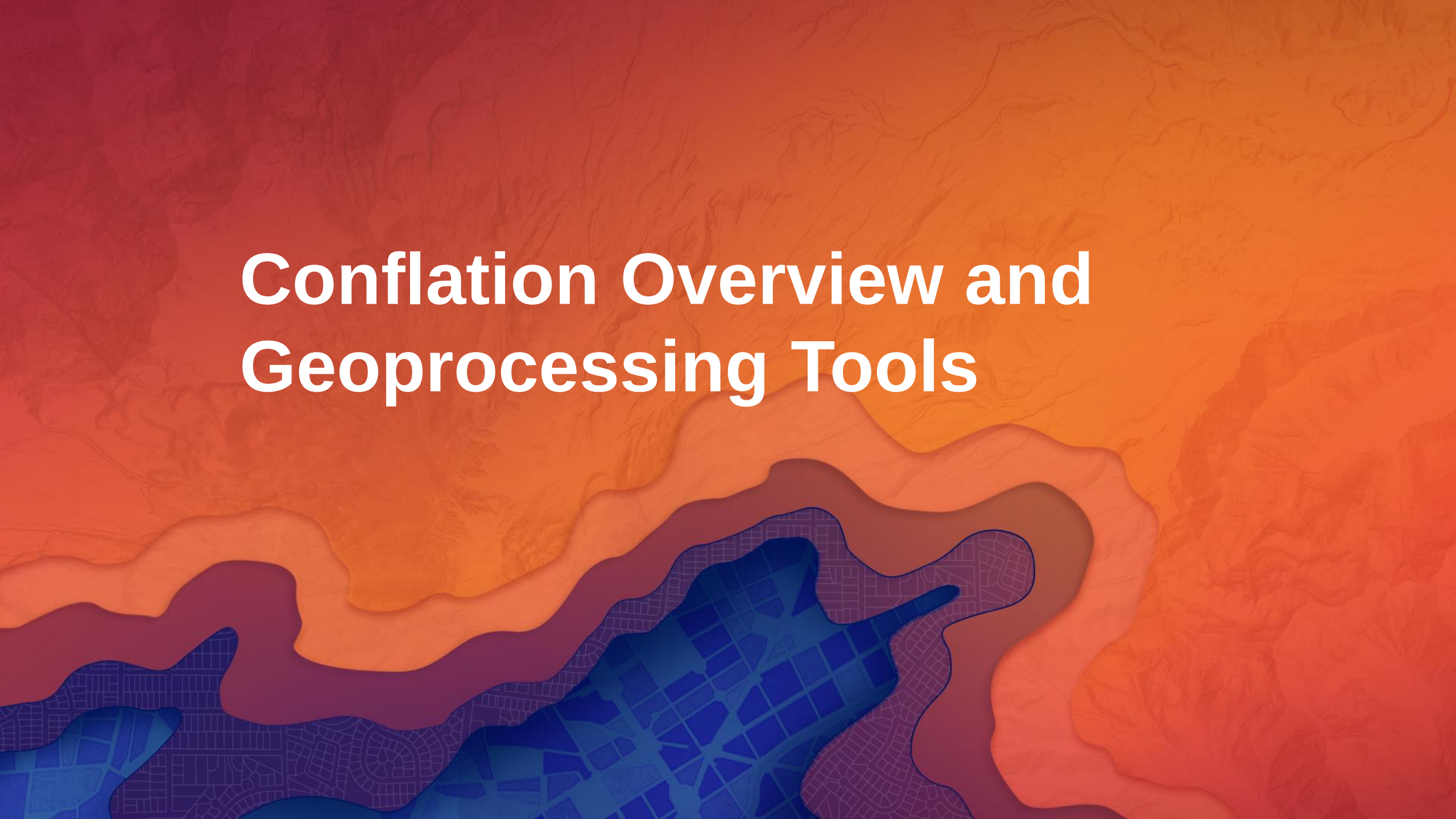
CONFLATION: EDGEMATCHING TOOLS AND WORKFLOWS

DAN LEE

Agenda

- Conflation Overview and Geoprocessing Tools
- Edge Matching Workflow
  - Demo
- Conclusions and Future Work

Conflation Overview and Geoprocessing Tools

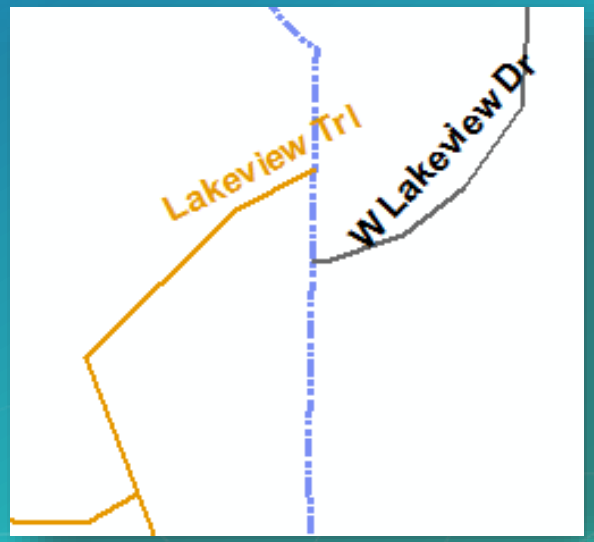


When using multi-source spatial data together

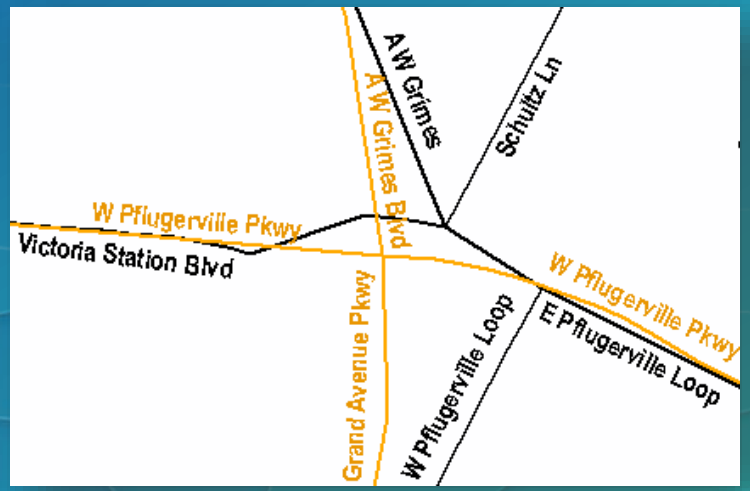
Common obstacles in analysis and mapping:

- Spatial and attribute inconsistency caused by differences in data collection and modeling
- High cost to fix the problems

Adjacent datasets



Overlapping datasets



Conflation reconciles multi-source datasets and optimizes data quality and usability

Between overlapping datasets:

- Detect feature changes (differences) through feature matching
- Make spatial adjustment and attribute transfer

Between adjacent datasets:

- Detect and resolve feature conflicts and disconnections through edge matching and alignment

Ultimately:

- Maintain an unified and seamless dataset – enriched and up-to-date
- No longer live with various imperfect datasets
- Rely on the data to perform analysis and quality mapping with confidence

What does it take to achieve the goals?

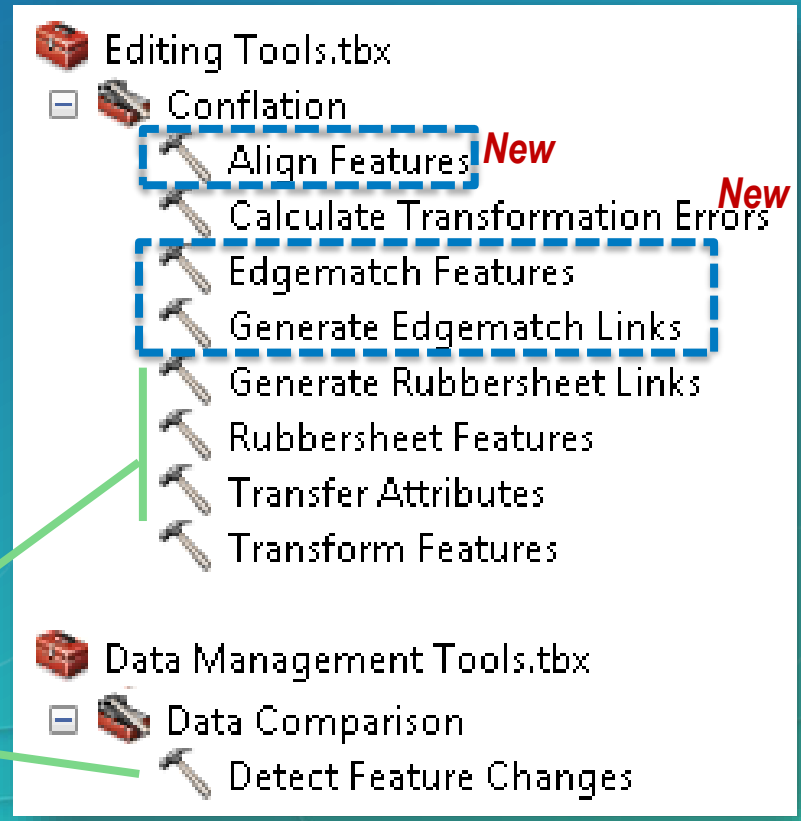
Our initial focuses

Have you used these tools in ArcMap?

Develop highly automated tools in Geoprocessing framework

- Starting with linear features (roads, parcel lines, rivers, etc.)
- Aiming at high accuracy (not promising 100%)
- Providing information to facilitate post-processing

In ArcGIS 10.5.1 and Pro 2.0



Build workflows

TW session (01:30pm – 02:45pm,Thursday, Room 31A) Conflation Tools and Workflows: An Introduction

Edge matching (EM) tools for adjacent datasets

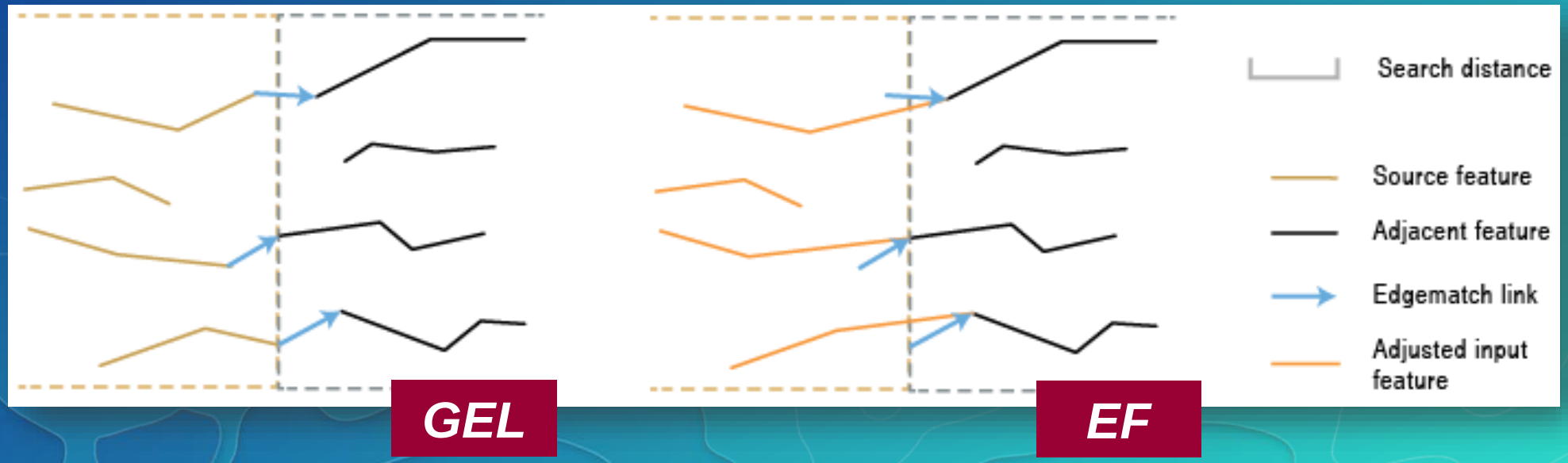
Based on proximity, topology, and continuity analysis, as well as attributes information

