

# GIS Community of Practice Monthly Forum

Last Wednesday of the month

Wednesday, September 25<sup>th</sup> , 2024  
1:30 – 2:30 PM



California  
DEPARTMENT OF TECHNOLOGY  
STRATEGY INNOVATION DELIVERY



# GIS Community of Practice (CoP)

- Welcome to the GIS CoP forum.
- For the best experience, please use your computer to join the meeting.
- Mute your audio.
- Turn off your video unless you're presenting or in active discussion.
- Use the raise hand button or the meeting chat for comments and questions.
- We will begin shortly.

# GIS CoP Agenda

## Welcome

- Lothar Petrik, State of CA GIO

## Main Topics

- Carmen Gonsalves ([carmen.gonsalves@state.ca.gov](mailto:carmen.gonsalves@state.ca.gov)), Geographic Information Services, California State Geoportal Enhancements
- Nick Santos ([nick.santos@state.ca.gov](mailto:nick.santos@state.ca.gov)), Geographic Information Services, Fully Automated Data Pipelines

## Announcements

- Lothar Petrik
- Open to participants – (Job openings, events, looking for assistance)

## Conferences/ Events

- GIS Day Celebration – November 21 – Sacramento, California
- LA County GIS Day – November 20, Gloria Molina Park, <https://gis-day-lacounty.hub.arcgis.com/>

# CA Geoportal Update

Presented by Carmen Gonsalves  
IT Specialist I  
Geographic Info Services

# Updates

- Changing the face of the Geoportal.
- Organizations and Featured Apps Pages remain the same – call for more apps!
- Added links to the CA GIS Community of Practice and the CA Open Data.
- Revising the back end- we will be reaching out to Organizations to move all your data to the CAOpenData groups.
- Adding a Local Government page containing links to the County Hubs or main pages.
- Updated the Training page to include directions for uploading the data to the groups.
- Looking for Data Analytics from users.

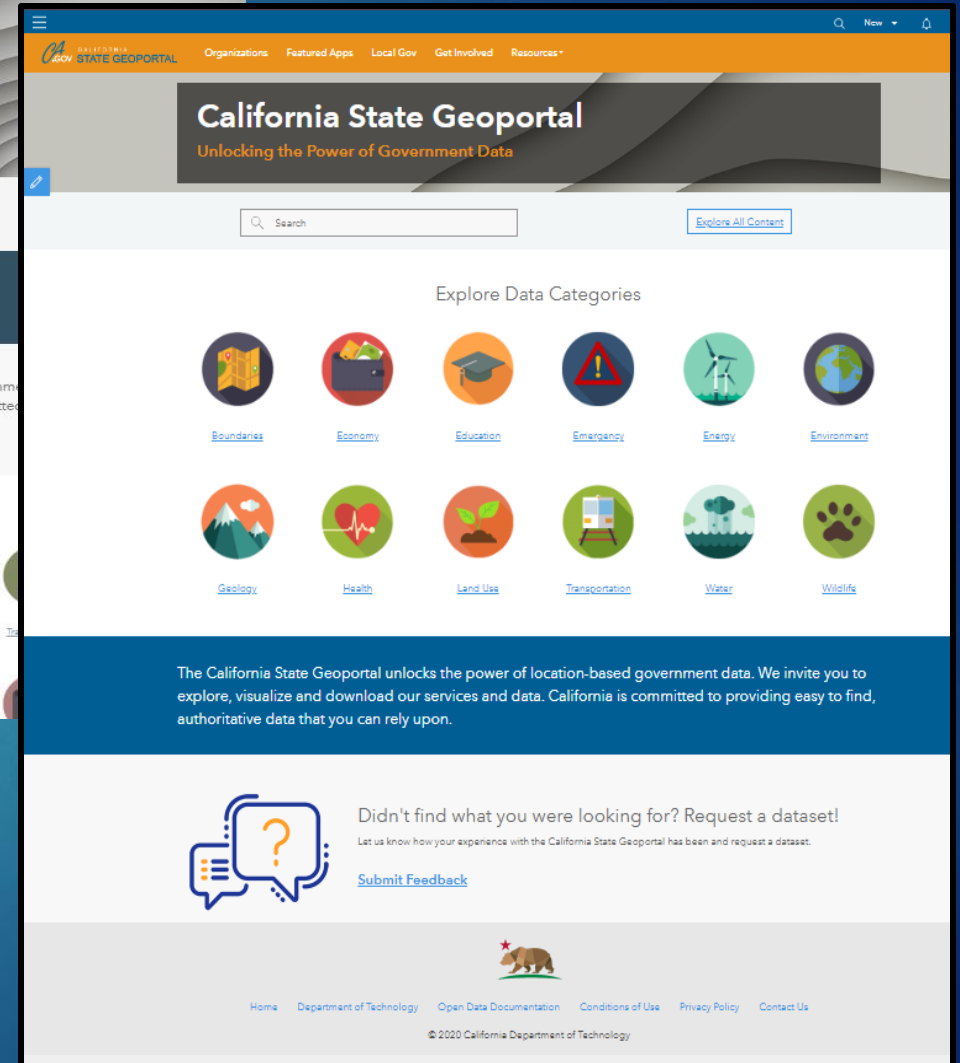
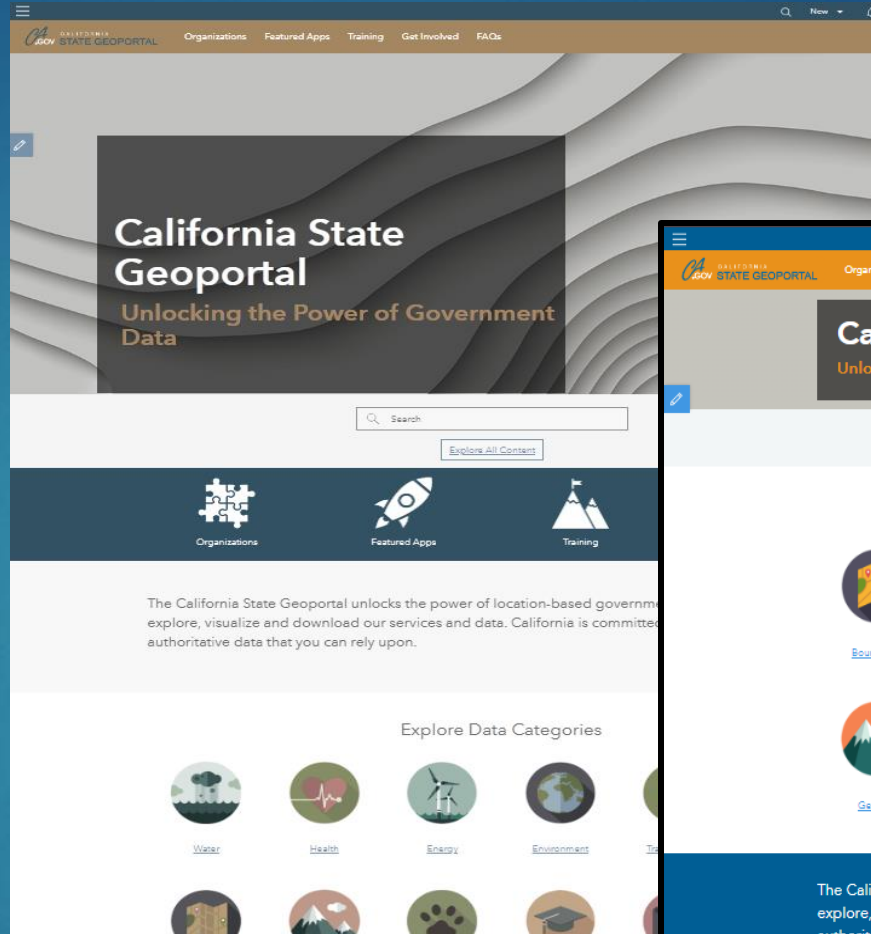
# Layout

- ▶ By changing the layout of the Geoportal landing page, it is data focused.

We removed this row of buttons because it was mirroring the tabs at the top of the page.

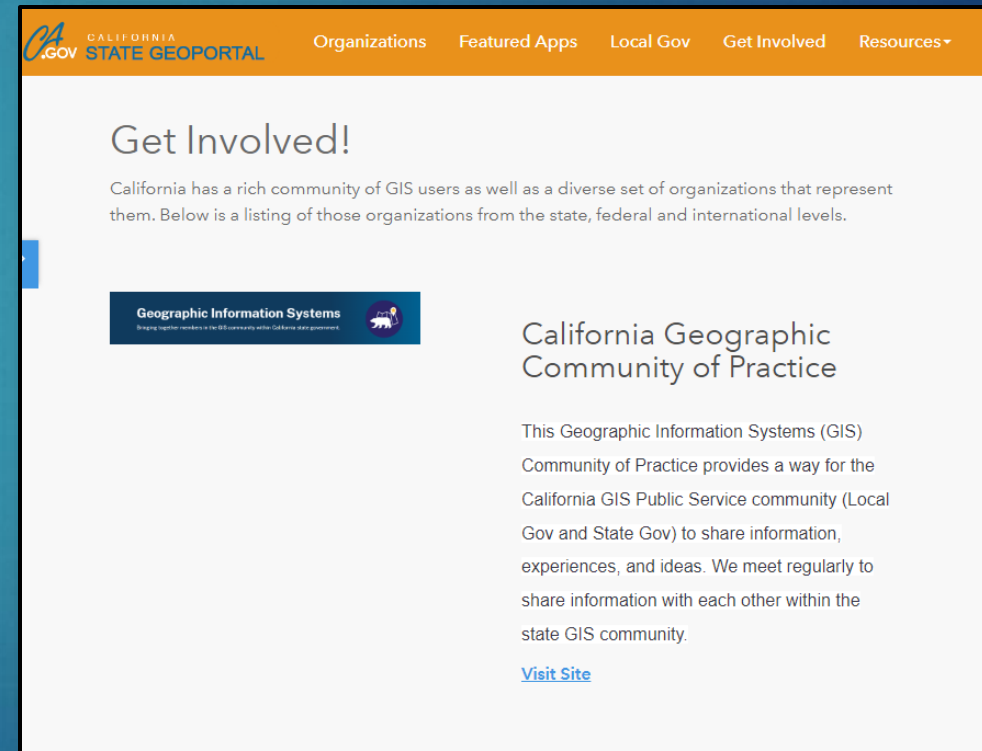


Then we moved the text below the categories icons.



