Introduction to SuperGlue

Overview

The California Community Colleges' (CCC) SuperGlue project, formerly known as Glue, is the codename for a grant-funded effort by the CCC Chancellor's Office. It supports the CCC's many initiatives that require workflows that span the entire system of colleges (currently 116) and serve more than 2 million students per year.

Project SuperGlue combines several technologies:

- application integration
- service orchestration
- API management
- data integration

These technologies mostly leverage existing open-source technology or cloud services from AWS, except in the case of ERP/SIS integration supporting the Student Information Systems in use by colleges in the California Community College system.

This area was either developed or co-developed by the software engineers at the California Community Colleges Technology Center (CCCTC).

These components are associated with SuperGlue:

Component	Definition
CCC Account	Application that manages student account information and is a data source that feeds in to SuperGlue.
CCC Apply	Application that enables prospective students to set up a profile and is a data source that feeds in to SuperGlue.
Orchestration Layer	Conductor workflows for delivering data.
College Adaptor	SIS-specific staging tables for the three supported ERPs (Banner, Colleague, PeopleSoft) and generic MSSQL data.
SuperGlue for Apply	Delivers student application data directly to the College's staging tables via the SuperGlue College Adaptor.
Multiple Measures Placement Service (MMPS)	Delivers placement recommendations based on high school student transcript information and/or self-reported transcript information directly to the college's SIS-specific staging tables.

SuperGlue serves many significant workflows. The CCCTC sees thousands of transactions flowing across the SuperGlue fabric every day.

Flow of Data

The diagram below illustrates how SuperGlue has integrated many CCCTC enterprise applications with the top three ERPs in use by our colleges.

SuperGlue has also been extended to provide support for writing transactional data into the CCCData data lake, which is then processed into the Data Warehouse.

The diagram below shows the most common usage of SuperGlue, with the Orchestration Layer feeding to the college's built-in tables.
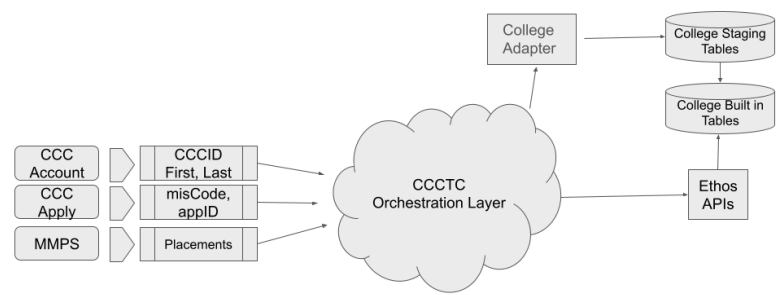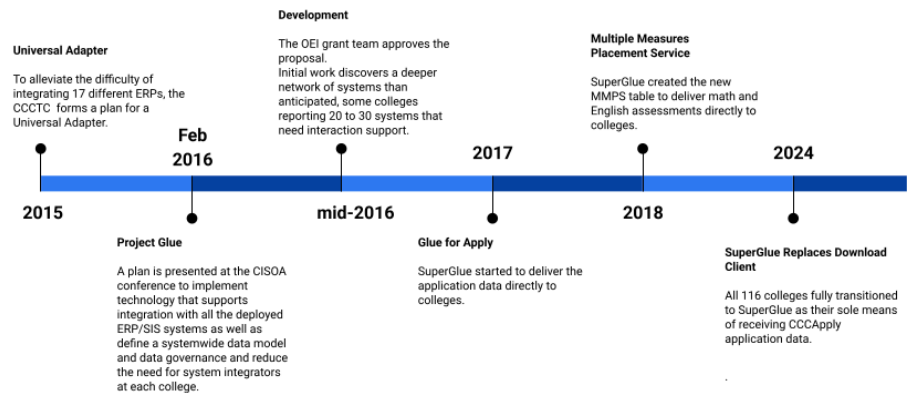

Diagram 1: Common Usage of SuperGlue

Project History

Project Glue started in 2016 in response to the growing need for colleges to be able to share data from different ERP and SIS systems. As new needs were identified, grant funding was secured and new features were added.

To fund the project, multiple vendors were considered and multiple grants were used to fund different initiatives of SuperGlue.

Grants Utilized	• Online Education Initiative (OEI) • EPI • CAI • CAP • DSP • SIP
Vendors considered for components of SuperGlue	• Informatica • Dell Boomi • Mulesoft • Datatel (Colleague) • Banner

	• CampusSolutions (PeopleSoft) • YOUnite • Tibco • Netflix Zuul • Netflix Conductor • Illuminate
Approvers	• The Director's Collaborative (CCC Project Directors and CCCCO staff) • CCCCO leadership • TTAC (Telecommunications and Technology Advisory Committee) • SAC (Systemwide Architecture Committee) • OEI Program Leadership and Steering Committee • CISOA conference attendees

To learn more about what systems were considered and why specific decisions were made with regard to the beginning of SuperGlue, see the Project History section in the document [History of SuperGlue](#).

