Conflation Tools and Workflows: An Introduction

Dan Lee

(dlee@esri.com)

ESRI USER CONFERENCE

SEE WHAT OTHERS CAN'T

Agenda

- Introduction
  - Multi-source data issues
  - Geoprocessing conflation tools
- Conflation Workflows
  - Automatic and interactive processes
  - Demo on Transfer Attributes workflow
- User Stories and Future Work

Introduction

Inconsistent multi-source data and common needs

For overlapping datasets:

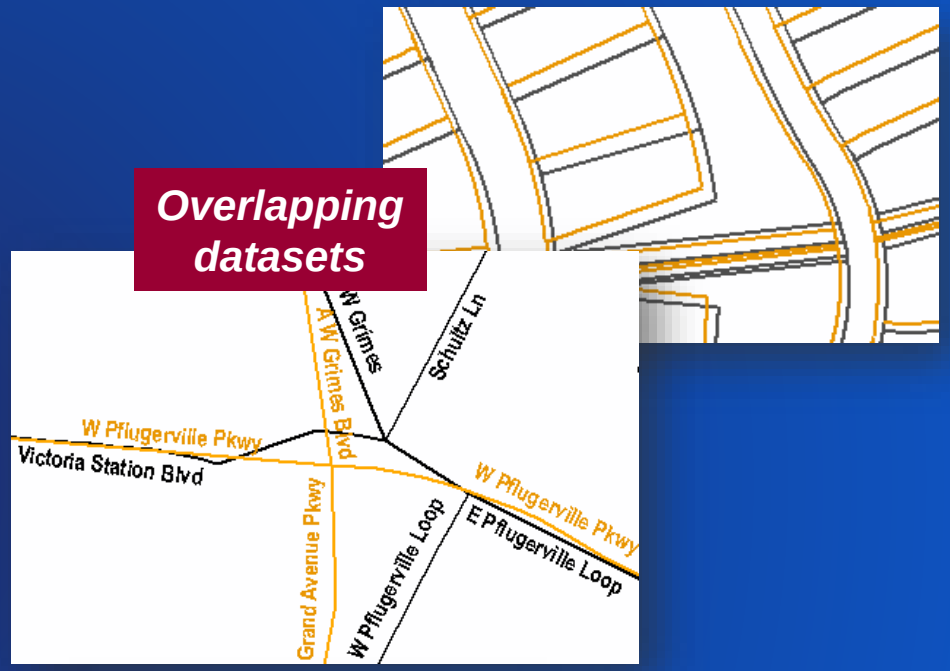
- Find spatial and attribute changes through feature matching
- Reconcile the differences

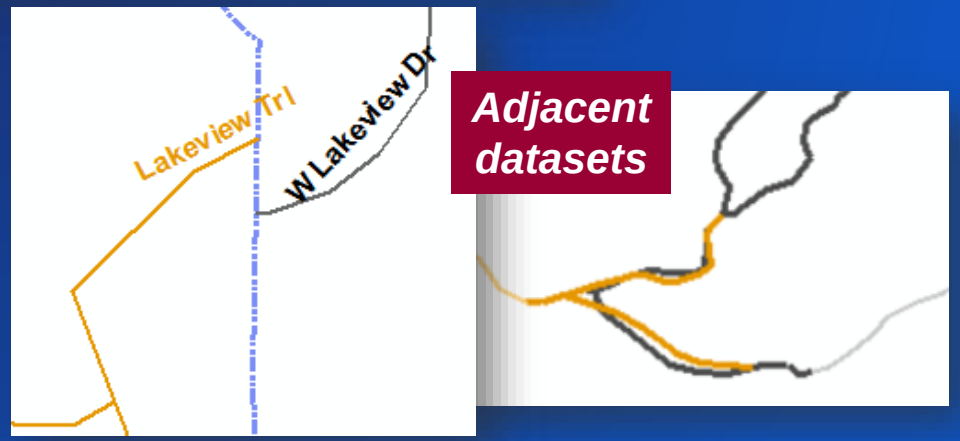
For adjacent datasets:

- Resolve disconnections and conflicts in bordering areas

Ultimate goal:

- Maintain a unified and seamless dataset for reliable analysis and quality mapping