Generate Edgematch Links (GEL)

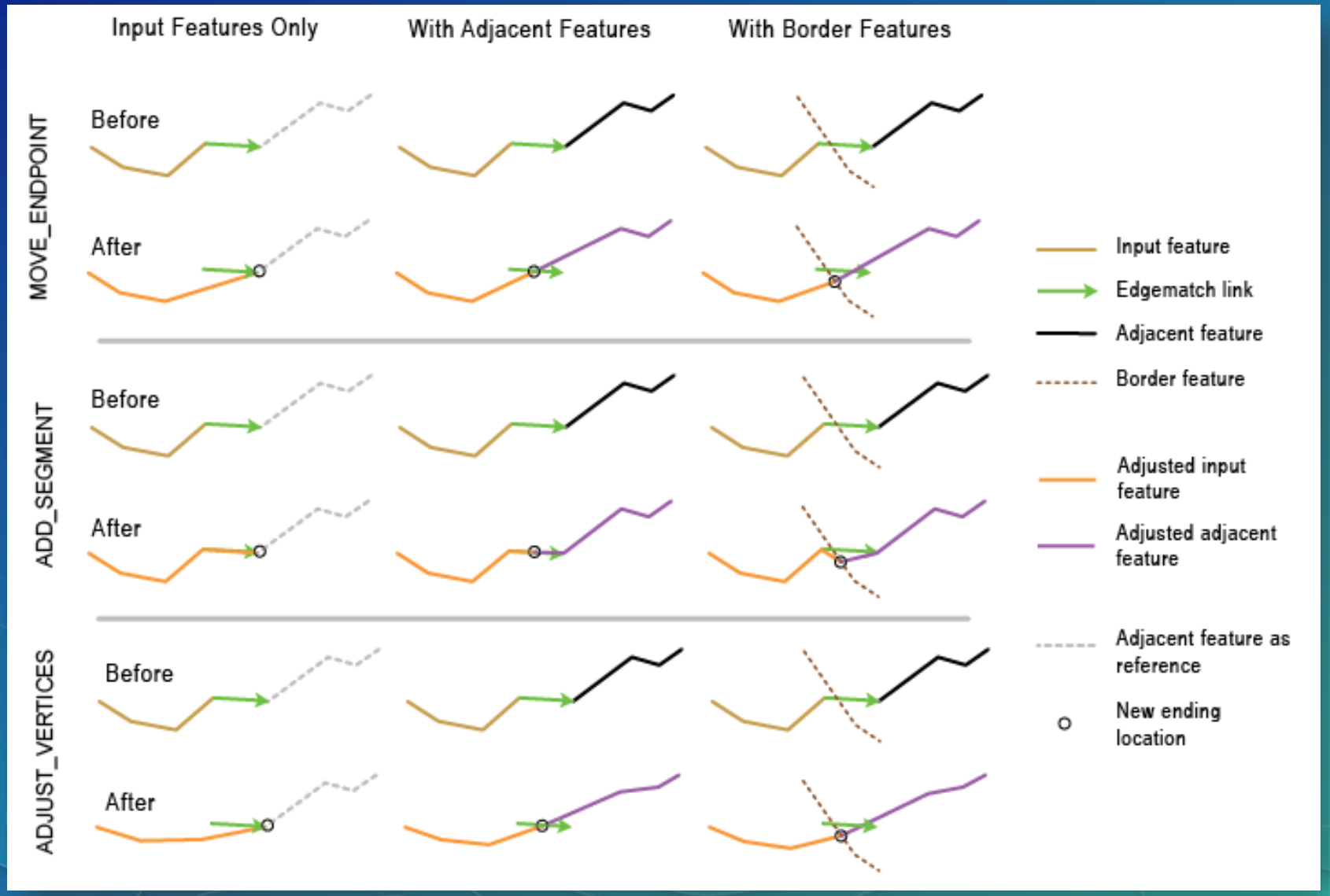
- From source features to adjacent features

Followed by Edgematch Features (EF)