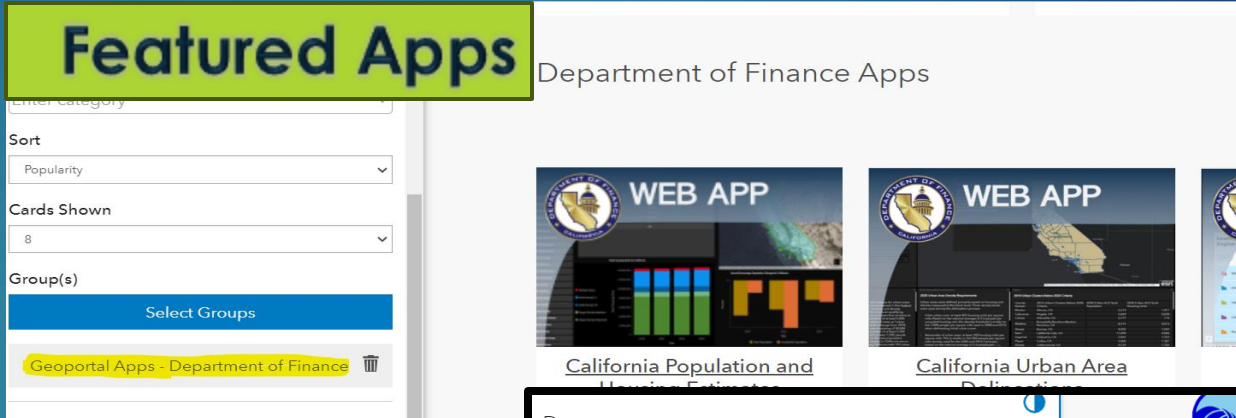


- ▶ Added a CA Open Data link under the Resources and a Community of Practice link on the Get Involved Page.



# Back end

- ▶ Users should move all data to the CAOpenData group.
- ▶ On the back end we will redirect the pages to also only access one group.
- ▶ This means users only have to post to one and not two!



**Featured Apps**

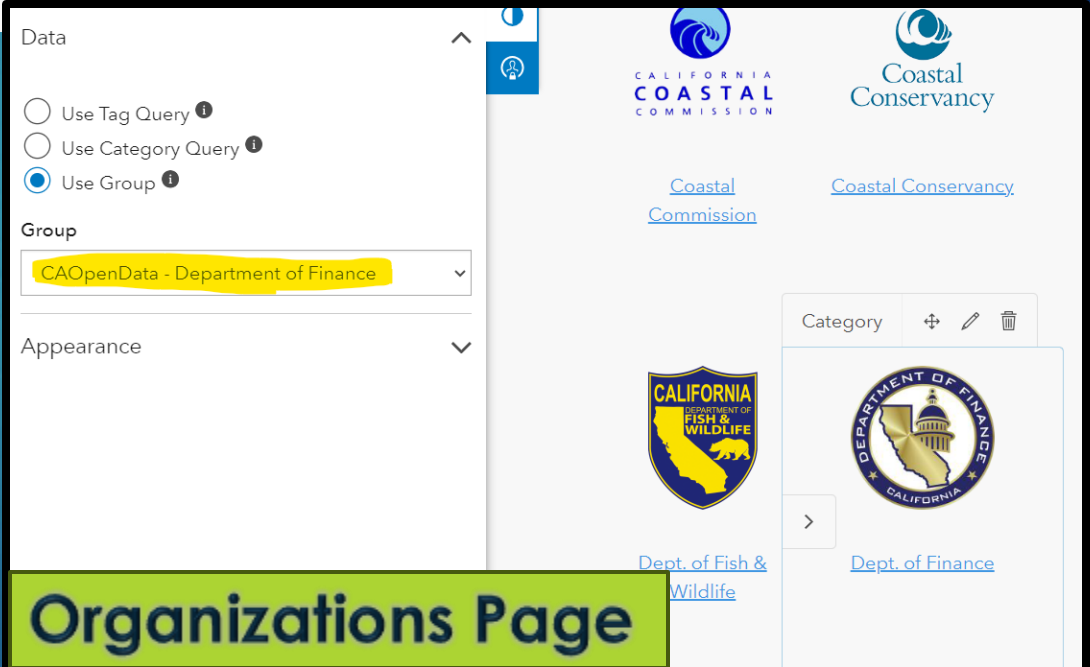
Department of Finance Apps

Sort: Popularity

Cards Shown: 8

Group(s): [Select Groups](#)

[Geoportal Apps - Department of Finance](#)



**Organizations Page**

Data

Use Tag Query

Use Category Query

Use Group

Group: [CAOpenData - Department of Finance](#)

Appearance

Coastal Commission

Coastal Conservancy

Dept. of Fish & Wildlife

Dept. of Finance



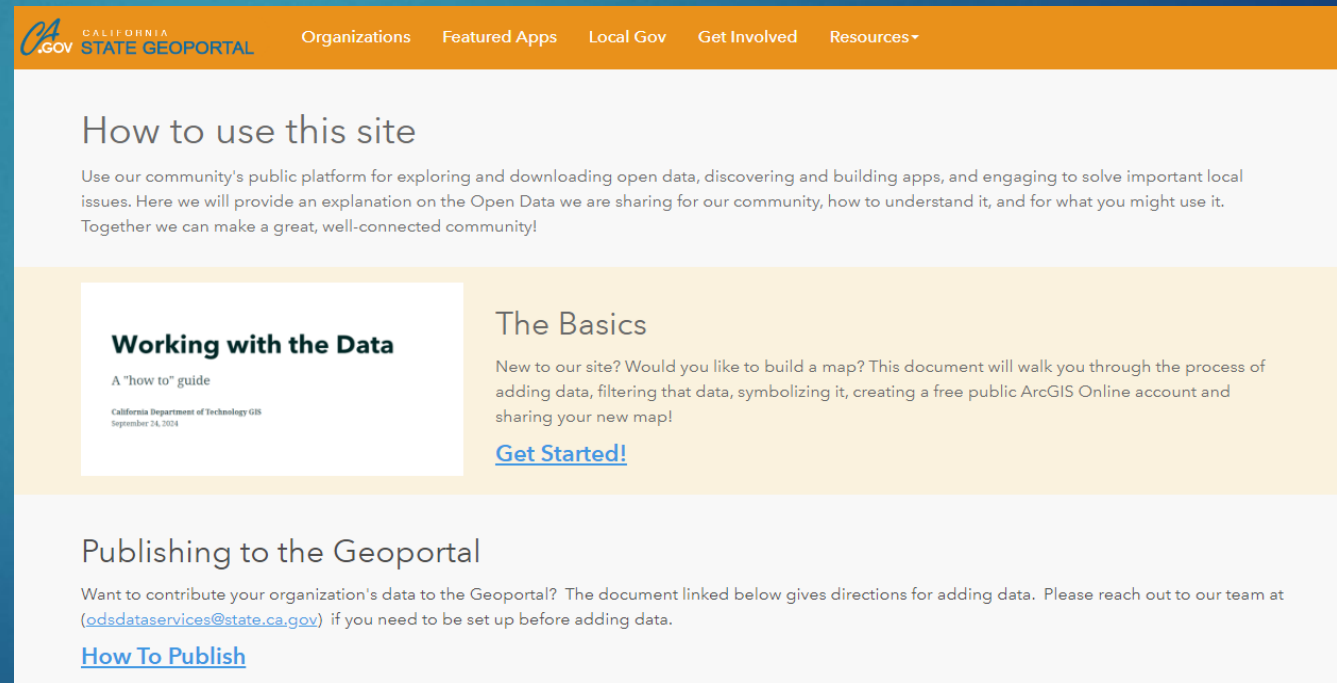
# Local Gov Links

- ▶ To include local Government, we created a page encompassing all 58 counties and links to their GIS pages, Hubs or Portals.



# Publishing to the Geoportal

- ▶ The Working with the Data Storymap has been updated to include a section about publishing.
- ▶ A separate Publishing to the Geoportal section has been added mirroring the information added to the Storymap.



The screenshot shows the California State Geoportal website. The top navigation bar includes the logo 'CA .GOV CALIFORNIA STATE GEOPORTAL' and links for 'Organizations', 'Featured Apps', 'Local Gov', 'Get Involved', and 'Resources'. The main content area is titled 'How to use this site' and contains introductory text. Below this, there are two columns of content: 'Working with the Data' (a 'how to' guide) and 'The Basics' (a document for new users). At the bottom, there is a section titled 'Publishing to the Geoportal' with a link to 'How To Publish'.

CA .GOV CALIFORNIA STATE GEOPORTAL Organizations Featured Apps Local Gov Get Involved Resources ▾

## How to use this site

Use our community's public platform for exploring and downloading open data, discovering and building apps, and engaging to solve important local issues. Here we will provide an explanation on the Open Data we are sharing for our community, how to understand it, and for what you might use it. Together we can make a great, well-connected community!

### Working with the Data

A "how to" guide

California Department of Technology GIS  
September 24, 2024

### The Basics

New to our site? Would you like to build a map? This document will walk you through the process of adding data, filtering that data, symbolizing it, creating a free public ArcGIS Online account and sharing your new map!

[Get Started!](#)

## Publishing to the Geoportal

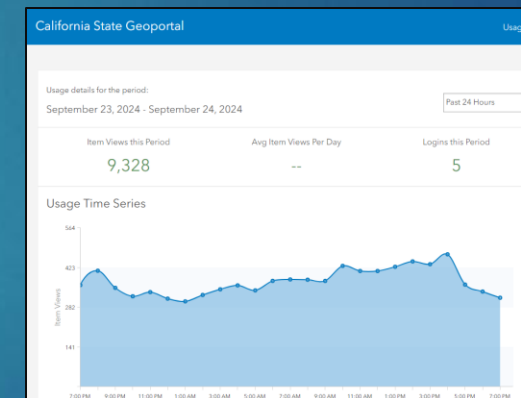
Want to contribute your organization's data to the Geoportal? The document linked below gives directions for adding data. Please reach out to our team at ([odsdataservices@state.ca.gov](mailto:odsdataservices@state.ca.gov)) if you need to be set up before adding data.