CVC-OEI Data Integration

SuperGlue has the ability to transmit student application information from the "home" college to the "teaching" college in support of a student registering for an online course. The storage of data in a college ERP is not standardized between different SISs. Furthermore, there are no agreed upon standards even among colleges using the same SIS.
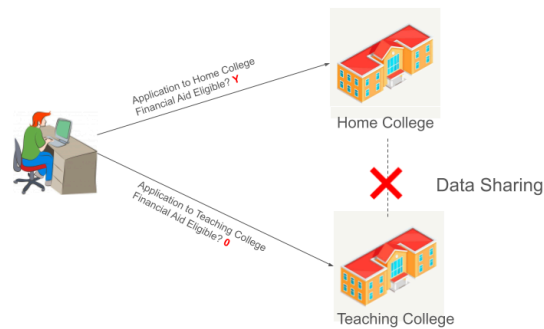

Diagram 2: Data Sharing

Data could be written into the teaching college ERP as a SQL write, but college concerns over data integrity and governance made that problematic. Instead, treating the transfer as a defined transaction via a web service was preferred. The transfer "across" the wire needs to be in a "standard" or canonical format. As such, utilizing staging tables within each Student Information System provides a standard way of delivery, while then allowing colleges to process each record into the built-in tables as desired to fit individual college needs.

College Adaptor

SuperGlue for Apply securely delivers CCCApply records - in near real-time - to a staging table in your college IT environment via the SuperGlue College Adaptor. Records include the entire CCCApply dataset, including the college's custom Supplemental Questions.
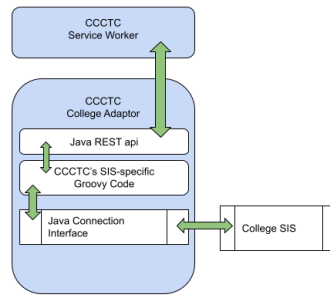
Diagram 3: College Adaptor

Data Lake and Data Warehouse

As a part of the evolution of SuperGlue, these features have been implemented:

- application to college
- determining placement based on multiple measures business rules
- bi-directional communication of fraud reports between colleges

The Data Warehouse is a shared space where researchers and the Chancellor's office can get data for all colleges.

Diagram 4 shows the state of some of the SuperGlue services. The elements are fully deployed and live supporting all colleges. The college Adaptor is available to all colleges today including schools that use Banner, Colleague, PeopleSoft, Ethos, and those that use a custom SIS.
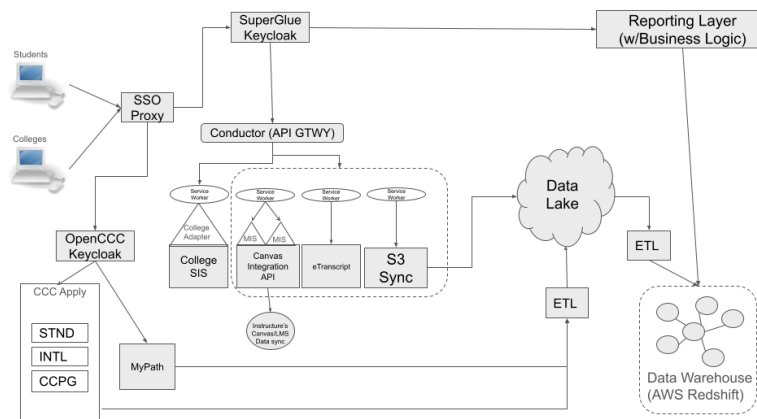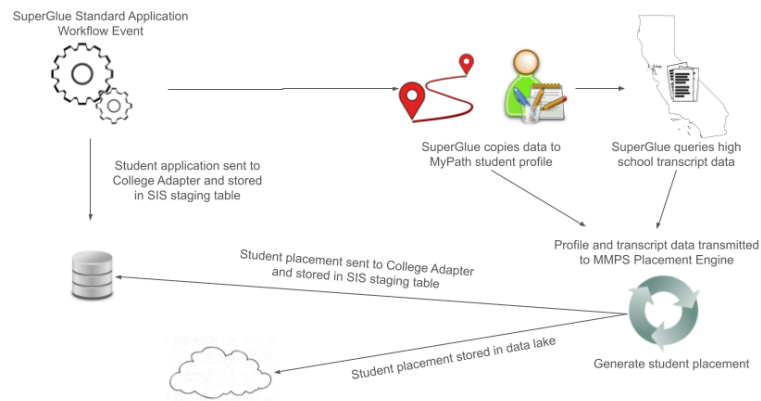

Diagram 4: SuperGlue Architecture

SuperGlue's impact can be summarized with an example of a potential student who creates an account with OpenCCC for the first time.