- Connects features guided by the established links

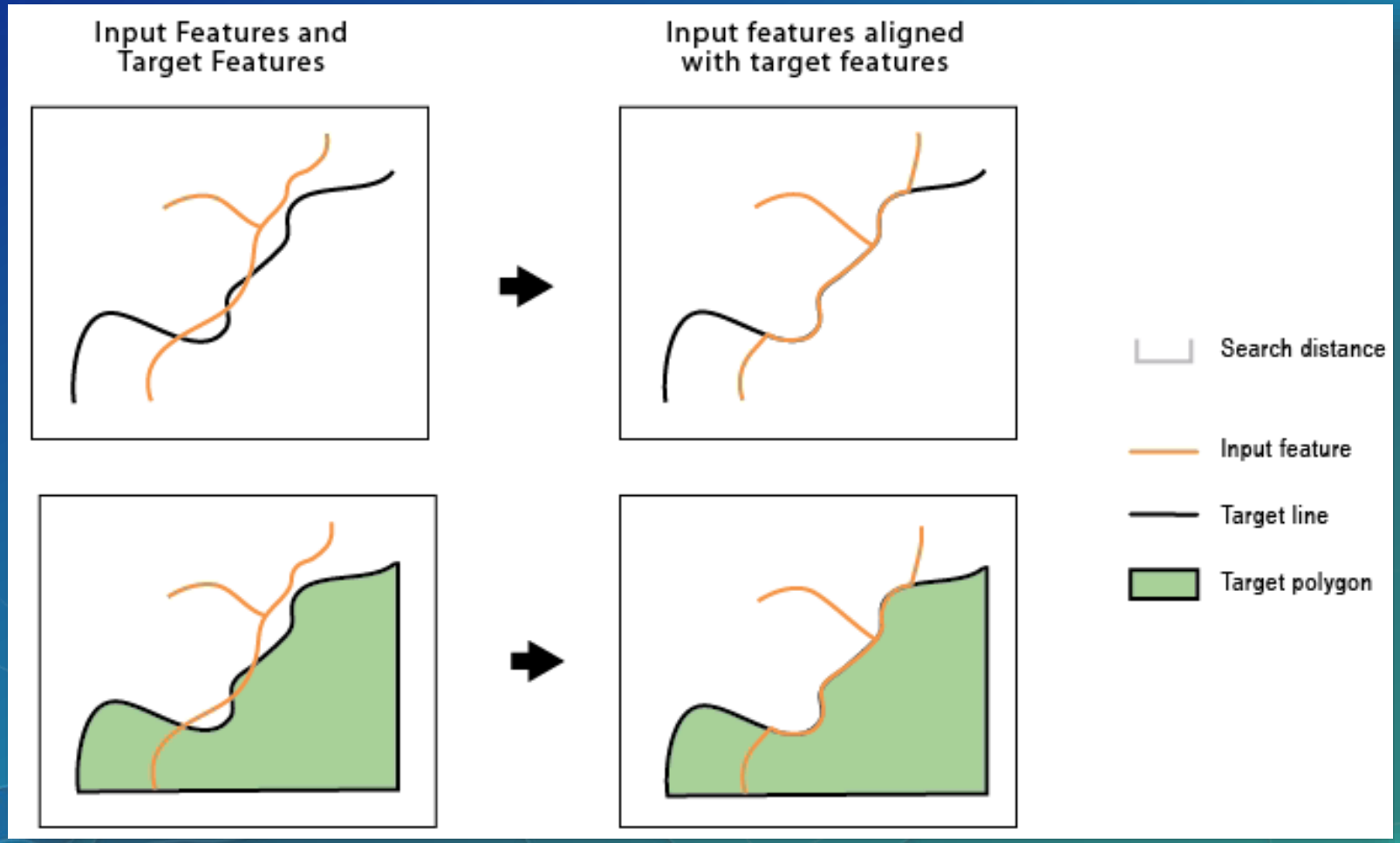


Options for connecting features

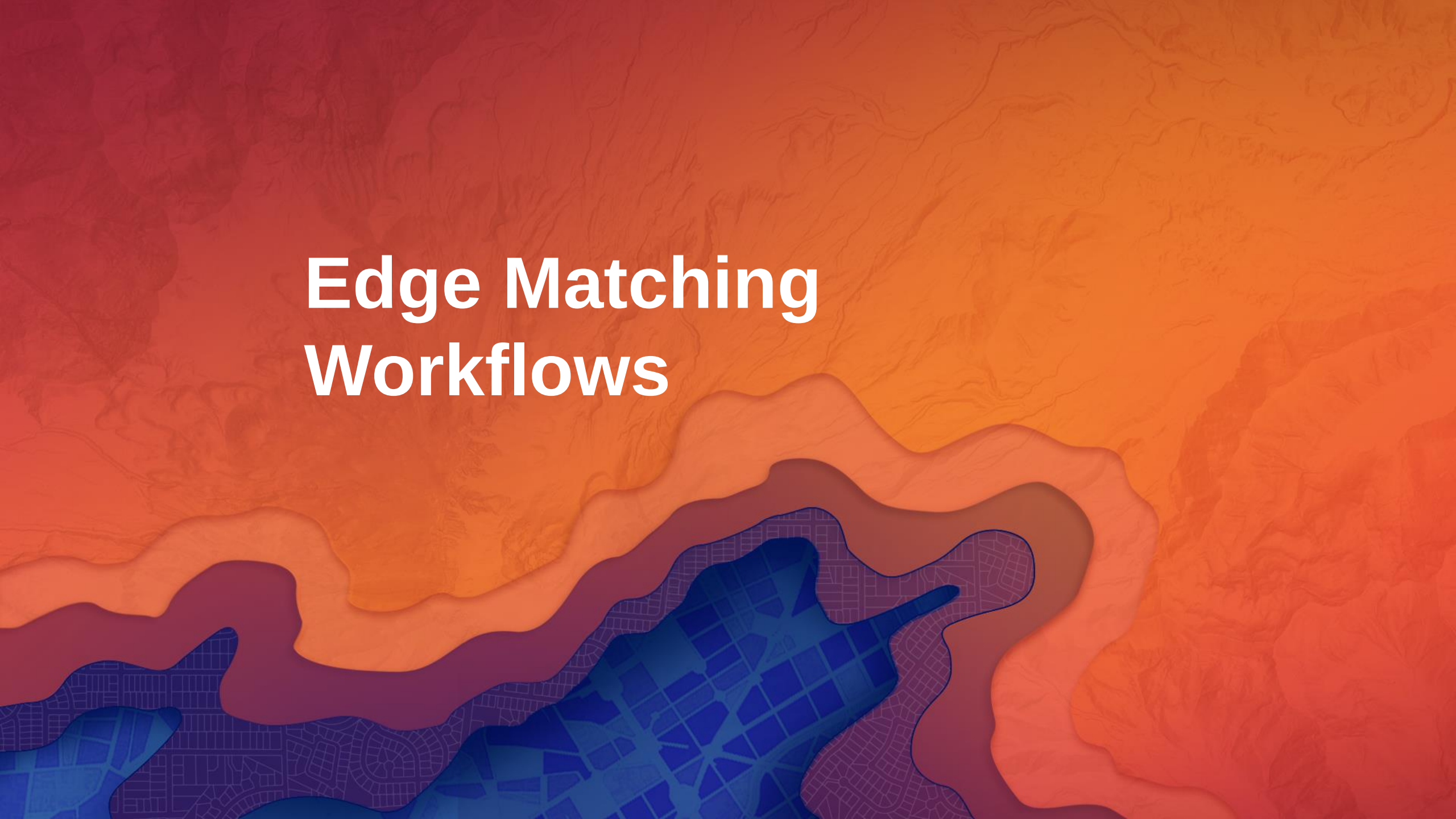


Align Features

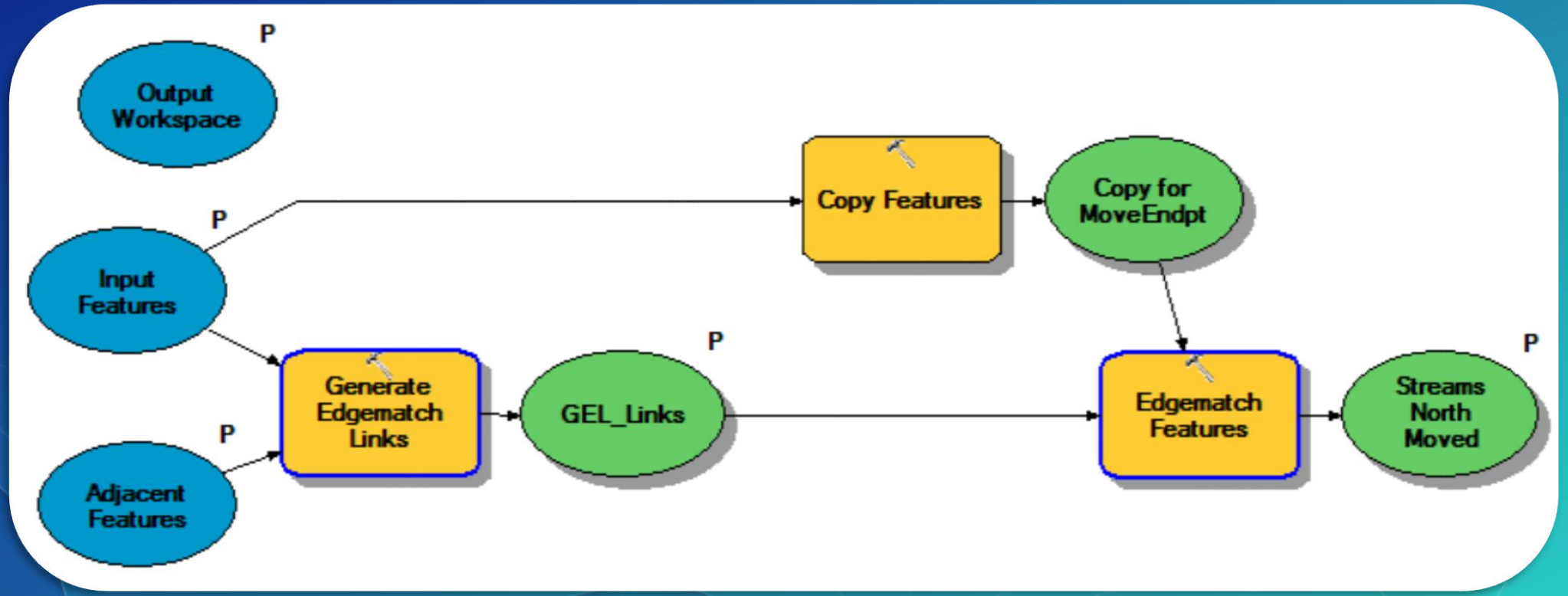
Based on proximity, topology, and similarity analysis, as well as attributes information



Edge Matching Workflows



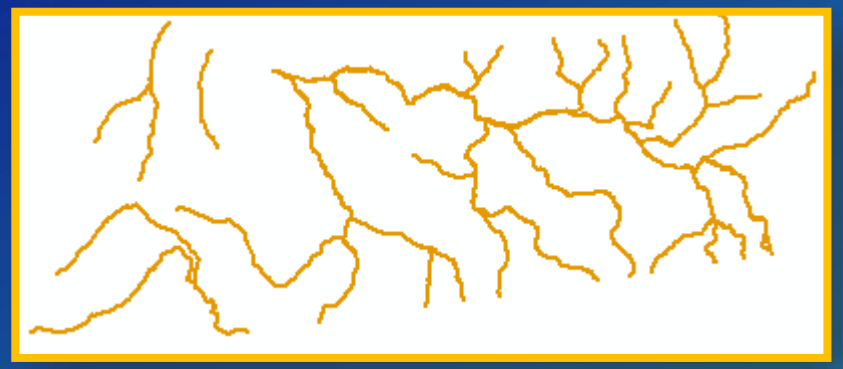
Conceptual workflow



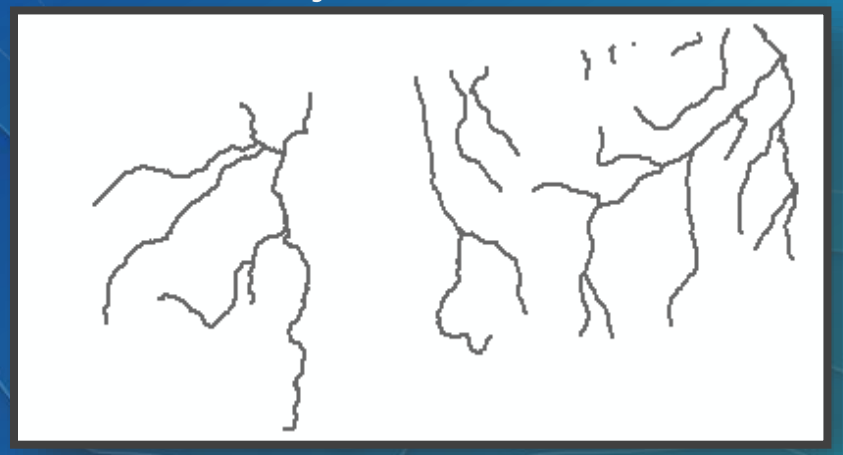
Example edgematching of adjacent datasets