[How To Publish](#)

# Analytics

- ▶ What are we looking for? A deeper understanding of data usage.
- ▶ What we need from you- since we aren't the layer owners, we need the admin from the Orgs to run statistics on your items in AGOL. Then share those stats with us.
- ▶ How do you run these reports? Run an item view count report. We can pull Item IDs from the metadata. We'll then filter by those IDs.
- ▶ What date range? Preferably for a year, 2023, if not then a few months at least.

Organization: California Department of Technology									
StartTime: Monday, 29 July 2024 00:00:00 UTC									
EndTime: Sunday, 04 August 2024 23:59:59 UTC									
Title	Item ID	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY	SUNDAY	Total view count
COVID-19	172f54a7d	1	1	1	1	1	1	1	7
mental-he	c683ebc06	1	0	0	0	1	0	1	3
Trailer Bill	d4ed28ad	0	0	0	0	0	0	1	1
GOBiz Wo	89463e28c	0	1	0	1	0	0	0	2
SD_Draft_	af05c09cff	0	0	3	0	0	0	0	3
EV-Capab	96ec9749f	0	0	0	0	0	0	1	1
Reopenin	8e1a871fc	0	0	0	0	0	0	1	1
USA State	2d12ea11f	0	3	0	0	0	0	0	3
TEST: Loc	7bd145e9f	0	0	4	2	2	0	0	8
California	91fd7ce8f	38	39	42	25	27	6	14	191
California	e01e7574f	0	2	0	1	0	0	0	3



For more information on reports:

[Supercharge your ArcGIS Online Organization Management with Reports](#)

# Questions?

Contact us at [odsdataservices@state.ca.gov](mailto:odsdataservices@state.ca.gov)

This presentation will be posted with the Meeting slides on the CoP webpage.

# Fully automated spatial data pipelines

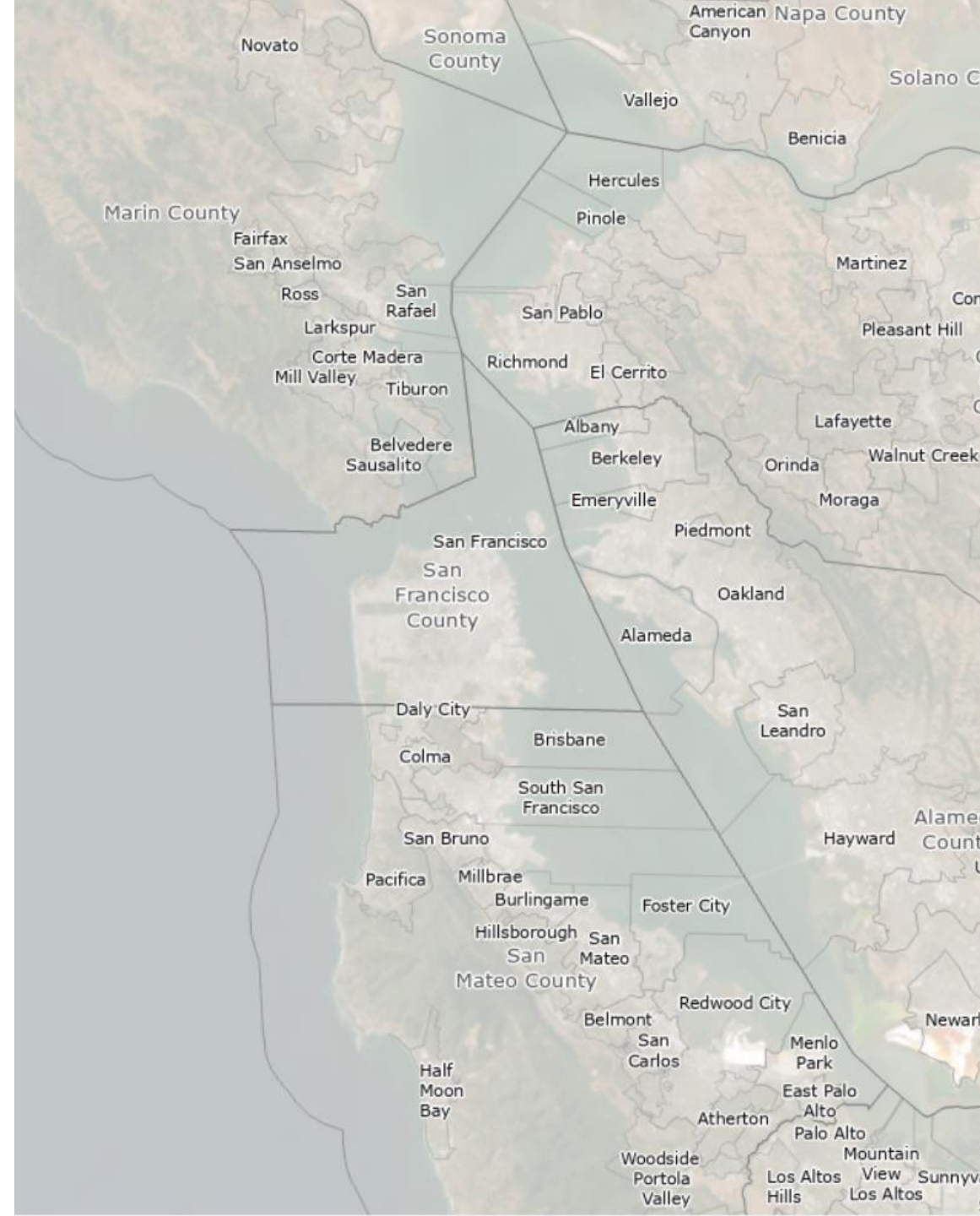
Nick Santos, Carmen Gonsalves, Lothar Petrik

California GIS CoP – 9/25/2024

[nick.santos@state.ca.gov](mailto:nick.santos@state.ca.gov)



California Department of  
**Technology**



# Topics

---

City/County Layer

---

ArcGIS Online Notebooks

---

Packages/Testing in GitHub and  
CalEnterprise (Azure DevOps)