1. The student explores careers using Career Coach, then completes the application in CCCApply.
2. SuperGlue kicks off a workflow event.

OpenCCC account creation   Career exploration using Career Coach   Complete application using CCCApply   SuperGlue Workflow Event

3. SuperGlue transfers application data to the associated SIS college Adaptor.
4. The Adaptor transforms the data to the SIS-specific system.

5. SuperGlue copies a subset of the application data to the MyPath student profile.

6. SuperGlue queries California College Guidance Initiative (CCGI) for that student's high school transcripts.

7. Profile and transcript data are transmitted to the MMPS Placement Engine.

8. Student placement is generated, then SuperGlue transmits the placement to the College Adaptor's staging table, as well as storing that data in the data lake.

College Adaptor Features

SuperGlue integrations deliver data to colleges via the SuperGlue College Adaptor, which is installed behind the college's firewall at the college's request. The SuperGlue College Adaptor provides these features:

- ability for colleges to receive student application data and verified transcript student placement records directly to their college's IT environment
- support for common workflows and data exchange between systemwide products such as CCCMyPath and the college SIS, and delivers student data to the CCC Data Warehouse

Glue for Apply

SuperGlue's Glue for Apply feature delivers student application data directly to the College's SIS via the SuperGlue College Adaptor. This includes the following CCCApply applications:

• Standard Application

• International Application

• California Promise Application

• Non-Credit Application

Glue for Apply delivers the full dataset for each of these applications directly to colleges within seconds of the student submitting them.

Multiple Measures Placement Service (MMPS)

SuperGlue's College Adaptor supports the CCCTC MMPS feature by delivering verified student transcript placement information and/or self-reported transcript placement information directly to the college SIS.
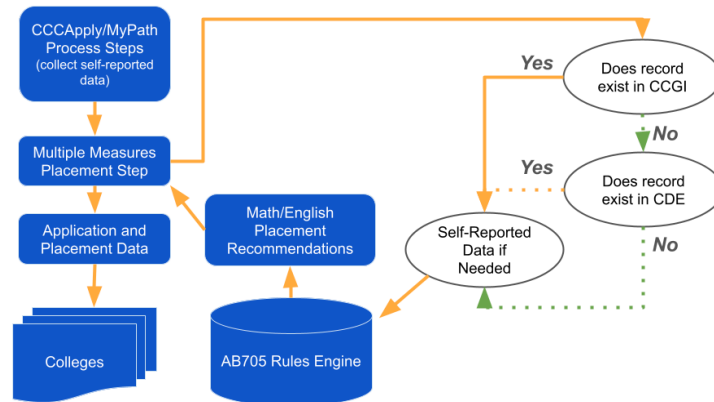
Diagram 7: MMPS Placement Logic

Data Lake and Data Warehouse Integration

CCCData is the fabric of data that follows the Gartner Group's Virtual Data Warehouse model. It is a combination of data sources stored, governed, and presented in ways which are the most effective for the intended use. A data lake is deployed to capture data, in its raw form, from operational data stores.

Multiple Measures (summarized above) also supports the use of Conductor to ensure its data is stored in the data lake as well as in the colleges' SIS systems. SuperGlue writes MMPS placement data to an S3 bucket using the s3Sync microservice.

Other data integrations, particularly those involving table replication, are not as well suited to SuperGlue. For those, we employ AWS services like the Data Migration Service (DMS) and AWS Glue.

The Data Lake is comprised of multiple S3 buckets, some of which SuperGlue writes to, some of which do not involve SuperGlue.

SuperGlue is not integrated with the Data Warehouse directly. Instead, SuperGlue passes data to the data lake where it is then processed by an AWS Glue job into the Data Warehouse.

What is GraphQL?

"GraphQL is an open-source data query and manipulation language for APIs, and a runtime for fulfilling queries with existing data. GraphQL was developed internally by Facebook in 2012 before being publicly released in 2015." - Wikipedia

GraphQL supports reading, writing (mutating), and subscribing to changes to data in a way which is a logical extension of the web service model. The benefit to SuperGlue is that a data model can be "published" in a way that is independent of how the data is retrieved, in effect turning all of the data at CCCTC into one large data repository. An application interacting with our GraphQL implementation does not need to know that it is pulling data from Colleague, Canvas, and perhaps CCCApply. It does not need to determine how to call the College Adaptor, or the Canvas REST APIs, or query the CCCApply database.