Goal - make two adjacent datasets properly connect

Source features

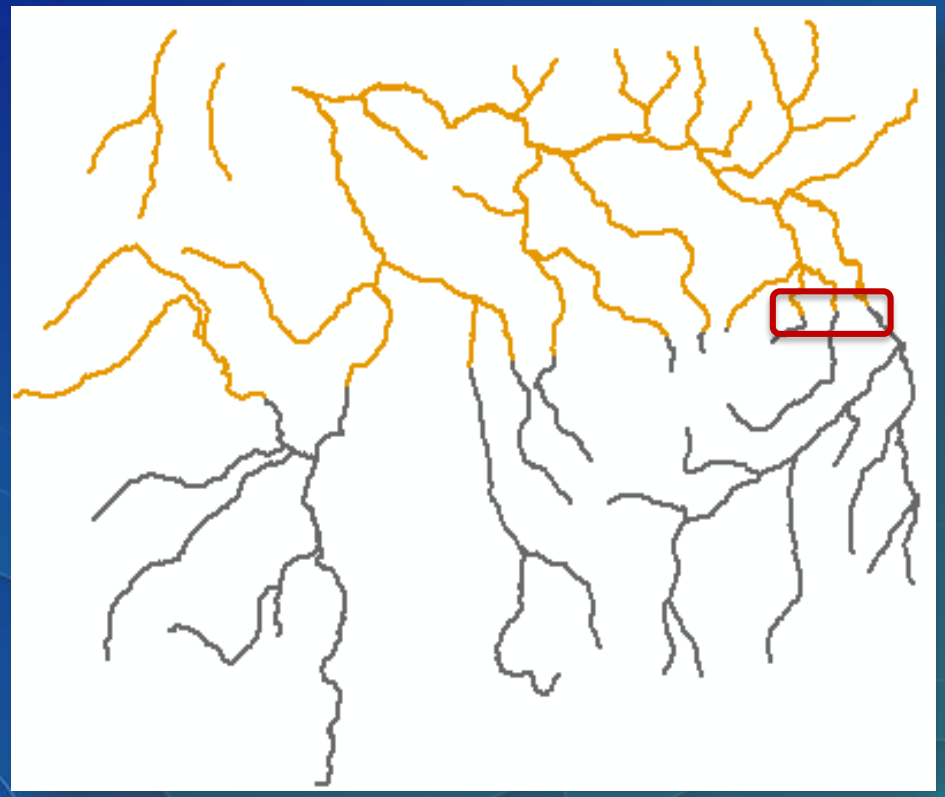


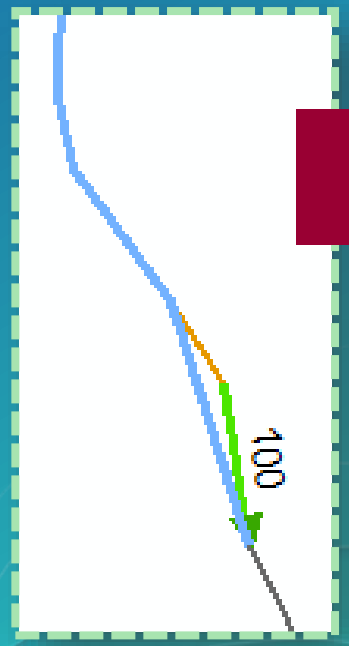
Adjacent features



Together

Results





Move endpoint

Conflation workflow in real world scenarios

- In same projection
- Data validation
- Selection of relevant features

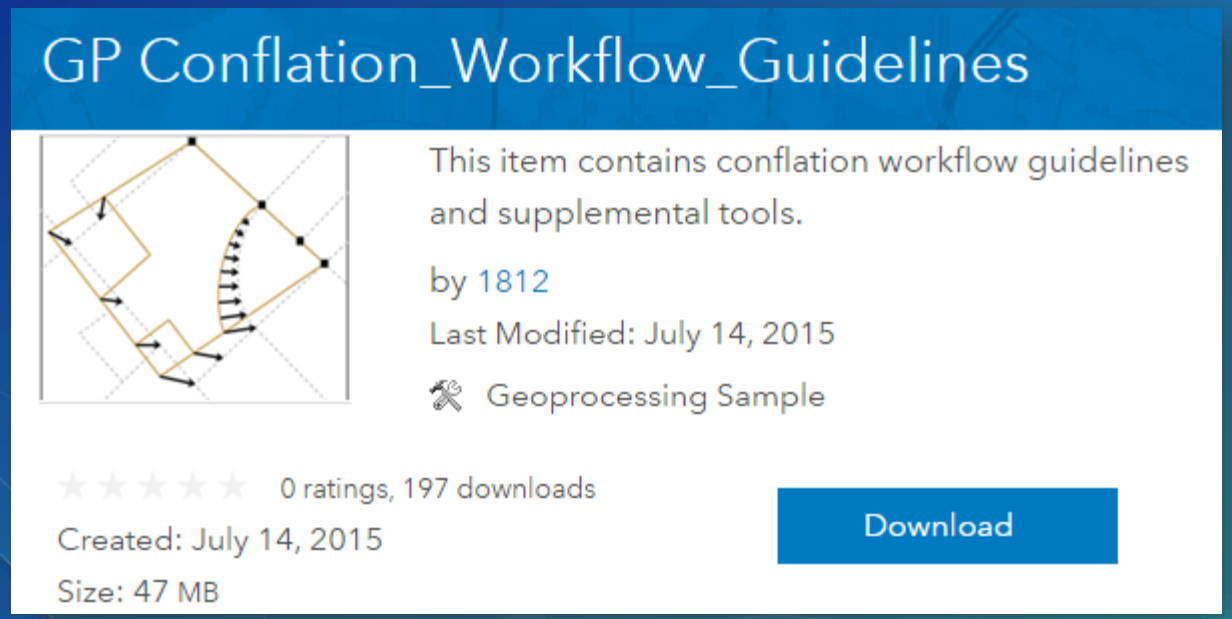
- Conflation tools
- Workflow tools

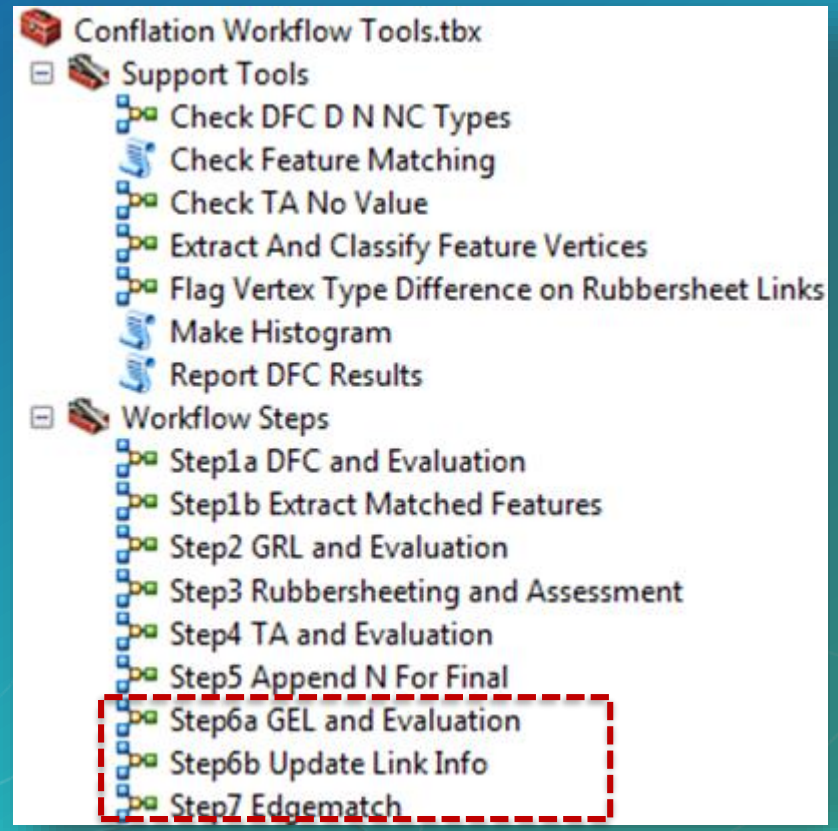
- Queued review
- Interactive editing

Supplemental tools and guidelines for download

http://www.arcgis.com/home/item.html?id=36961cde1b074f1f944758f6abec87cc

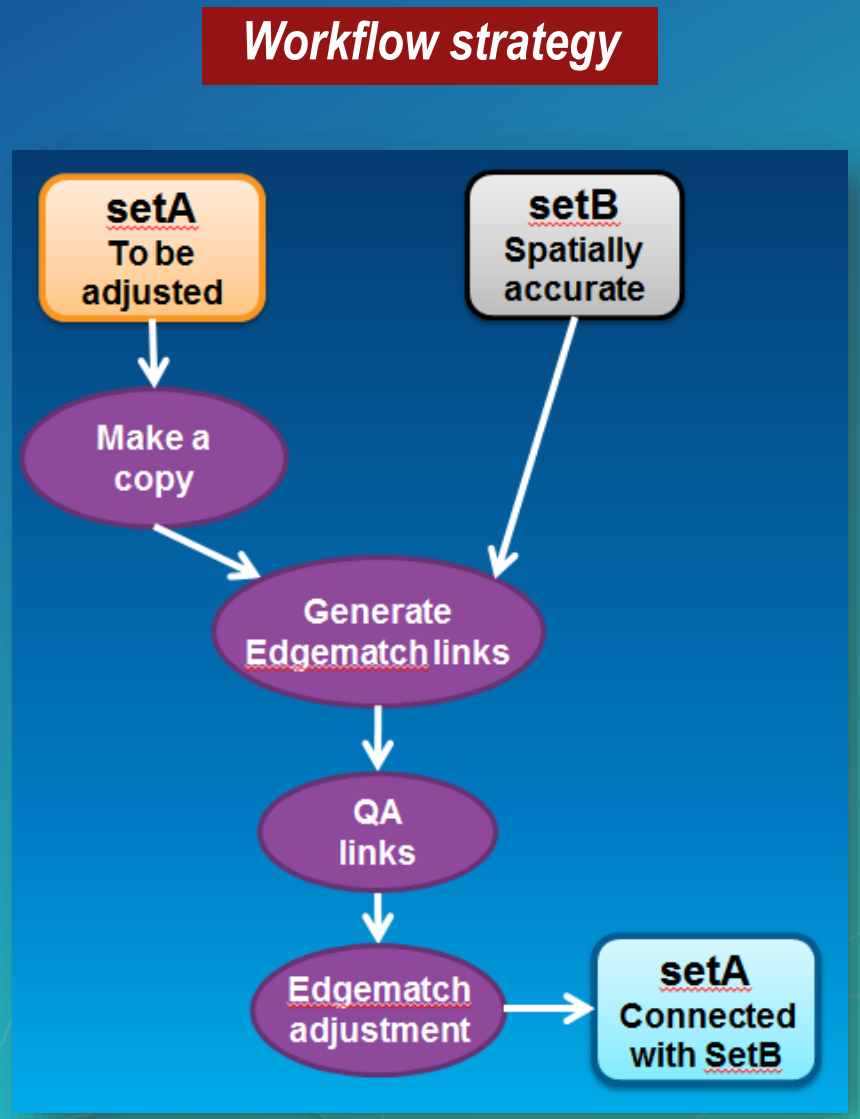
You can also search by “conflation” at arcgis.com to find the download.



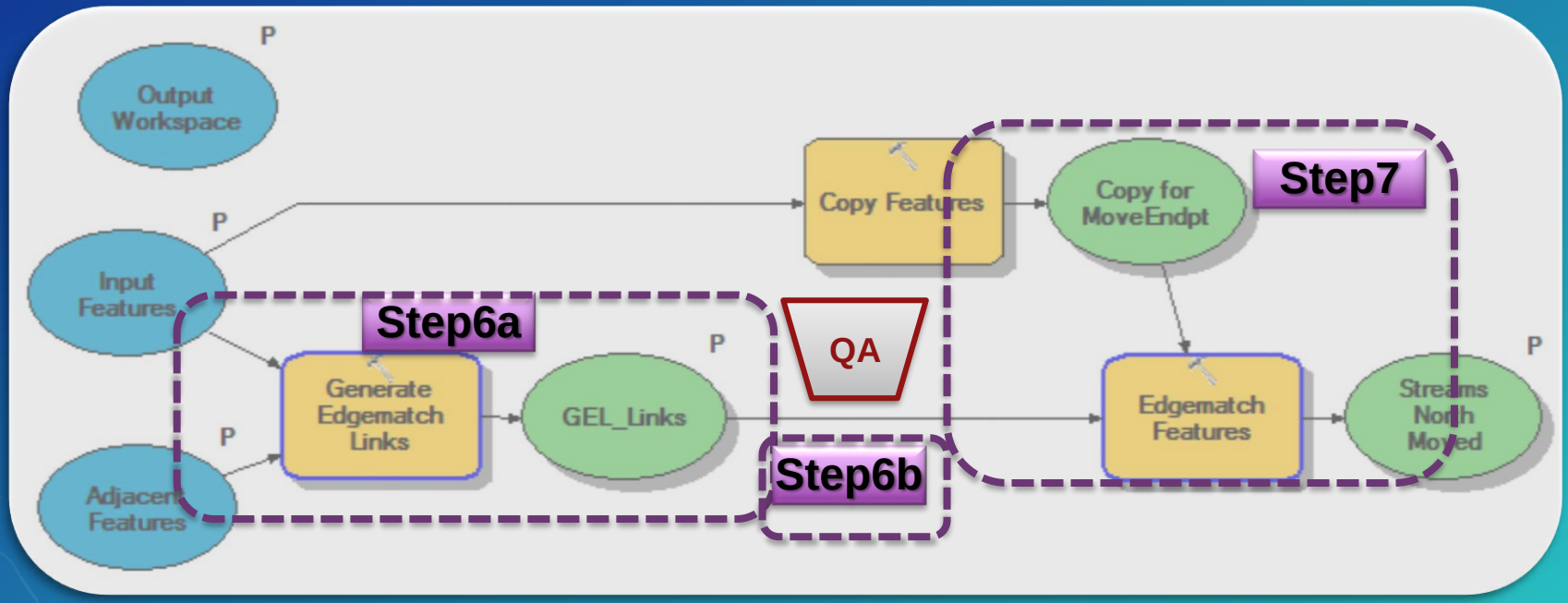


We are improving the add-in toolbar and workflow tools.

Demo:
Real world scenario



Breakdown of the conceptual workflow into sub-steps



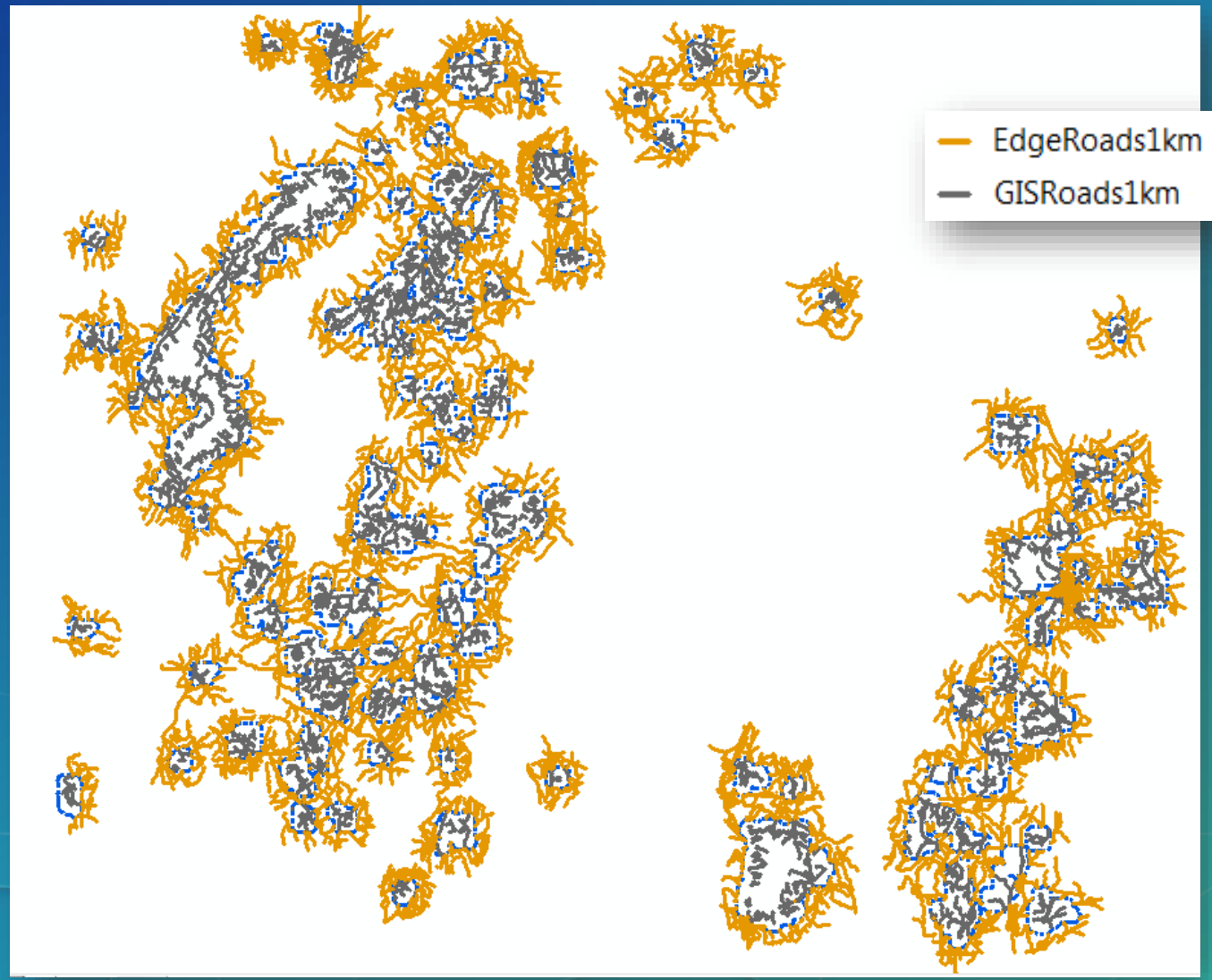
Demo data overview

Two road datasets (an area in Alabama):