---

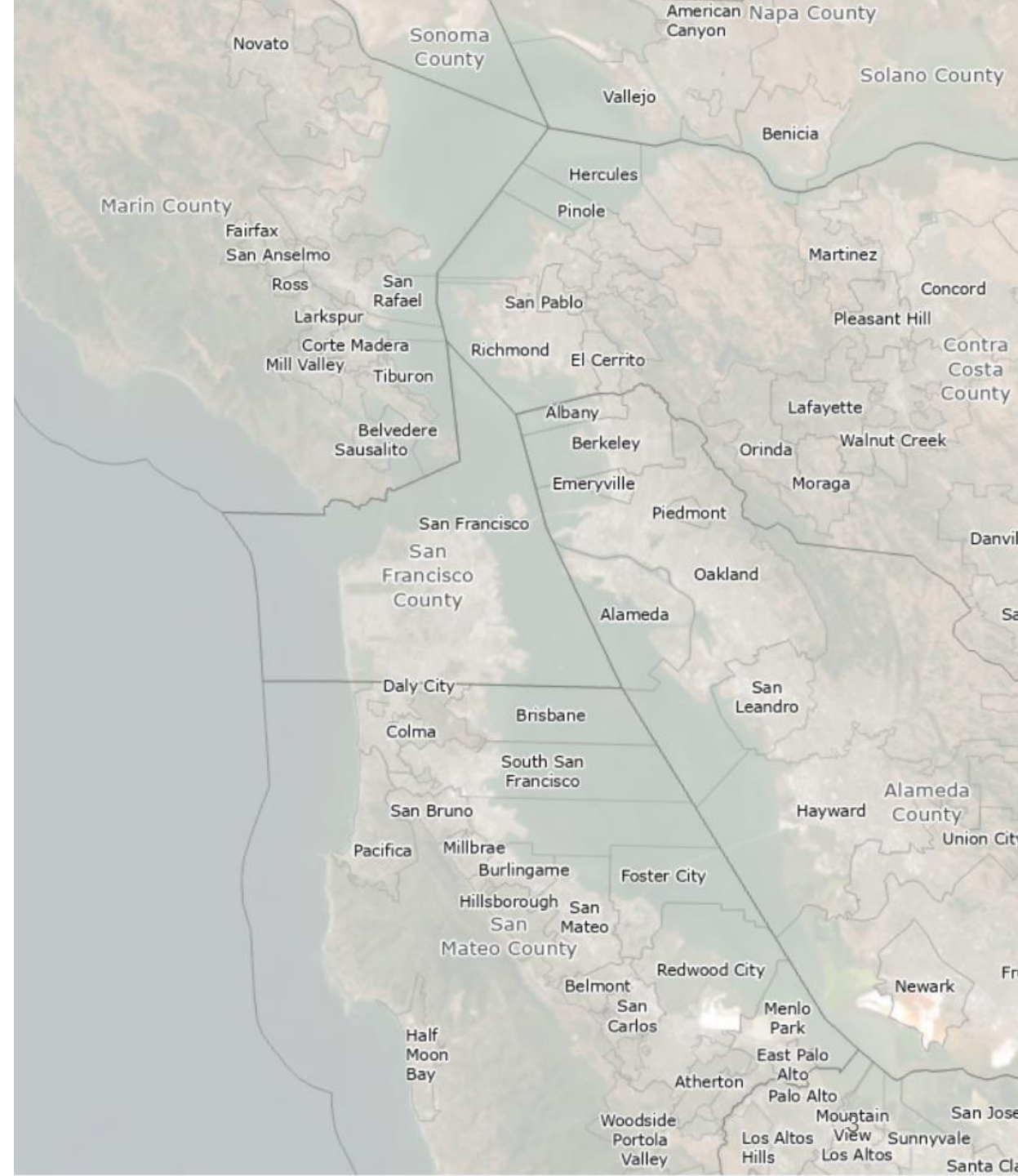
Automating Metadata Updates

---

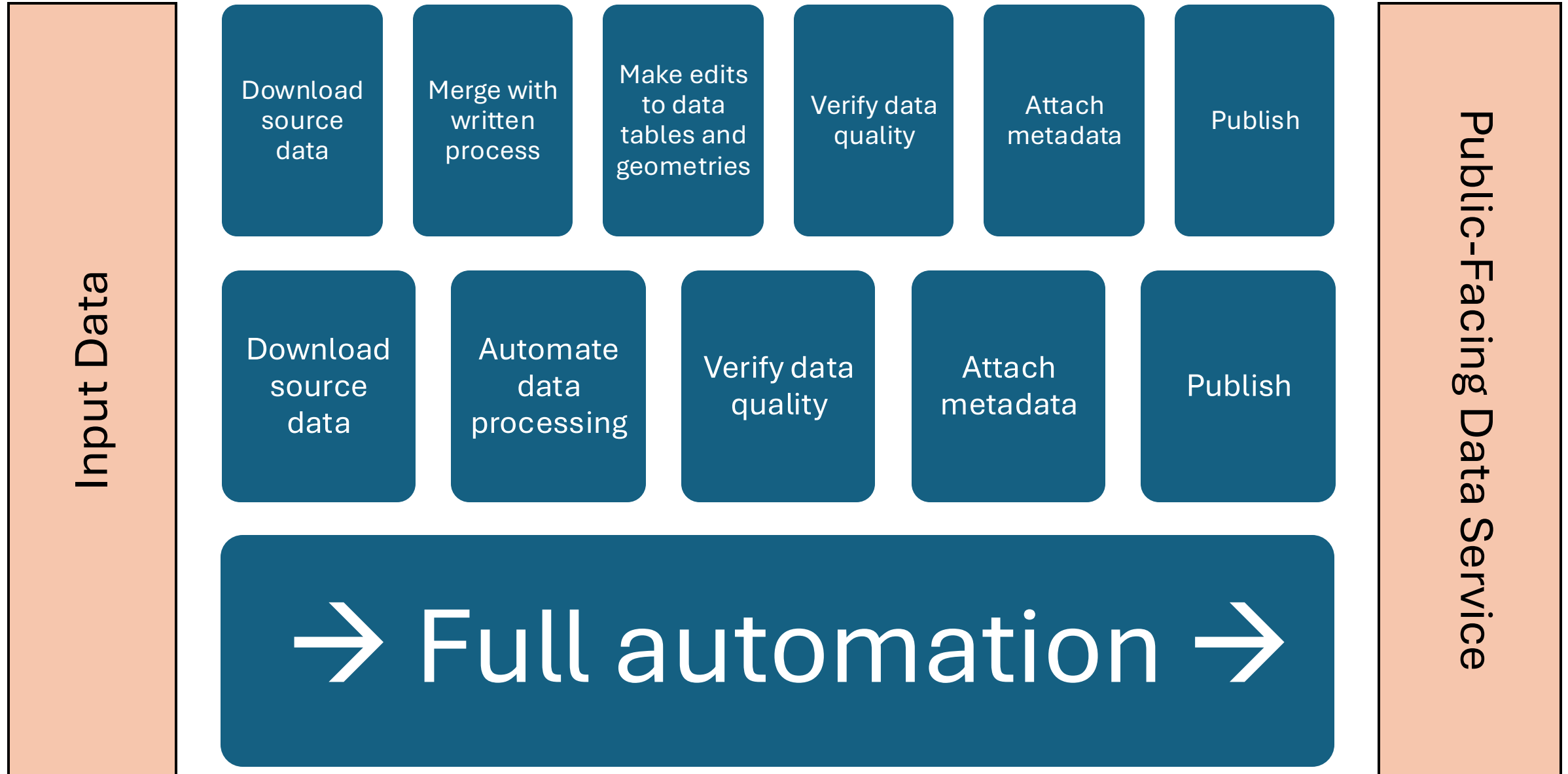
Highly Reliable Blue-Green Data  
Services

# City/County Data Layer

- CDTFA publishes an authoritative layer for internal use managed by Anna Price (thank you!).
- Produce a data service fusing CDTFA's city and county boundary data with identifiers from other agencies to make the data broadly applicable
  - GNIS, Census, and CalTrans IDs
  - Coastline cut
  - Room for other nice-to-haves and common needs like areas.
- Updated whenever the source datasets are updated



# Example Pipelines





Task

Download  
source data

Automate data  
processing

Verify data  
quality

Attach  
metadata

Publish

Public-Facing  
ArcGIS Feature  
Service (View)

Tool

GitHub + ArcGIS Online Notebooks

arcpy\_metadata

Blue-green data  
services

# The Task

- **We need to:**

- Access ESRI's Python libraries and APIs
- On-demand and when-scheduled
- In code that reads and publishes internal data services.

- **Common solutions:**

- Run code on ArcGIS Server
- Maintain an ArcGIS Pro virtual machine for scripts

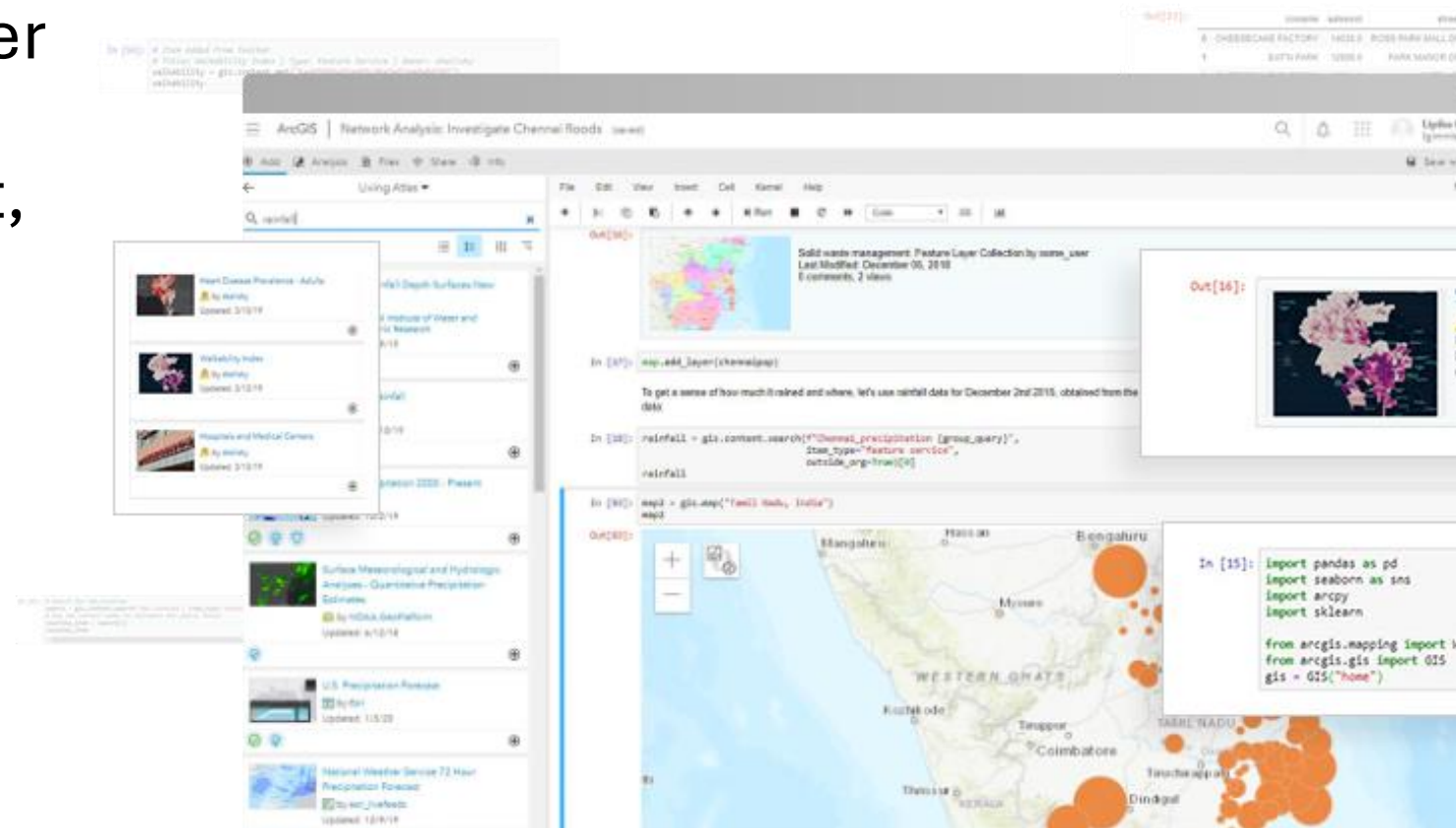
- **Challenges:**

- Expensive to license and run the VMs
- Maintaining logins for the ArcGIS Pro VM is non-trivial



# ArcGIS Online Notebooks

- Includes a standard Jupyter notebook interface for Python code development, preloaded with:
  - ArcGIS API for Python
  - Optionally arcpy
  - Many other standard packages – published in different runtime versions.
  - Ability to install your own packages, write your own code, access uploaded resources



# Why notebooks?

- **Low overhead** – can be scheduled, cost almost nothing when not running and cost cents per hour to run, even for analytical workloads.
- **ArcGIS environment is already ready to go** – no additional administration costs
- **Already logged into your organization** – no credential management when retrieving or publishing layers to AGOL
- **Can be debugged interactively**
- **Can be scheduled to run on their own**

# Other usage scenarios for notebooks

- On-demand geoprocessing tools with parameters
- Data analyst notebooks for exploratory analysis, modeling, and more – similar to standard ArcGIS notebooks, but nothing to install.
- Manually run repeated actions in your ArcGIS Online organization
- **Limitations**
  - By default, only organization administrators can create/edit/run notebooks, but this can be assigned to custom roles.
  - All notebooks for a user share a filesystem – isolate carefully.

# ArcGIS Online Notebooks - Credits

## Storage

- Storage cost for the notebook of 1.2 credits/1GB/month, but file size is *\*very\** small.
- Data stored for use *within* the notebook billed at 12 credits/1GB/month.

