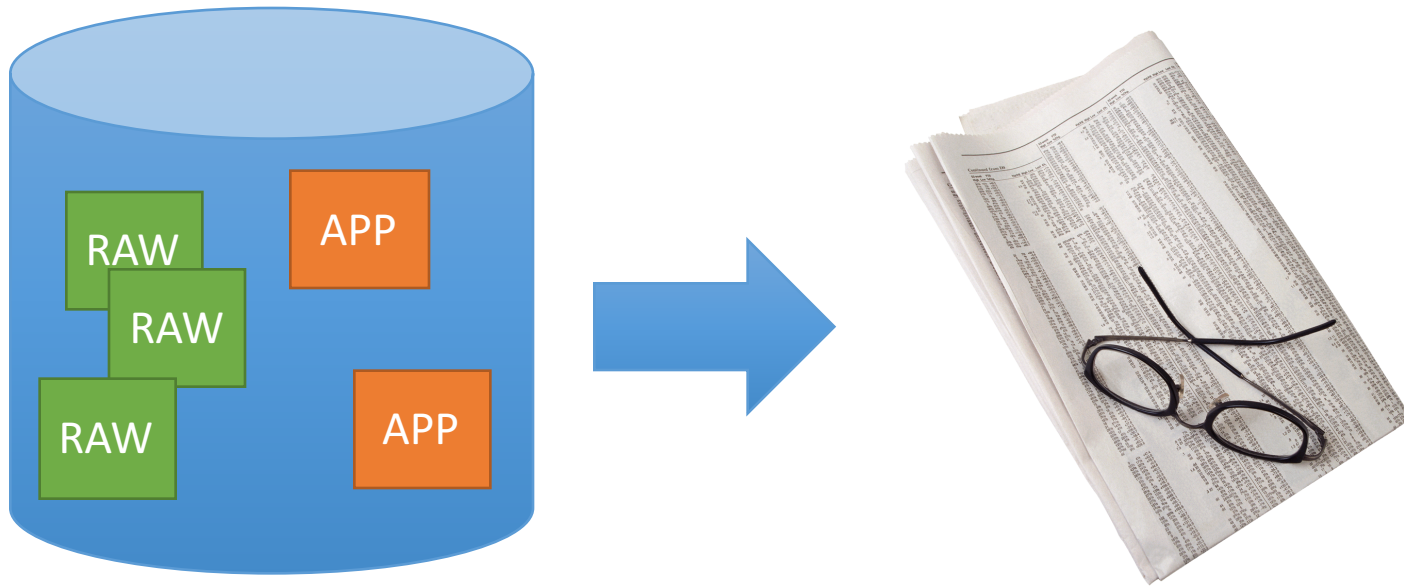
Technology for each phase

Source	Raw Canonical Model	Aggregation Data Science	Application Model
Load data from partner institution.	Mapping Source onto our Base model	Computations to do statistical analytics.	Prepare tables for final output to the application database:
Ingest using Libraries	SQL	SQL/Statistical Modeling	SQL



Extract

- As long as the data is in the tables, there are export commands that can simply dump the data to a file.