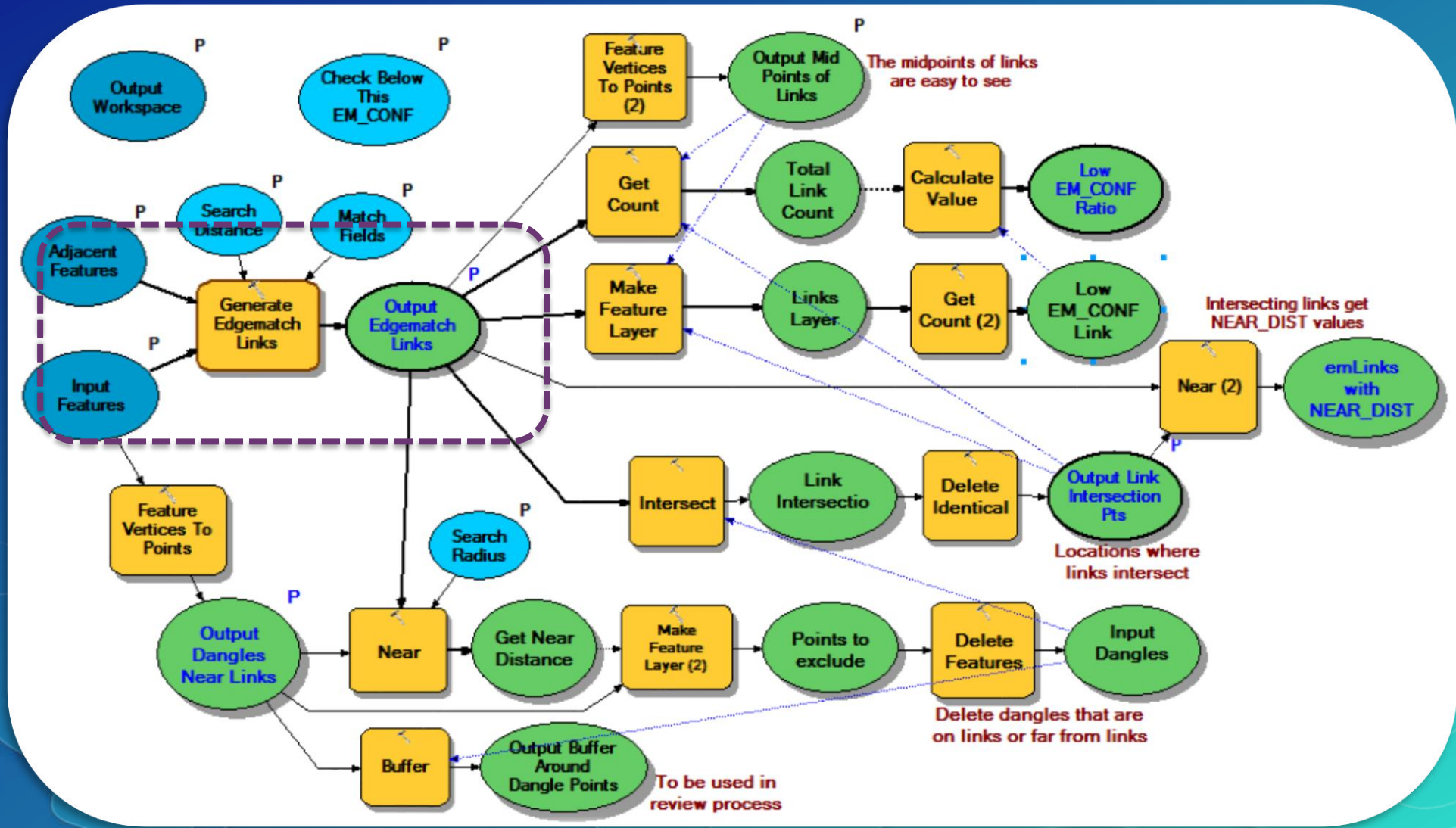
- EdgeRoads – 7576 features
- GISRoads – 3634 features

Both datasets:

- Contain roads that are within 1 km to borders
- Have inconsistent road names

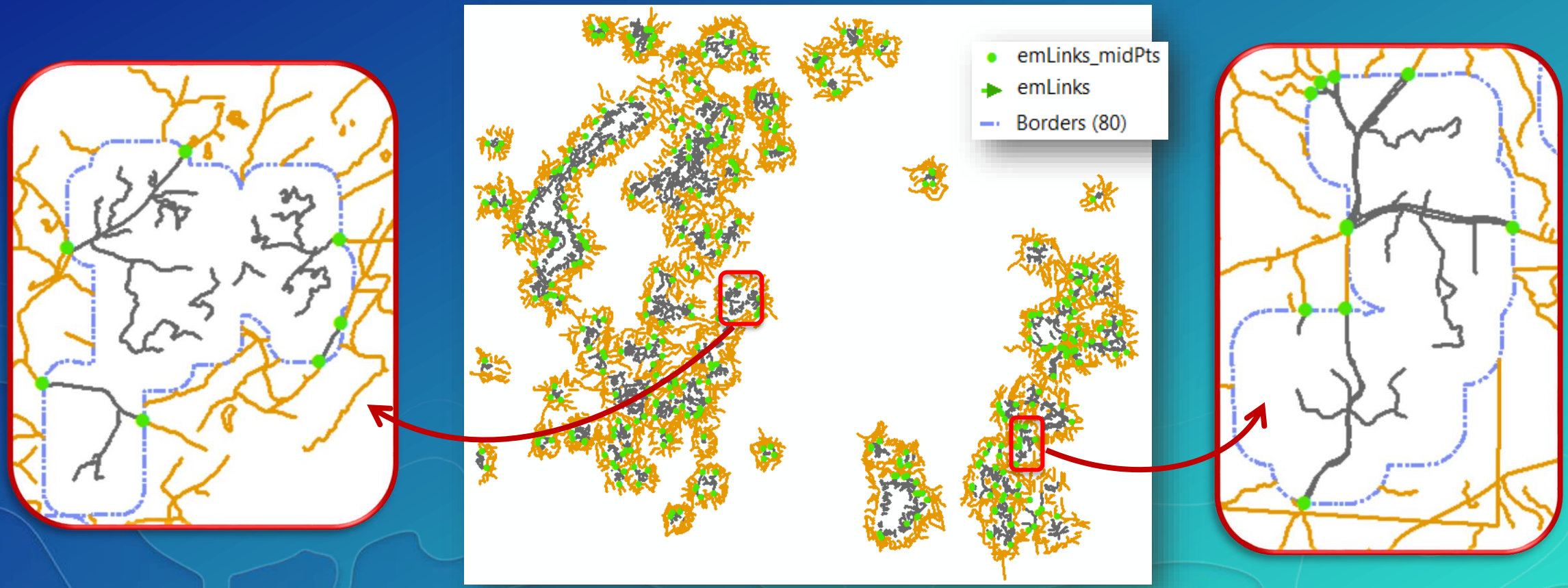


Step6a GEL and Evaluation



GEL result

Generated 454 links; midpoints of links were created for visualization purpose.
Borders were not in the process, but displayed for reference.

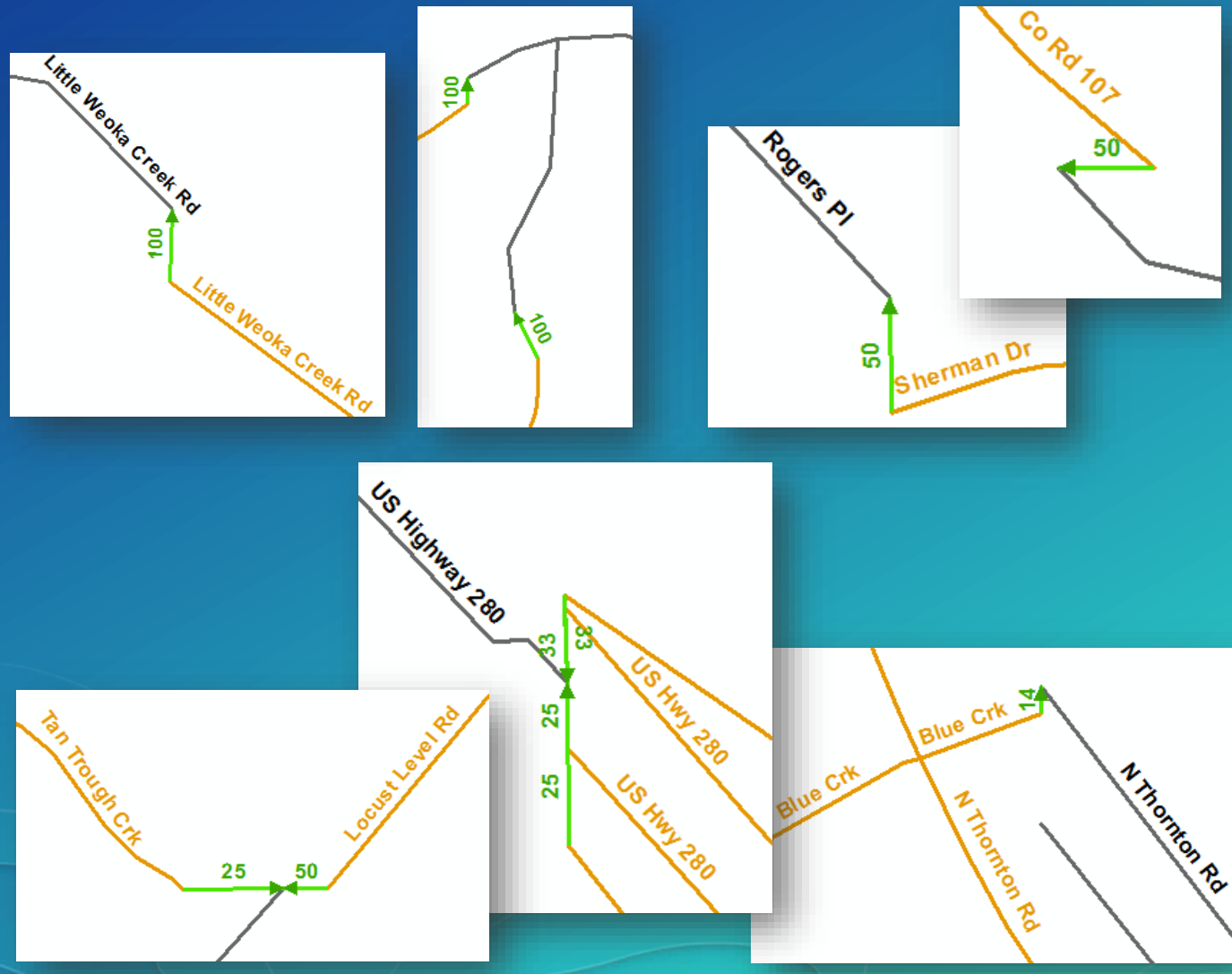


EM_CONF in output

- 100 (matched with no ambiguity)
- 50 (spatially matched with unmatched attributes)
- < 50 (spatially matched with some ambiguity and weak continuity)