## Execution

<p>ArcGIS Notebooks: Interactive</p>	<p>Create, open, and run notebooks</p>	<p>3 credits per hour, per notebook using the Advanced runtime, calculated per minute (10-minute minimum)</p> <p>30 credits per hour, per notebook using the Advanced with GPU runtime, calculated per minute (10-minute minimum)</p>
<p>ArcGIS Notebooks: Automated Workflows</p>	<p>Run a <a href="#">scheduled notebook task</a></p> <p>Run a <a href="#">notebook web tool</a></p>	<p>1.5 credits per hour, per run using the Standard runtime, calculated per minute</p> <p>4.5 credits per hour, per run using the Advanced runtime, calculated per minute</p> <p>45 credits per hour, per run using the Advanced with GPU runtime, calculated per minute</p>

# Comments during CoP Presentation

- During the live presentation, members noted the following about ArcGIS Online notebooks:
  - Limitation of 5 scheduled notebooks – have one notebook kick off multiple processes if need more.
  - Multiple people expressed some frustrations with getting notebooks to behave reliably
  - Some moved back to running processes on ArcGIS Enterprise

Task

Download  
source data

Automate data  
processing

Verify data  
quality

Attach  
metadata

Publish

Public-Facing  
ArcGIS Feature  
Service (View)

Tool

GitHub + ArcGIS Online Notebooks

arcpy\_metadata

Blue-green data  
services



# Why a Python package?

*(i.e. why not put the code in the notebook)*

- We can move core pieces of code out for modularity and testing
- Use any/all other development and testing workflows that work with the team, then bring it into the ArcGIS Online Notebook only when it's time to run the full pipeline
- Azure DevOps can publish private packages we pull into the notebook if the code passes tests.



Task

Download  
source data

Automate data  
processing

Verify data  
quality

Attach  
metadata

Publish

Public-Facing  
ArcGIS Feature  
Service (View)

Tool

GitHub + ArcGIS Online Notebooks

arcpy\_metadata

Blue-green data  
services

# Metadata

- At the end of a data pipeline, we need to code to change the metadata based on the data run:
  - What were the input datasets, URLs, and versions?
  - When did the code run? With what version of the code?
- **Challenge:**
  - ArcGIS Pro provides scripting access to *basic* metadata for automated processes via Python – title, description, etc. Good for many uses
  - Little access to structured ISO 19139 (FGDC) metadata
  - Different Python libraries/interfaces for local data (`arcpy`) versus online services (`arcgis`)

# Metadata – the solution

- Revived and enhanced a Python library we'd built in the ArcMap era to provide access to read and edit metadata.
- ArcGIS Pro initially lacked the mechanisms – now has them
- Unified API for local and online services
- Access to 34 structured metadata elements (plus more sub-elements) – extensible to most others.
- Useful for updating elements like Last Update, Points of Contact, and code versions used in producing the dataset.

# Example

```
import arcpy_metadata as md
metadata = md.MetadataEditor(path_to_some_feature_class)

from datetime import date
today = date.today()
metadata.last_update = today

metadata.point_of_contact.contact_name = "Nick Santos"
metadata.point_of_contact.email = nick.santos@state.ca.gov

metadata.finish() # save it all back to the source
```

Task

Download  
source data

Automate data  
processing

Verify data  
quality

Attach  
metadata

Publish

Public-Facing  
ArcGIS Feature  
Service (View)

Tool

GitHub + ArcGIS Online Notebooks

arcpy\_metadata

Blue-green data  
services

# Blue Green Data Services

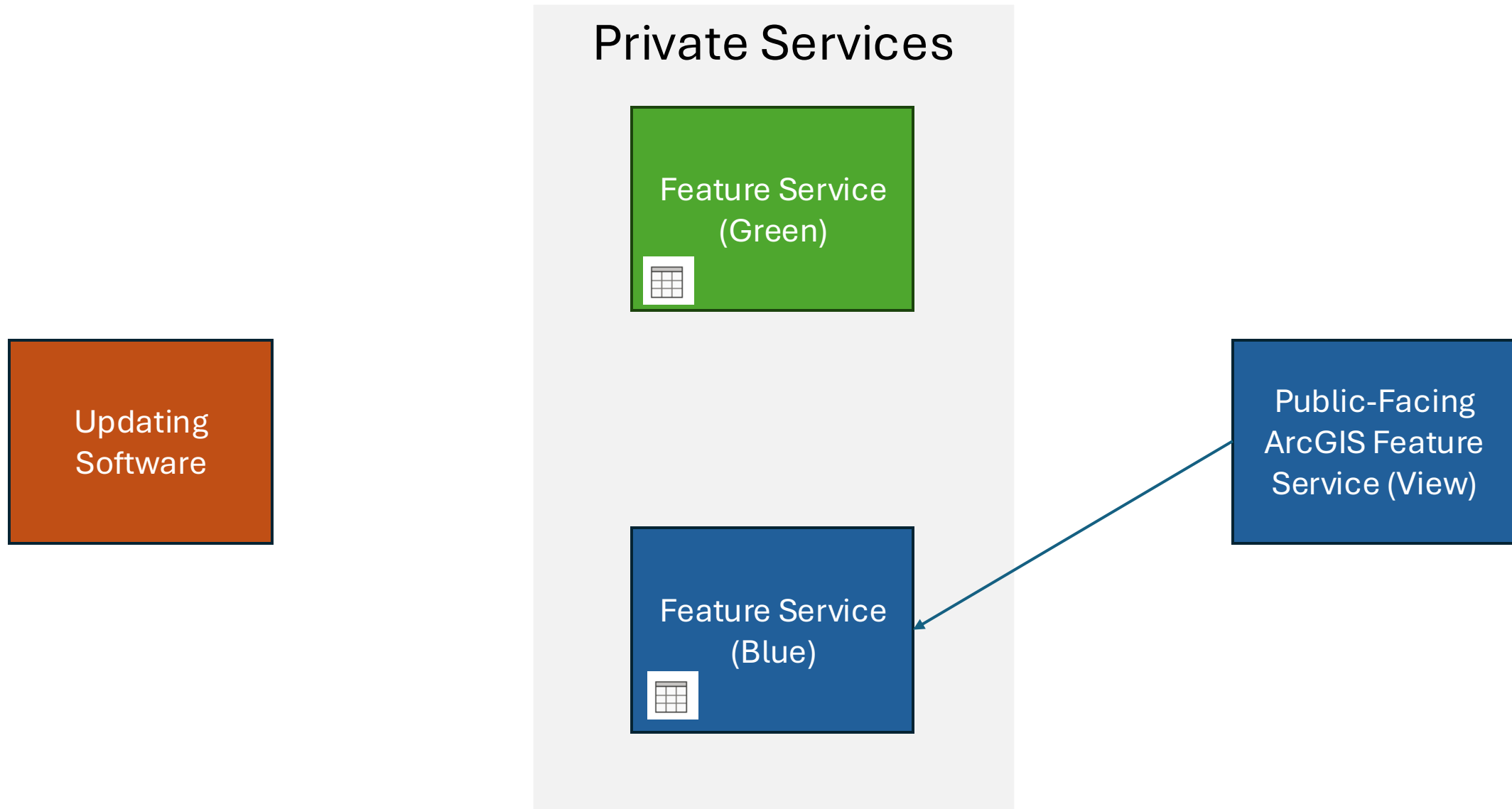
- Need a feature service that is highly reliable, but which can be safely updated
- **Challenge:**
  - Overwrite Web Layer modifies more than just data, could fail and leave the service in an unknown state
  - Upsert (Truncate/Append) workflows are simpler, but still leave the service without data temporarily and could also fail, leaving service in an unknown state needing recovery

# Blue-Green Data Services

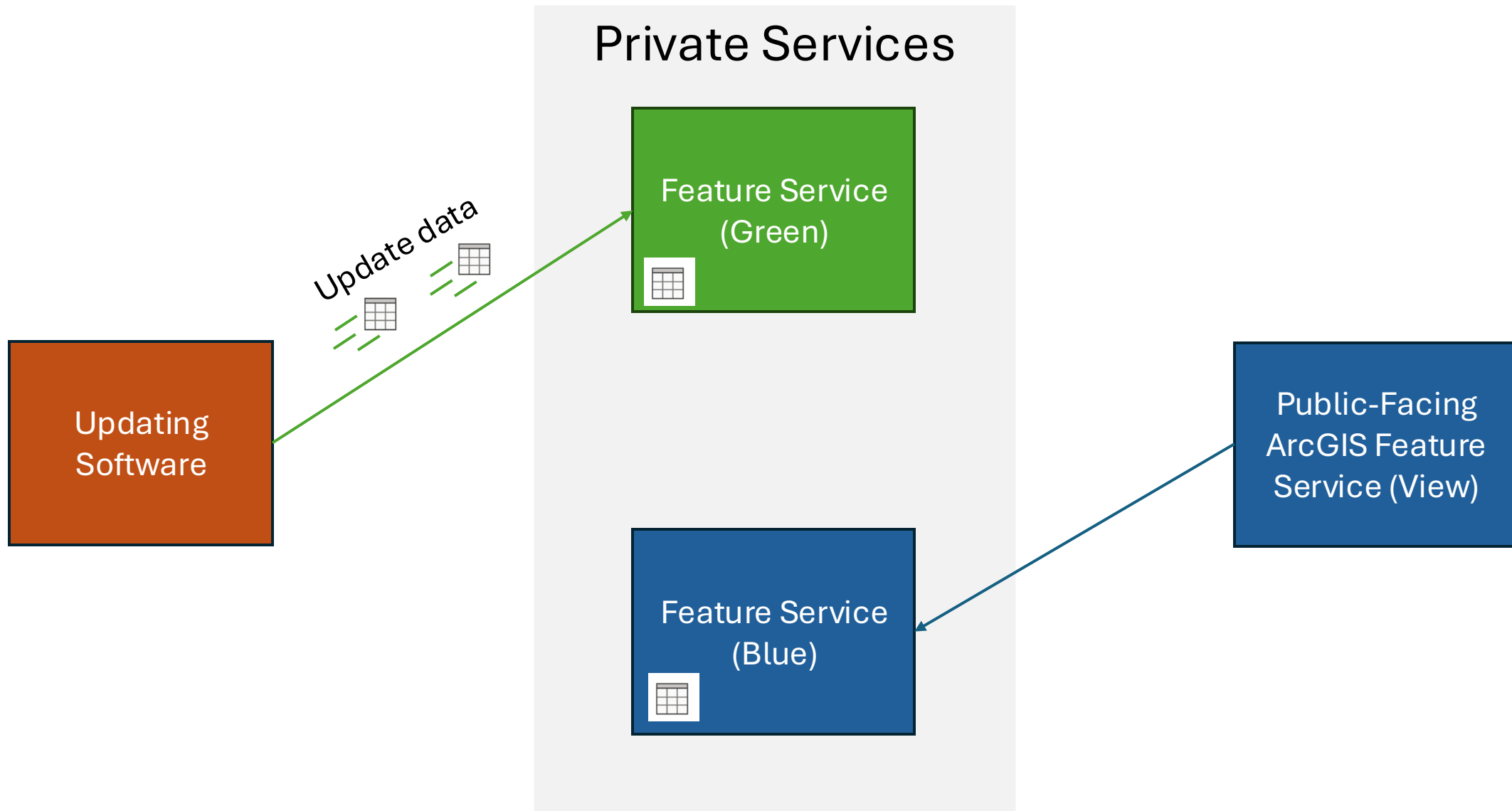
- The term comes from DevOps' Blue-Green Deployments
  - Two equal services – one is live for the public, the other for staging future changes. Swap them to deploy!
- Pair a public feature service view with two private feature services
- Only one of the private services is the data source for the view
- The other is available for changes and updates.
- Helpful for:
  - Seamless, atomic updates at a known, stable URL
  - Comparing between the previous and current versions
  - Reverting to the last known good version if an issue is identified



