

California Airports - 2020

Data producers make important work that can lead to further scientific advancement through data sharing and reuse after initial publication. Unfortunately, data producers' work often possesses significant impediments to reuse due to technical language or limited exposure in scientific databases. Digital publishers bridge the discovery gap between data producers and data users by transforming data and providing easy access options for users. One such example of a valuable dataset currently in need of a digital publisher is the Public Airports 2020 dataset published by the California Department of Transportation (2021). Through extensive work, the Public Airports 2020 dataset appeals to a new set of data users by simplifying, restoring functionality, and increasing usability.

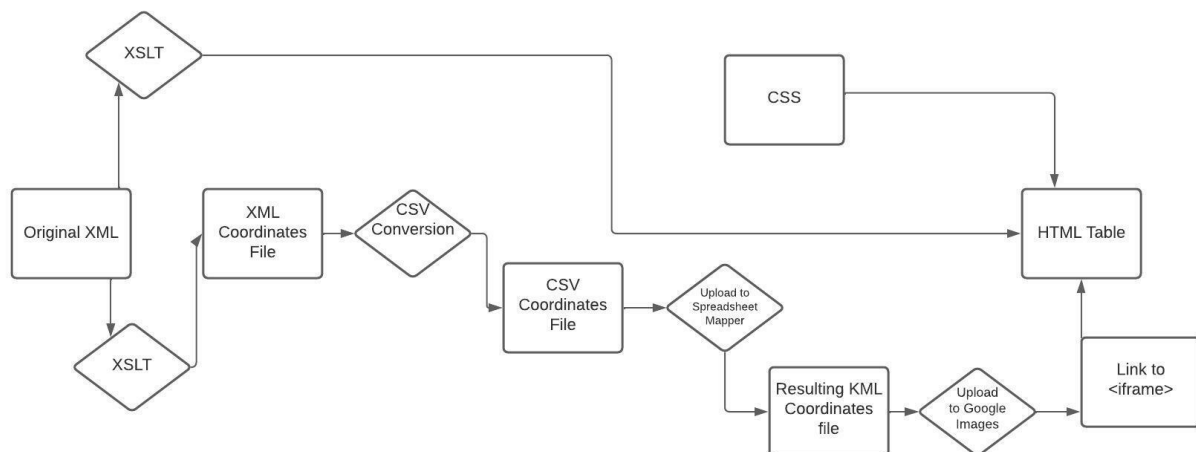
The project aimed to transform a public dataset of California airport locations and data that was made and published online by the state of California. The published data contains several inhibitors to meaningful usability by larger audiences, including: technical jargon (e.g., latitude and longitudinal data, runway shape) and relational database terminology (e.g., Object ID). Additionally, since its release, the dataset's column linking to the airports' websites has stopped working, frustrating users wanting to know more about a specific airfield. In its current state, the dataset has limited appeal to the public, and reflects this with the low metrics of 66 views and 7 downloads since publication (2021).

The new audience for the dataset will be local laypeople interested in airplanes. For instance, local Californians wanting to explore skydiving trips, or those curious about pursuing flying lessons. Others may be photographers wanting quality locations to take pictures of planes taking off or landing. Finally, a small user base may have private airplanes and are in need of planning trip logistics for their pilots or party. All of these users need a simplified dataset to

quickly and conveniently obtain essential information; one that limits technical jargon, unnecessary information, and is easy to access.

The project possessed a unique workflow as two separate processes took place which were not connected until reaching the finished product. As can be seen in **Figure 1**, no overlap was present between the goals of creating an <iframe> and the making of a stylized HTML file until the end. Remarkably few problems occurred when connecting the <iframe> to the HTML. A description of the project workflow follows, with a discussion of challenges encountered.

Figure 1



Transformation of the data included a plethora of steps, leading to a highly redesigned website. For starters, an XSLT transformation occurred on the originating XML file to achieve two ends: (1) the elimination of technical or non-functioning data and (2) the change from XML to HTML for web publication. In this process, columns: “OBJECTID,” “shape,” “classcode,” “district,” “f5010url,” “latdms,” “pmtlat,” “longdms,” “arplatdms,” “arplongdms,” “arplats,” “arplongs,” and “pmlong,” were not carried over into the new HTML file. Additionally, “latdd” and “longdd” entities were combined through a concatenate function to produce a workable Google Maps link. By cutting and condensing the dataset, it became less intimidating for users to approach. Moreover, the addition of Google Maps URLs provides a realistic use of GIS data, as laypeople are unlikely to understand longitude and latitude coordinates in a meaningful way. The

resulting HTML table also keeps the email link to airport managers by the creation of an additional concatenate function. A CSS file was later created to add color and readability features (e.g., sticky table head, tr hover color; table column colors) to the HTML file for ease of reading.

Another XSLT file was created from the originating XML file to pursue the making of an embeddable Google Images map with pins representing the airports. This XSLT transformation kept the result as an XML file, but grabbed just the name of each airport and its corresponding longitude and latitude data. The resulting XML file was converted to CSV for accessibility with Google Spreadsheet Mapper v3.2. The CSV file was uploaded into the spreadsheet, and a KML file resulted. This KML file was then added to Google Maps, which established an <iframe> to copy and embed within the HTML file discussed above.

Few significant issues impacted the publication process. One issue was sizing and aligning the table and the embedded map in the final HTML file. This proved to be a challenge, but a viewable final form took hold after consulting references. Another challenge was the creation of the correct and usable Google Maps URL. The concatenate function for this had to include two separate data sources with a URL string between them. Moreover, the URL contained an ampersand, creating complications with validity. Luckily research and testing solved these problems.

The 2020 California Airports data publishing endeavor showcased the value of data publishing for the scientific community, through adapting datasets to appeal to broader audiences. With a few simple XSLT transformations the dataset was simplified of technical language unnecessary for most potential users. Further, a basic CSS file added color and design to make the dataset a pleasant reading experience. Finally, the use of concatenate in the XSLT and the embedding of Google Images with pins added critical functionality for geospatial understanding that did not exist before (outside of technical latitudinal and longitudinal data).

Critically, the skills honed here can easily apply to other dataset and research opportunities in the future.

References

California Department of Transportation. (2021, February 8). Public Airports 2020.

<https://data.bayareametro.gov/Operation-Use/Public-Airports-Point-2020-/dncy-egiu>