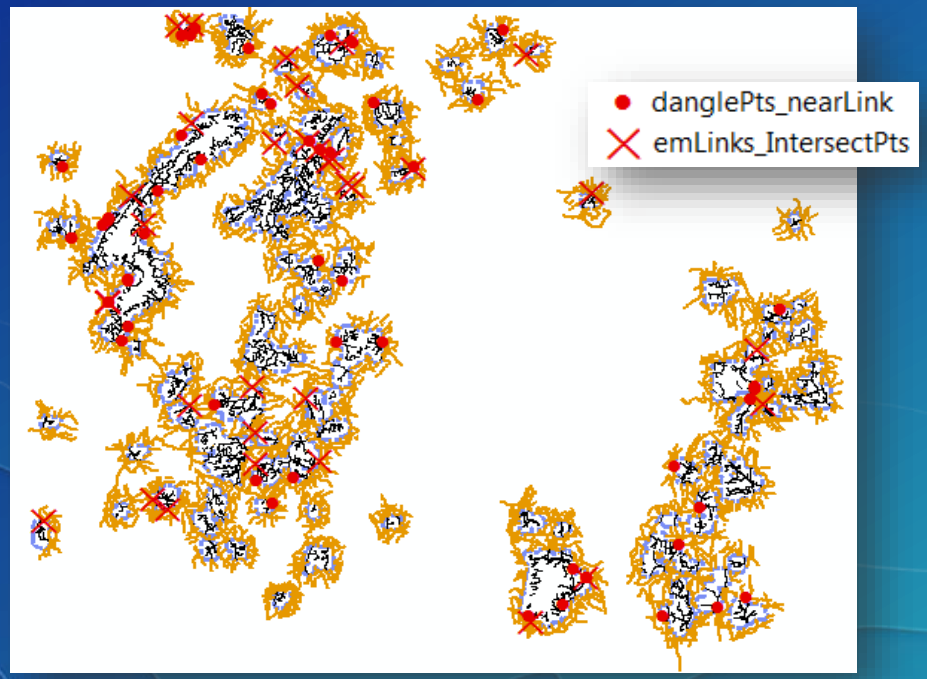
emLinks

OBJECTID *	SHAPE *	SRC_FID	TGT_FID	EM_CONF
402	Polyline	6763	346	100
403	Polyline	6768	607	100
409	Polyline	6854	3047	100
418	Polyline	6946	2898	100
421	Polyline	7019	2053	100
425	Polyline	7148	3343	100
442	Polyline	7420	2240	100
450	Polyline	7532	1390	100
2	Polyline	51	46	50
3	Polyline	51	42	50
5	Polyline	207	21	50
6	Polyline	227	51	50
7	Polyline	359	1618	50
10	Polyline	397	890	50
12	Polyline	404	1366	50



GEL evaluation results

EM_CONF < 33: 134 links
Intersecting links: 33 locations
Potential missing links: 62 source dangle locations



Table

emLinks

OBJECTID *	SHAPE *	SRC_FID	TGT_FID	EM_CONF	SHAPE_Length	NEAR_FID	NEAR_DIST
107	Polyline	2093	1444	4	40.738594	19	0.000043
326	Polyline	5493	1439	6	23.537258	19	0.00002
1	Polyline	20	21	33	7.172191	1	0
5	Polyline	207	21	50	6.374133	1	0
8	Polyline	364	3041	20	37.807129	3	0
14	Polyline	442	3041	50	11.227967	3	0
23	Polyline	754	3139	25	32.194831	5	0
26	Polyline	828	3598	10	53.134568	7	0
36	Polyline	986	3071	100	0.625383	9	0
37	Polyline	987	3117	11	41.232301	59	0
44	Polyline	1068	2613	33	46.451843	11	0
54	Polyline	1326	2870	20	7.738393	62	0
57	Polyline	1350	3386	100	48.961098	13	0
63	Polyline	1407	2000	25	33.527426	15	0
65	Polyline	1414	1113	25	59.985441	17	0
67	Polyline	1458	2870	17	21.200973	62	0
73	Polyline	1581	3117	50	0.0806	59	0

It's time for inspection …

Inspection and editing of edgematch links

Reviewed:

- 33 locations of intersecting links

  NEAR_DIST >= 0
- 98 low EM_CONF links

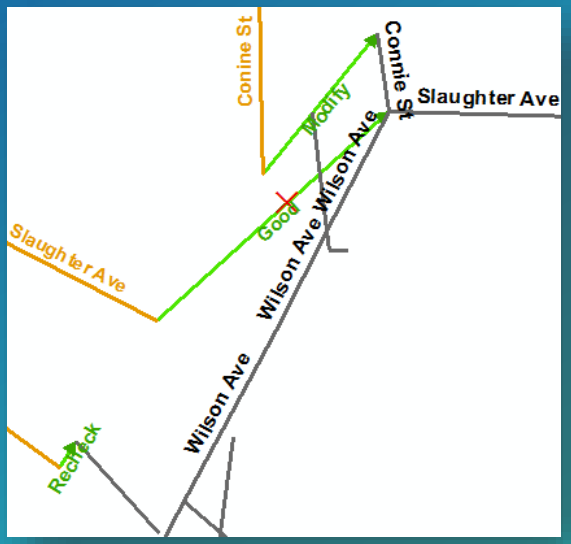
  (EM_CONF < 33) AND (REV_FLAG IS NULL)
- 62 source dangle locations (near links)

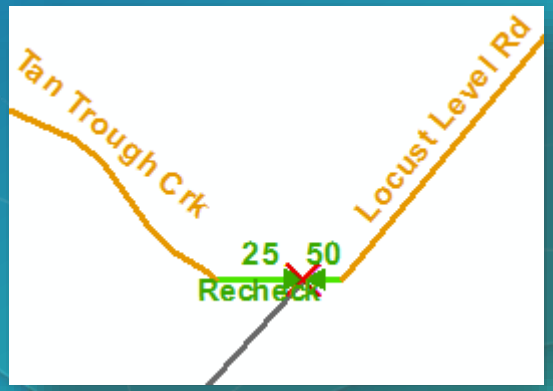
Summary:

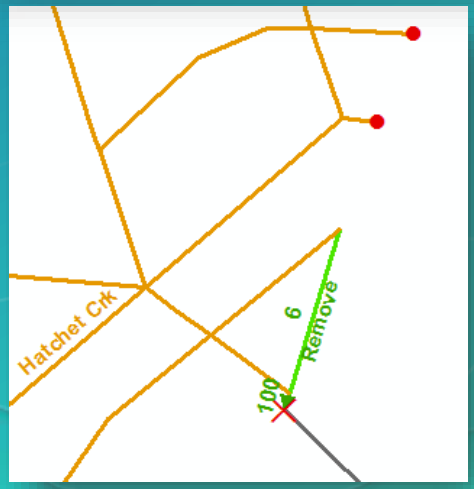
- 388 (~85%) of total 459 links were good (54 were flagged for recheck)
- 71 (~15%) of total links were modified, removed, or added

emLinks_freqREVFLAG

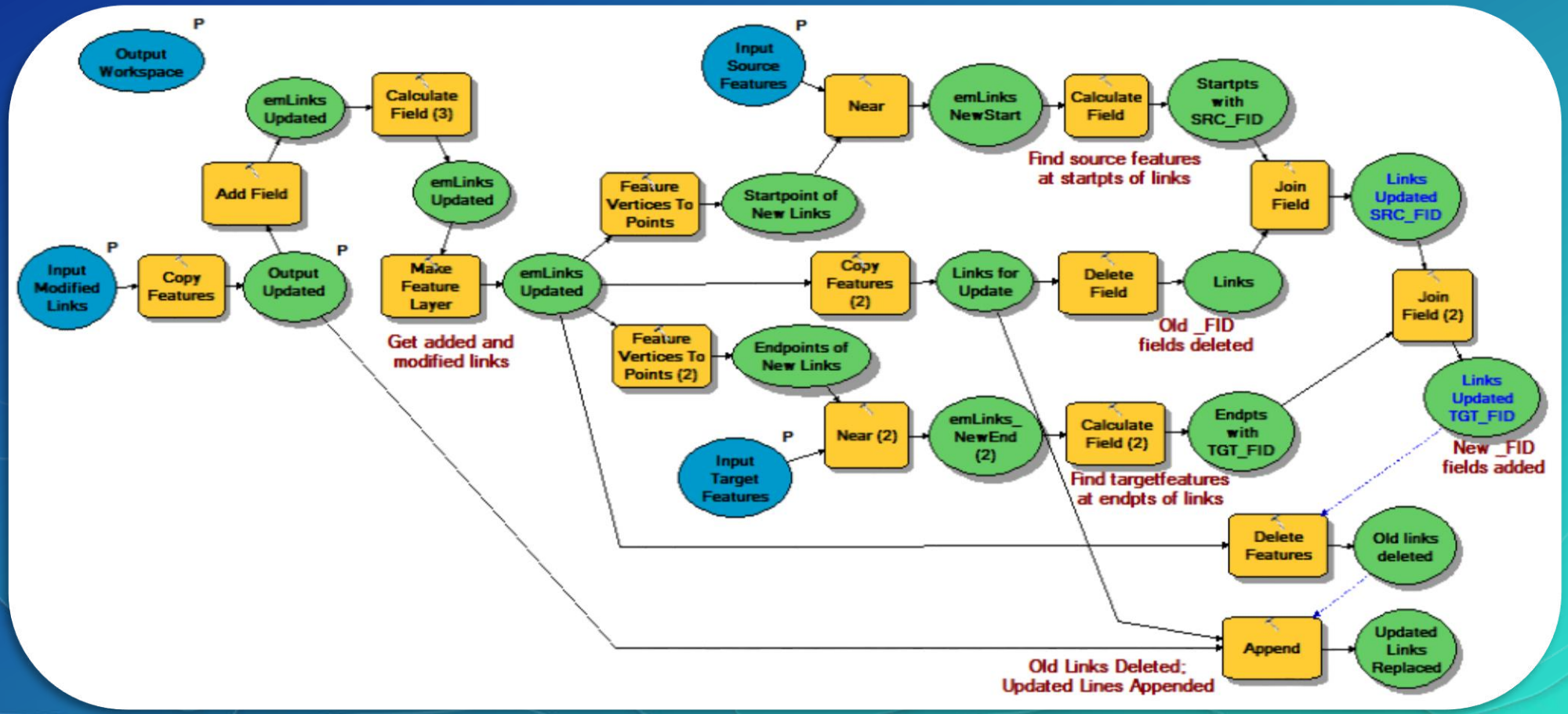
OBJECTID *	FREQUENCY	REV_FLAG
1	267	<Null>
2	5	Added
3	66	Good
4	46	Modify
5	55	Recheck
6	20	Remove