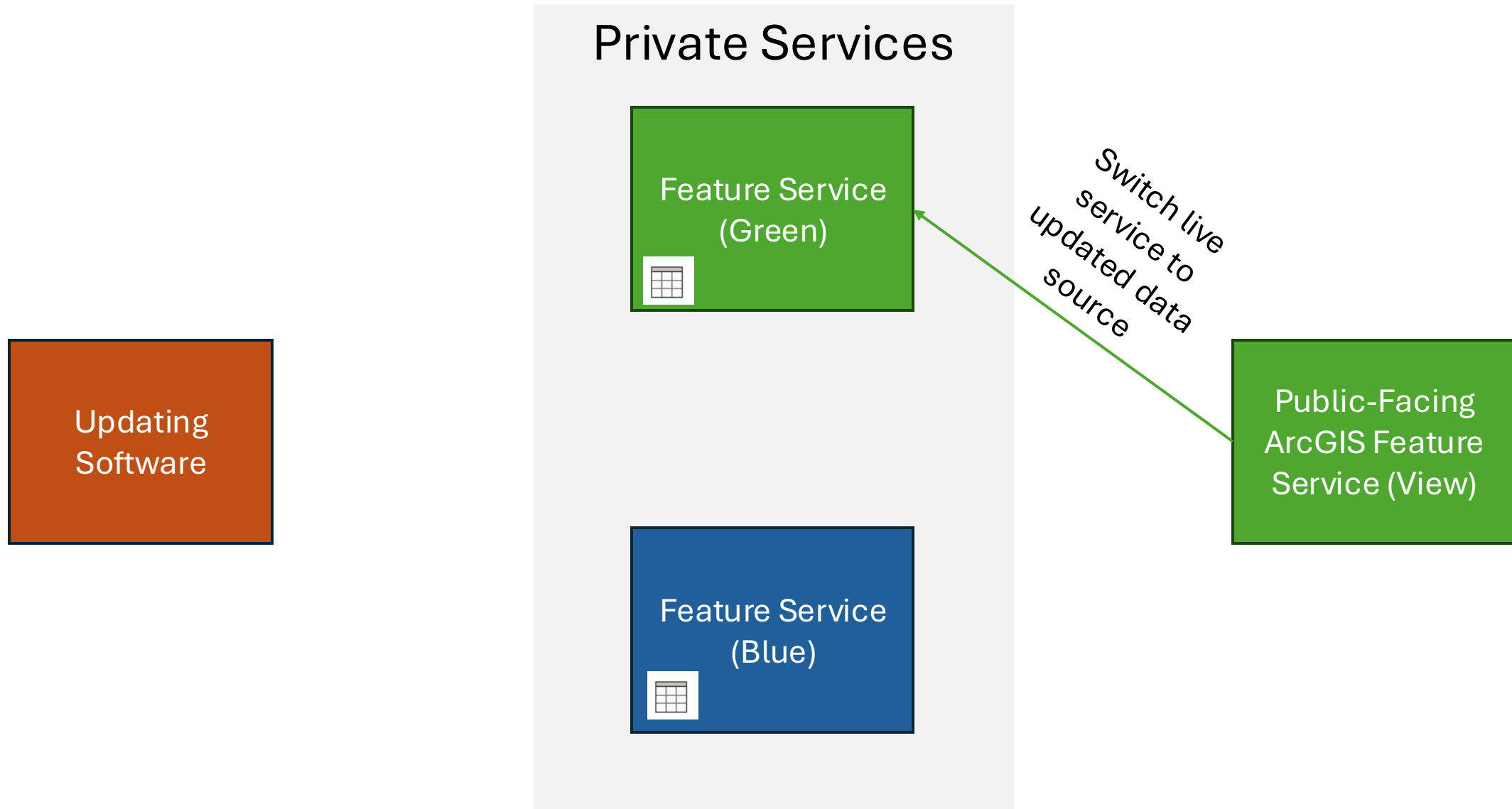
# Blue-Green Data Services



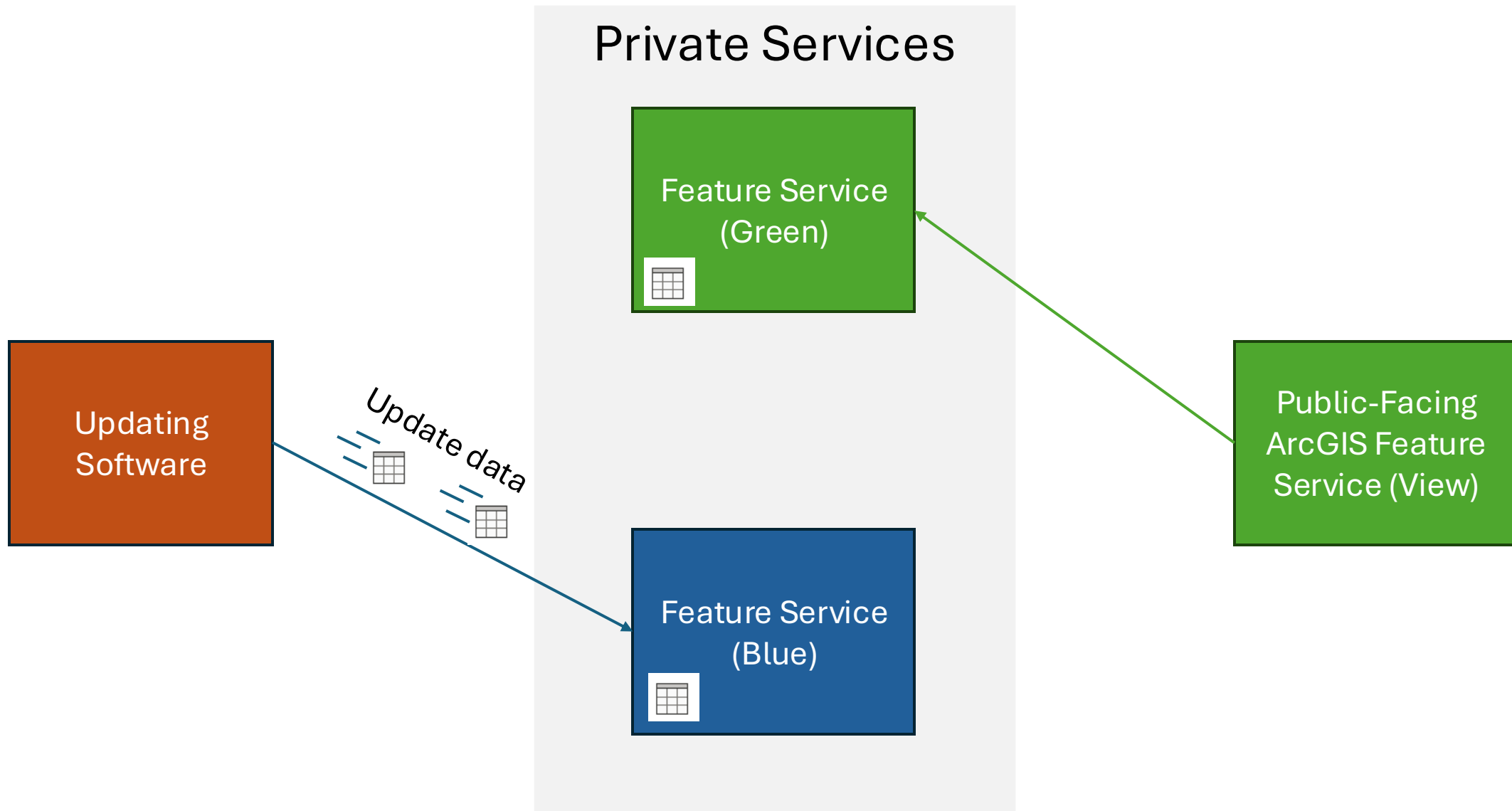
# Blue-Green Data Services



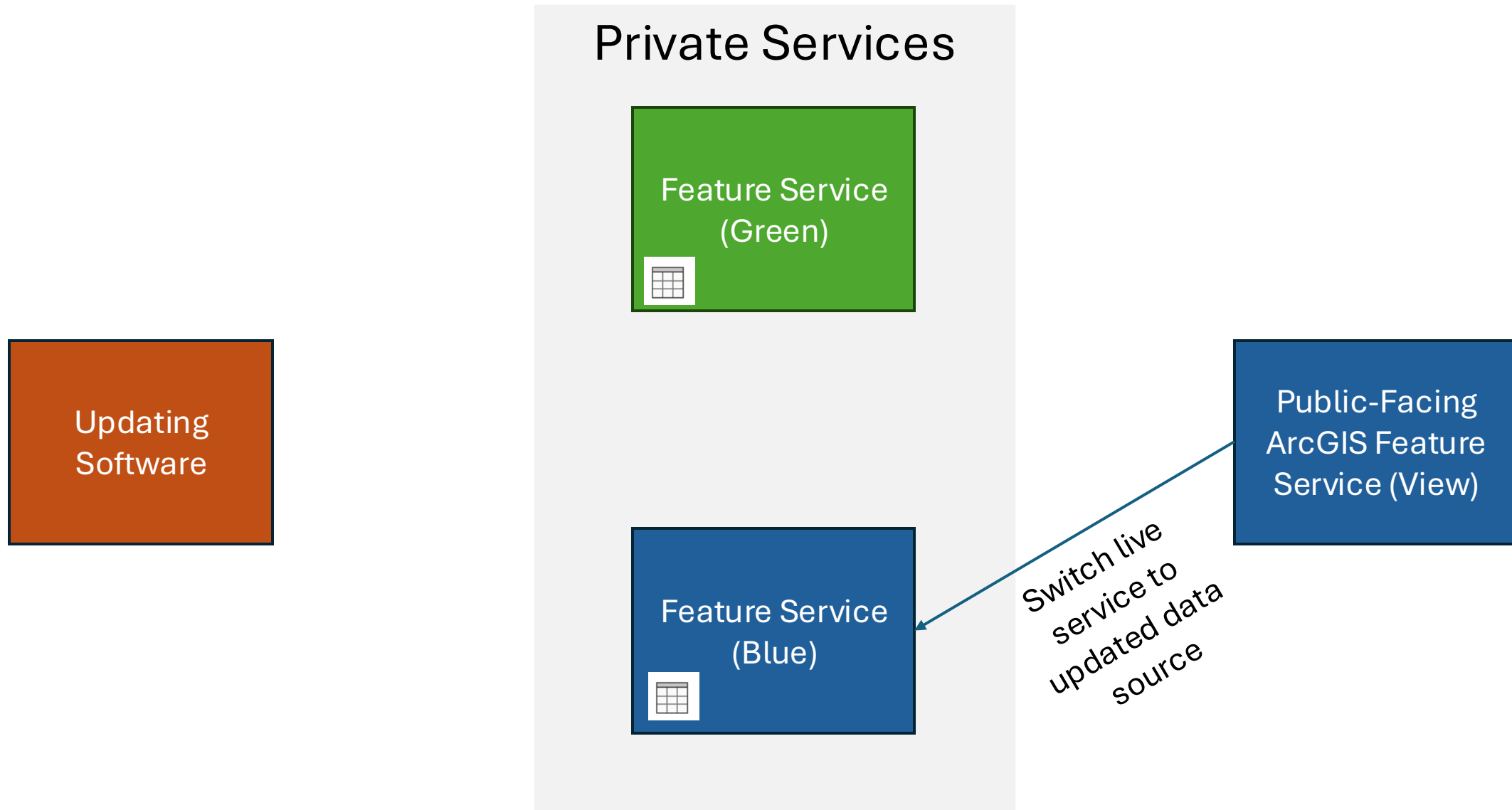
# Blue-Green Data Services



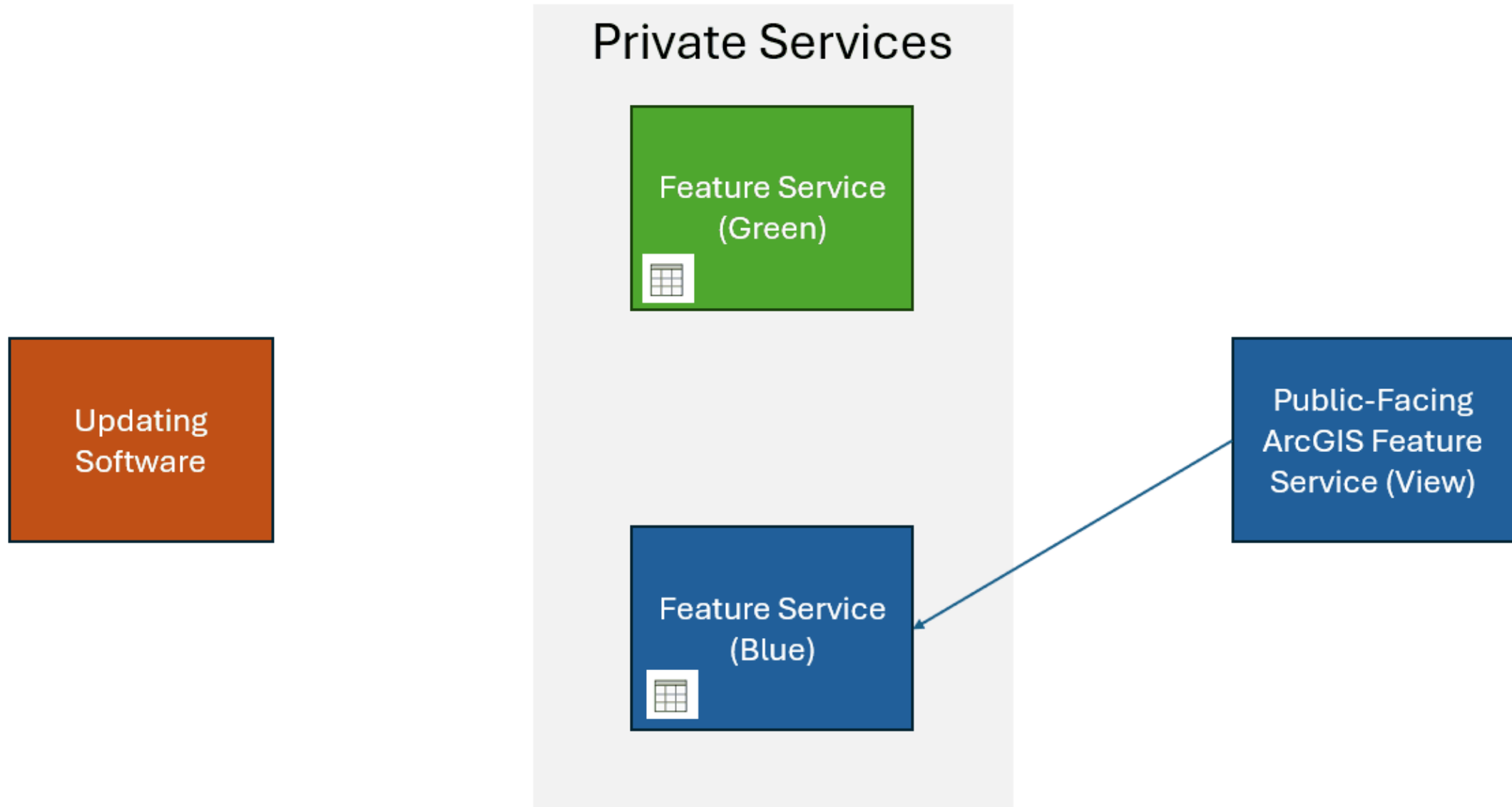
# Blue-Green Data Services



# Blue-Green Data Services



# Blue-Green Data Services



# The resulting services in ArcGIS Online

Two feature services and a view

Only the view is public

Title		Modified		
<input type="checkbox"/> California City Boundaries and Identifiers	Feature Layer (hosted, view)	Sep 16, 2024		Preview ...
<input type="checkbox"/> California City Boundaries and Identifiers Blue Version	Feature Layer (hosted)	Sep 16, 2024		Preview ...
<input type="checkbox"/> California City Boundaries and Identifiers Blue Version	Service Definition	Sep 16, 2024		Preview ...
<input type="checkbox"/> California City Boundaries and Identifiers Green Version	Feature Layer (hosted)	Sep 16, 2024		Preview ...
<input type="checkbox"/> California City Boundaries and Identifiers Green Version	Service Definition	Sep 16, 2024		Preview ...