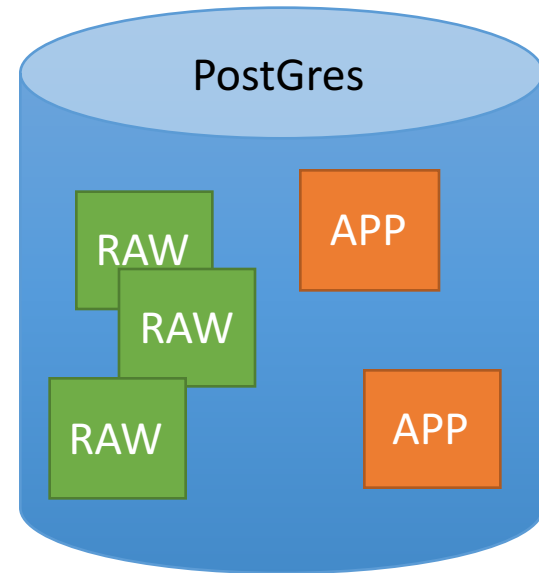


Exported as FLAT FILES

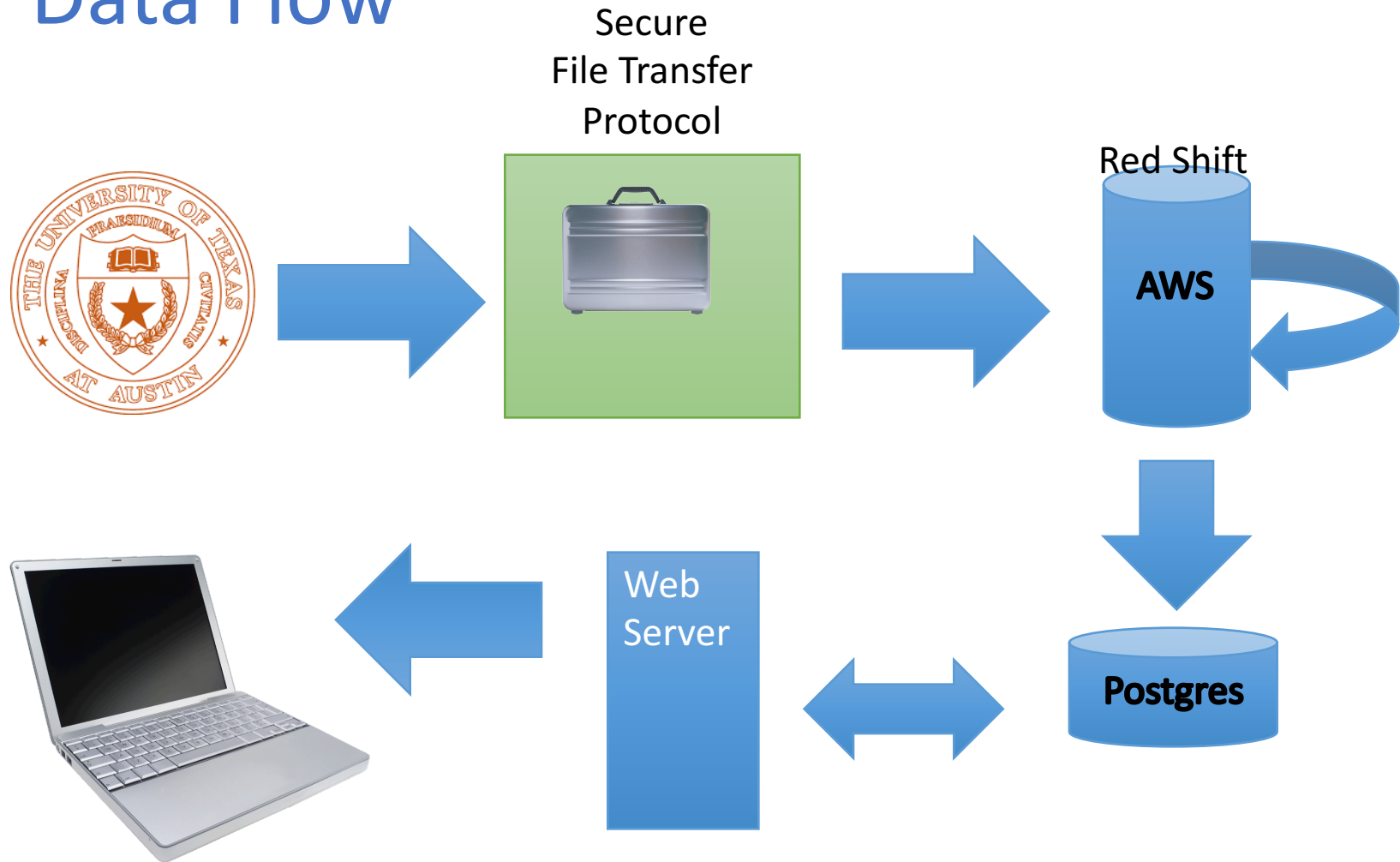


Extract

- As long as the data is in the tables, there are export commands that can simply dump the data to a file.



Data Flow



Summary

- Jobs in the data industry.
- Process Civitas uses to manipulate data.
- Methods of loading data into a database.

- Careers at Civitas Learning.



Questions?