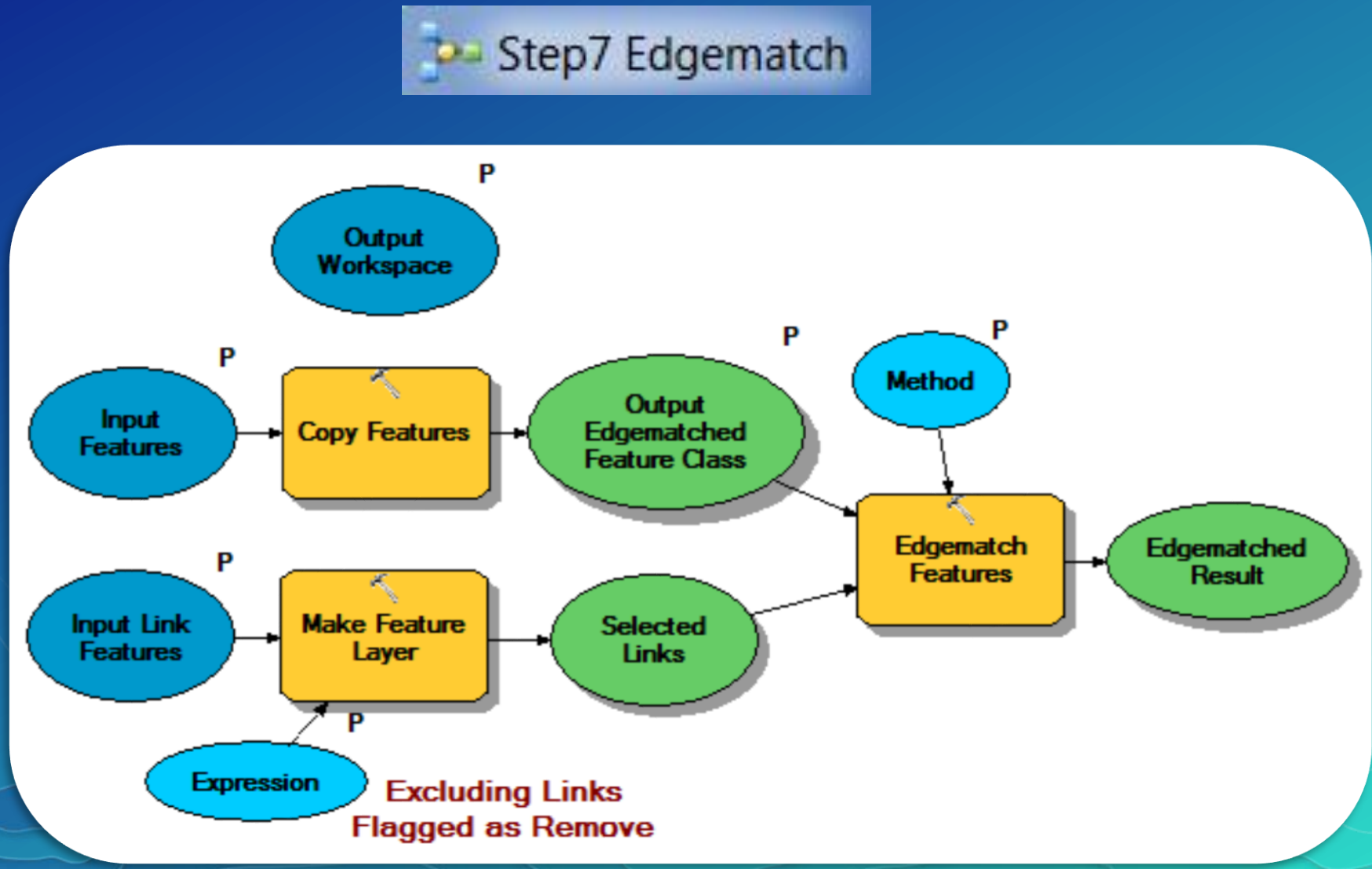




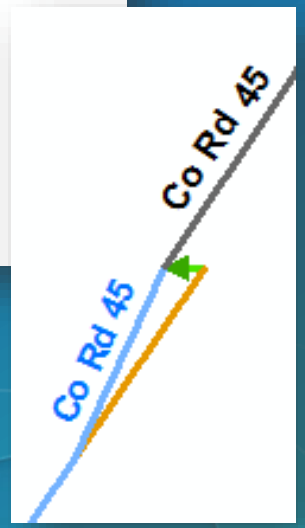
What happened to the SRC_FID and TGT_FID of the added or modified links?

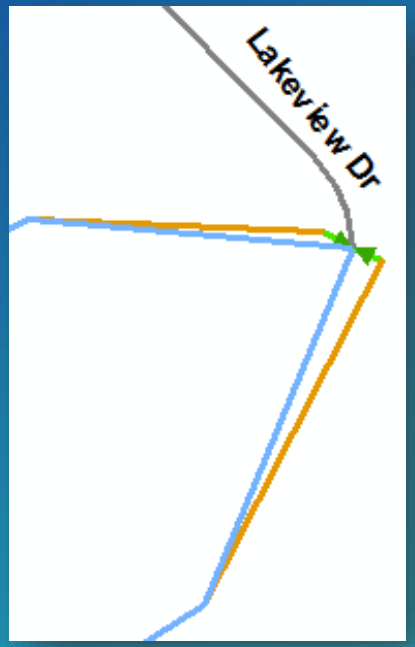


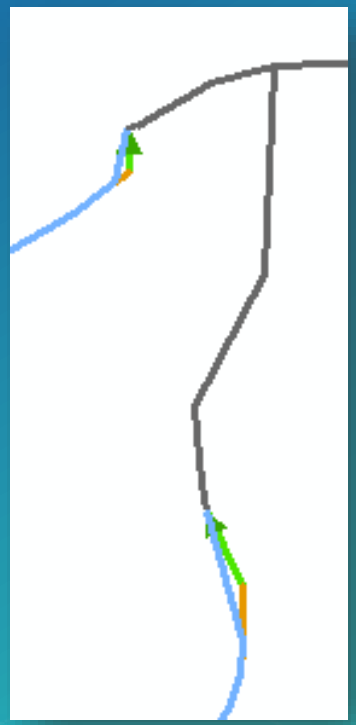
Edgematch Features



Edgematch result







Review flagged locations …

Edgematching of adjacent datasets workflow completed!

Automated processing

	Processing Time
Step 6a	6.52 sec
Step 6b	4.09 sec
Step 6c	2.15 sec
Total	12.76 sec

Interactive processing (not counting final review)

	QA Links	Time (2-3 review counts per minute)
Review Count (locations or feature groups)	~ 193	~ 1 - 1.6 hrs.
Edit Count (field values)	192	

Conclusions and Future Work

Thanks to:

- Department of Public Works (DPW), Los Angeles County, USA.
- Resource Management Service, LLC, Birmingham, AL, USA.
- All others who supported us along the way.

Edge matching can be done more efficiently now

It takes a workflow:

- **Use the best practice in preprocessing.**
- **Run automated tools to obtain highly accurate results and evaluation information.**
- **Interactively review and edit the results. The time is worth-investing.**

Consider conflation a higher priority

Study the tools and workflows; understand the results

- Start with small test areas

Customize the workflows for your organizations

- Improve data quality and usability
- Bring new live and value to your data

Work with broader communities

- Data sharing and collaboration
- Seamless analysis and mapping

Please send us your feedbacks and share your stories … ☺

Future work

New tools and enhancements

- Split Line By Match tool (for overlapping datasets)
- Better feature matching
- Tools for other feature types

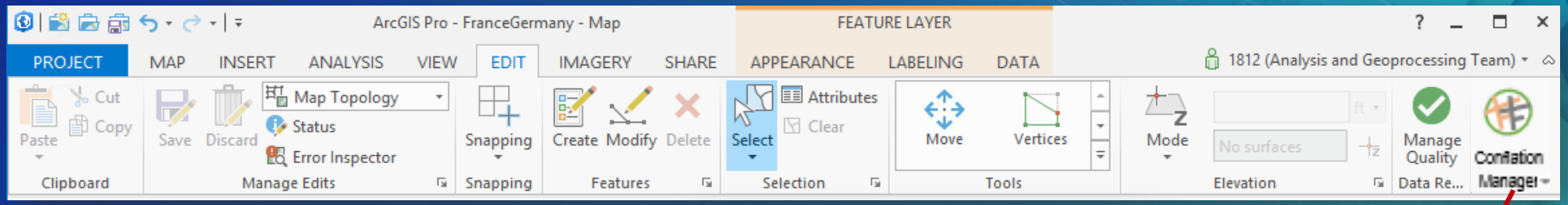
Integrated processing and inspections

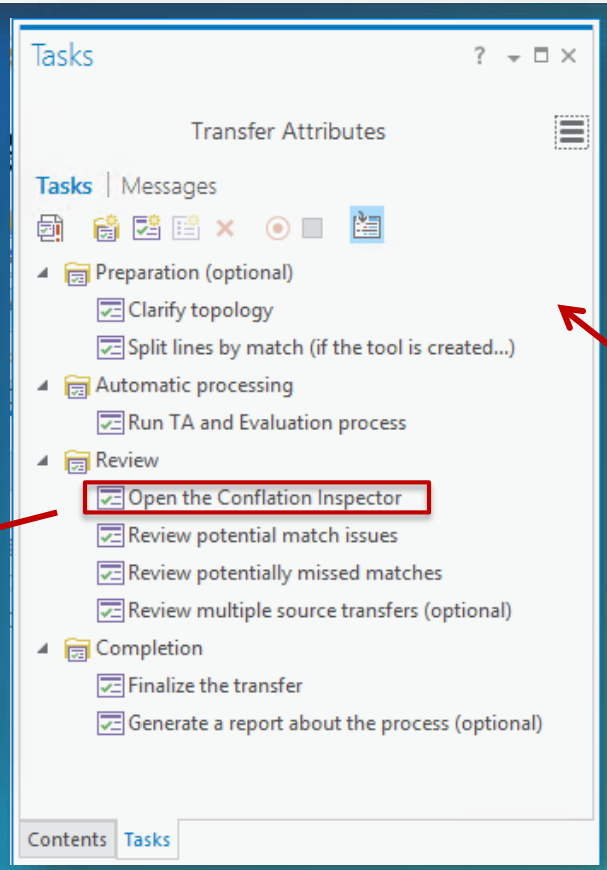
- Design of Conflation Manager is underway

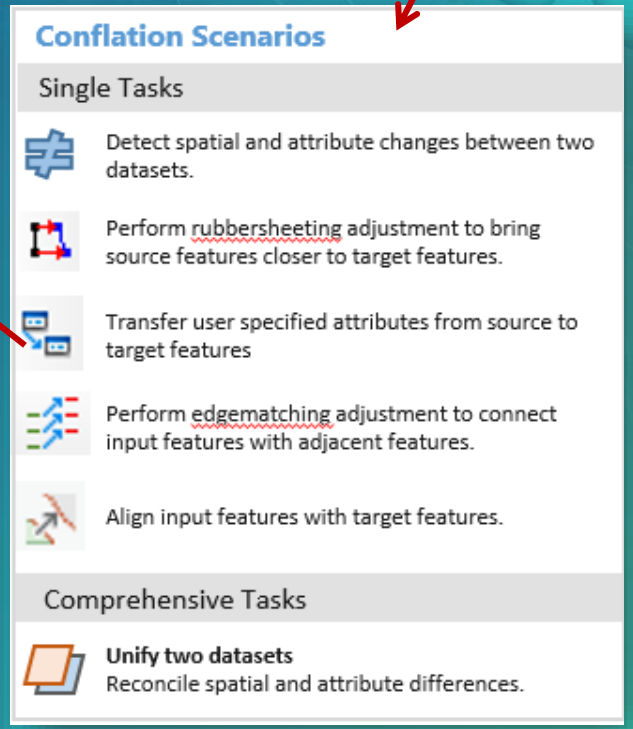
Formalization of workflows

- Common scenarios (e.g. multi-scale data updating, linking buildings of different scale)
- Incorporation of other data sources (imagery, lidar, GPS)
- Contextual conflation (spatially related features)