# How?

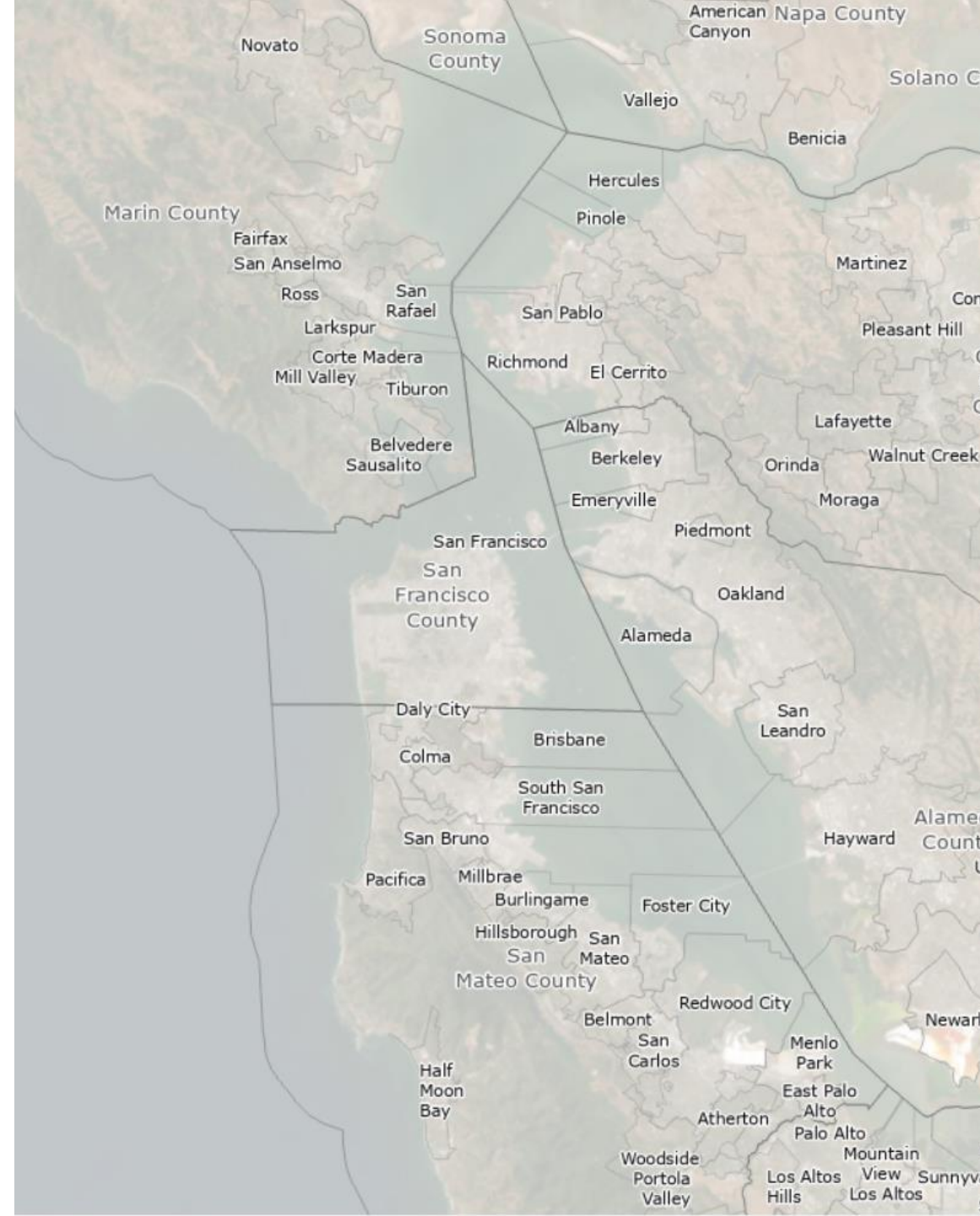
- Set up and configure initial services manually.
- Blue/Green must have same schema (among other requirements)
- This pattern can be used for manual/interactive updates with a human controlling publishing *or* controlled via code.
- We're building out a code library to support this pattern so it can be as simple as (in pseudocode):

```
import AGOLBlueGreen
service = AGOLBlueGreen(blue_id, green_id, view_id)
service.upsert(my_new_features) # empty staging service and insert new records
service.promote_staging() # - swap the services - demote live and promote staging
```



# City/County Status

- Coastline cut upcoming will change geometries and areas, add fourth coastal buffers service.
- Ideally keep the current schema – considering a few field name adjustments (please send us feedback)
- More metadata!
- May add more fields in the future – let us know what you need.



# City/County Links

- **City/County Feature Layers:**

- Original CDTFA Data: <https://gis.data.ca.gov/datasets/CDTFA::city-and-county-boundary-line-changes/>
- Counties with identifiers: <https://california.maps.arcgis.com/home/item.html?id=60b7e0f3d33b4064a4b43bf14589bfe3>
- Cities with identifiers: <https://california.maps.arcgis.com/home/item.html?id=8cd5d2038c5547ba911eea7bec48e28b>
- Overlapping cities and counties with identifiers: <https://california.maps.arcgis.com/home/item.html?id=894e9cda46bb45c2a0c7b5534b9a6b4a>

# Summary

- **ArcGIS Notebooks** are a cheap, simple-to-manage resource for data service automation
  - Optionally, use your existing development toolchains and deploy your packages into the notebooks
- **Blue-green data services** can help with data publishing and verification workflows while keeping high uptime
- **Metadata automation** with custom Python libraries unifies editing methods and adds ability to provide structured metadata

# Links

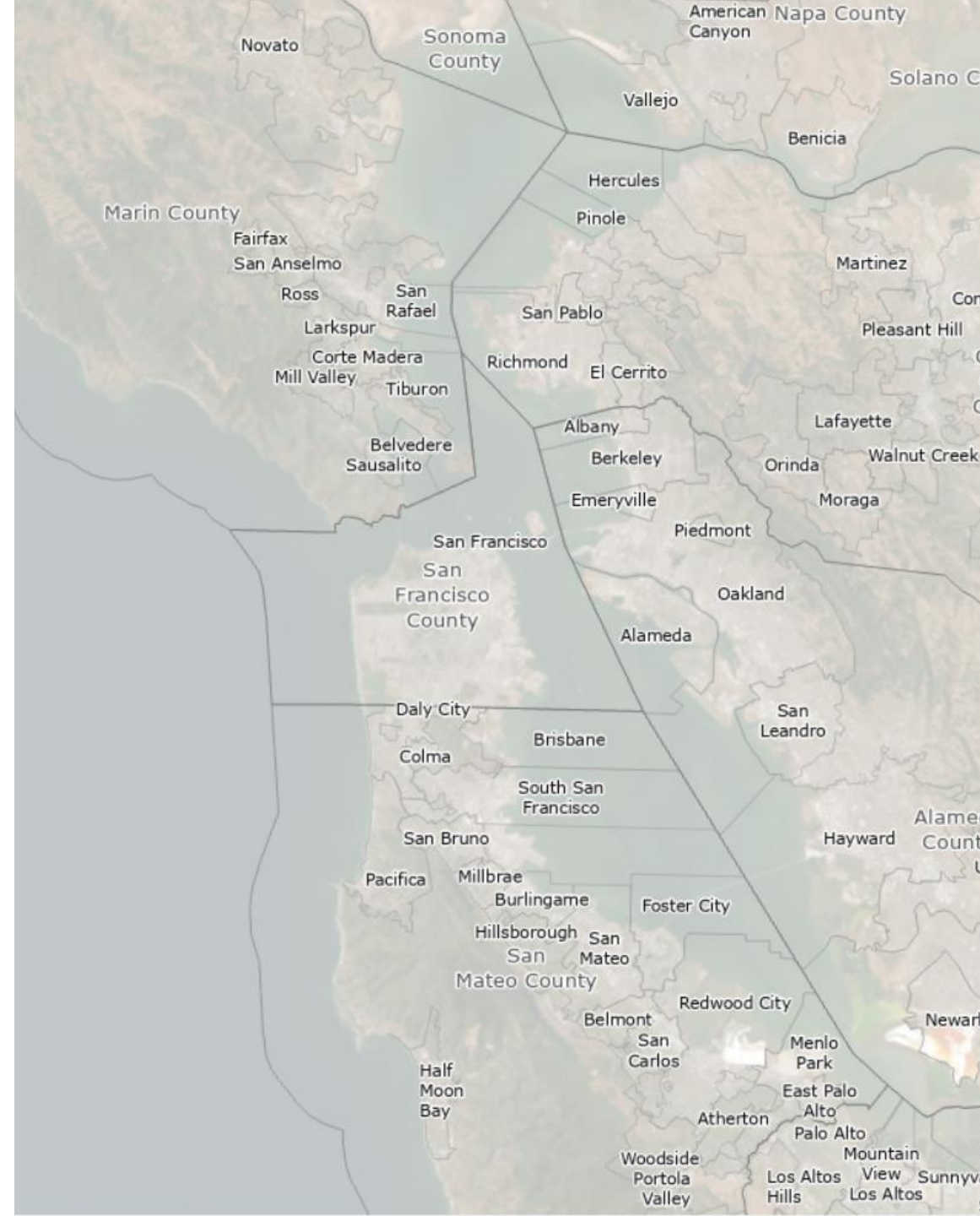
- **ArcGIS Notebooks:** <https://www.esri.com/en-us/arcgis/products/arcgis-notebooks/overview>
- **arcpy\_metadata:** [https://github.com/ucd-cws/arcpy\\_metadata/](https://github.com/ucd-cws/arcpy_metadata/) - still finalizing the update to ArcGIS Pro/Server 11
- **AGOLBlueGreen:** <https://github.com/Office-of-Digital-Services/cdt-ods-gis-agolbluegreen> – Pre-release, in development
- **CalEnterprise (Azure DevOps) access:** submit a DevSecOps request at <https://services.cdt.ca.gov/>
- Get in touch if you'd like to work together and make these tools useful for your team.

# What tooling are you using for automation?

- Please share in the chat!
- Contact us to set up your own presentation of your data pipelines at a future CoP.
  - New? Old? Perfect? Broken? Solves a problem? We'd love to start a conversation and hear how you're moving data around for your department's needs and constraints.

# Credits and Acknowledgements

- City/County Data:
  - CDT GIS team – current and former, including Liana Huang, Carmen Gonsalves, Mark Tyler, and Lothar Petrik
  - Anna Price at CDTFA for maintaining the authoritative dataset
  - Emma Lyons at CalTrans for workflow design with Liana
- arcpy\_metadata:
  - Thomas Maschler at the World Resources Institute coauthored the package



# Additional References

- Metadata:
  - Esri's Python package - arcpy.metadata: <https://pro.arcgis.com/en/pro-app/latest/arcpy/metadata/metadata-class.htm>
  - FGDC/ISO Metadata Standards: <https://www.fgdc.gov/metadata/geospatial-metadata-standards>
- ArcGIS Online Notebooks:
  - Overview: <https://www.esri.com/en-us/arcgis/products/arcgis-notebooks/overview>
  - Credit usage: <https://www.esri.com/en-us/arcgis/products/credits/usage#accordion-1-8>
- Blue-Green:
  - Blue-Green deployments: [https://en.wikipedia.org/wiki/Blue%E2%80%93green\\_deployment](https://en.wikipedia.org/wiki/Blue%E2%80%93green_deployment)
  - Feature Layer Views: <https://doc.arcgis.com/en/arcgis-online/manage-data/create-hosted-views.htm>
  - Esri article about swapping views: <https://www.esri.com/arcgis-blog/products/arcgis-online/data-management/swapping-layers-a-great-way-to-build-and-maintain-your-feature-layer/?srsltid=AfmBOoqXIZDffZfOrU-5GH9bREtc3ykoiXJUAcBkV6P30rmf4s1bpgO7>
  - Updating View Schema: <https://community.esri.com/t5/arcgis-online-blog/feature-layer-views-updating-parent-feature-layer/ba-p/1053668>

# Thank you

## Next GIS CoP Monthly Forum

### Wednesday, October 30<sup>th</sup> , 2024

Questions or comments send to: [gio@state.ca.gov](mailto:gio@state.ca.gov)