Conflation Manager (ConfMgr)

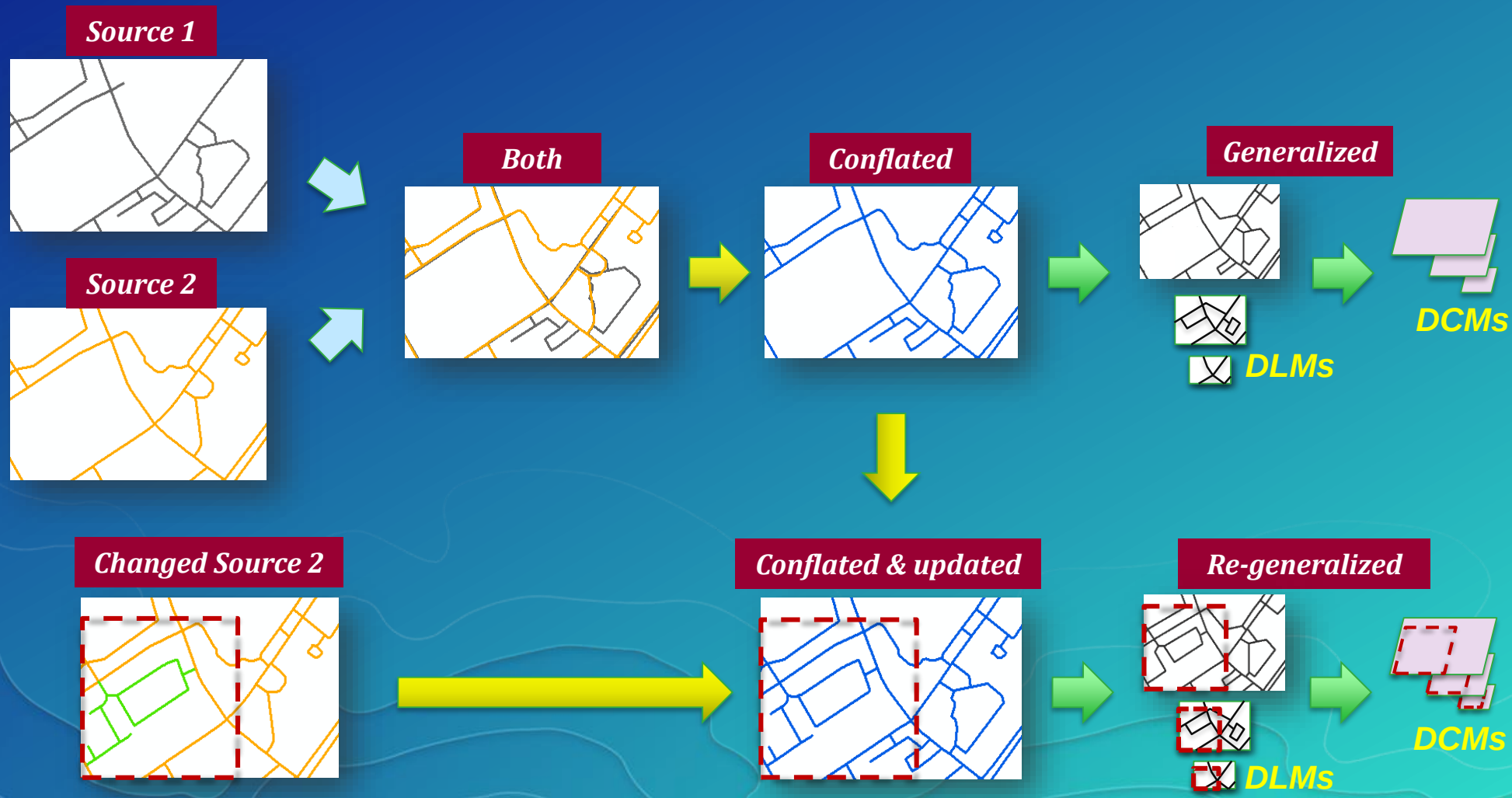






Conflation in multi-scale data updating and mapping

DLM – digital landscape model; DCM – digital cartographic model



Recent papers

- Baella B, Lee D, Lleopart A, Pla M (2014) ICGC MRDB for topographic data: first steps in the implementation, The 17th ICA Generalization Workshop, 2014, Vienna, Austria.
  https://kartographie.geo.tu-dresden.de/downloads/ica-gen/workshop2014/genemr2014_submission_8.pdf

- Lee D, Yang W, Ahmed N (2017), Road data conflation – the key step to geospatial data enhancement, The 27th International Cartographic Conference, 2017, Washington DC, USA.

- Lee D (2015), Using Conflation for Keeping Data Harmonized and Up-to-date, ICA-ISPRS Workshop on Generalisation and Multiple Representation, 2015, Rio de Janeiro, Brazil.
  https://kartographie.geo.tu-dresden.de/downloads/ica-gen/workshop2015/genemr2015_submission_8.pdf

- Lee D, Yang W, Ahmed N (2015) Improving Cross-border Data Reliability Through Edgematching, to be presented at The 27th International Cartographic Conference, 2015, Rio de Janeiro, Brazil. http://www.icc2015.org/abstract,670.html

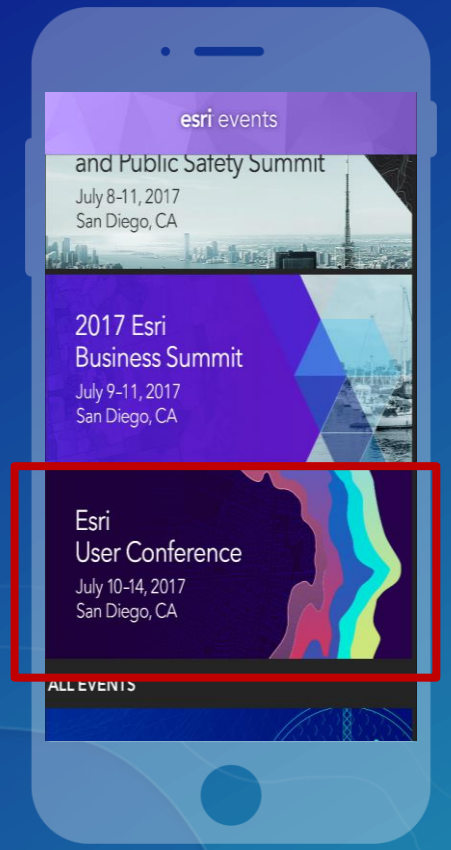
- Lee D, Yang W, Ahmed N (2014) Conflation in Geoprocessing Framework - Case Studies, GEOProcessing, 2014, Barcelona, Spain. http://goo.gl/iOoSGV

- Yang W, Lee D, and Ahmed N, "Pattern Based Feature Matching for Geospatial Data Conflation", GEOProcessing, 2014, Barcelona, Spain. http://goo.gl/JKGJbo

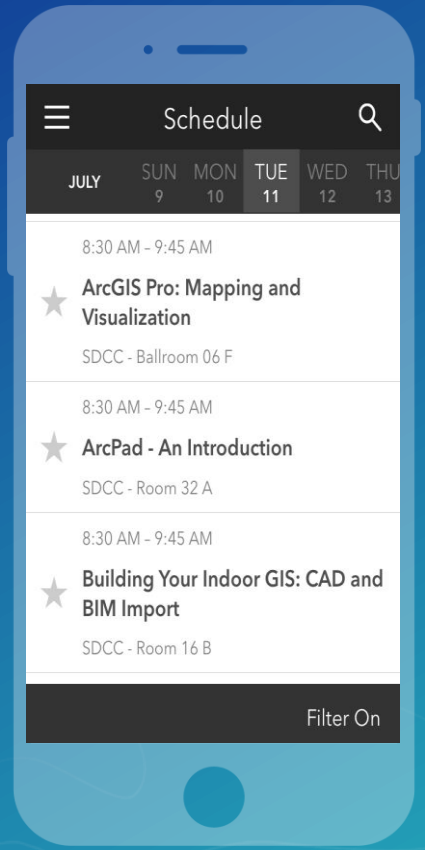
Please Take Our Survey on the **Esri Events App!**

Your feedback allows us to help maintain high standards and to help presenters

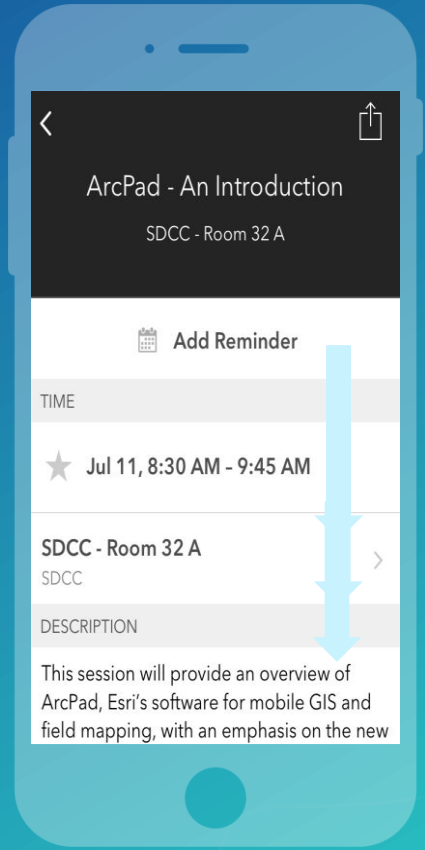
Download the Esri Events app and find your event



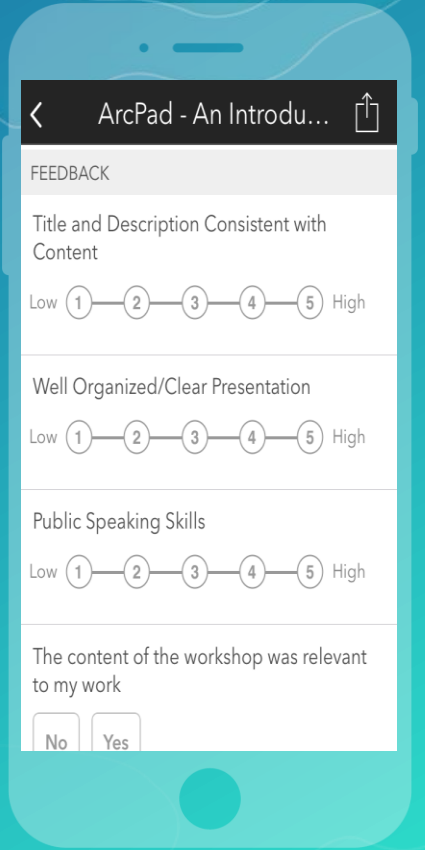
Select the session you attended



Scroll down to find the survey



Complete Answers and Select "Submit"



Thank you for attending!

esri
THE
SCIENCE
OF
WHERE