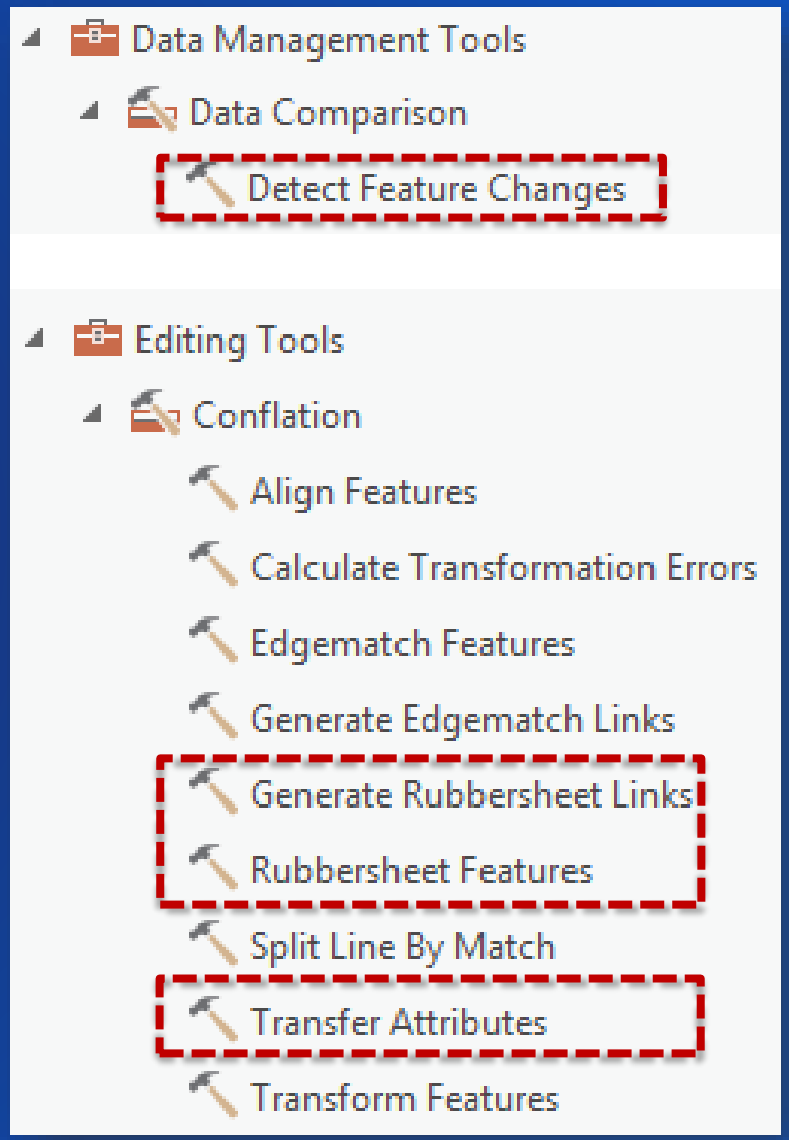
Conflation tools in geoprocessing

In ArcGIS Pro 2.4 and 10.7.1

System tools for overlapping datasets

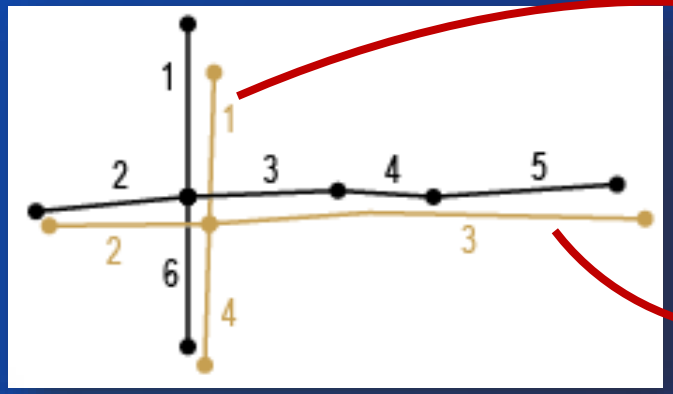
- Focus on common use cases
- Aim at high matching accuracy (not promising 100%)
- Provide information to facilitate post-review and edit

Supplemental workflow tools



Feature matching (FM) for overlapping datasets

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

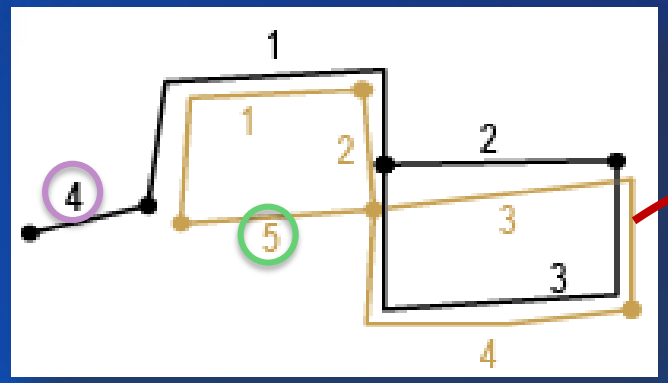


Source feature

Target feature

1:1 and 1:m matches

OBJECTID *	SRC_FID	TGT_FID	FM_GRP	FM_MN	FM_CONF
1	1	1	1	1:1	100
2	2	2	2	1:1	100
3	3	3	3	1:3	100
4	3	4	3	1:3	100
5	3	5	3	1:3	100
6	4	6	4	1:1	100



m:1 and m:n matches

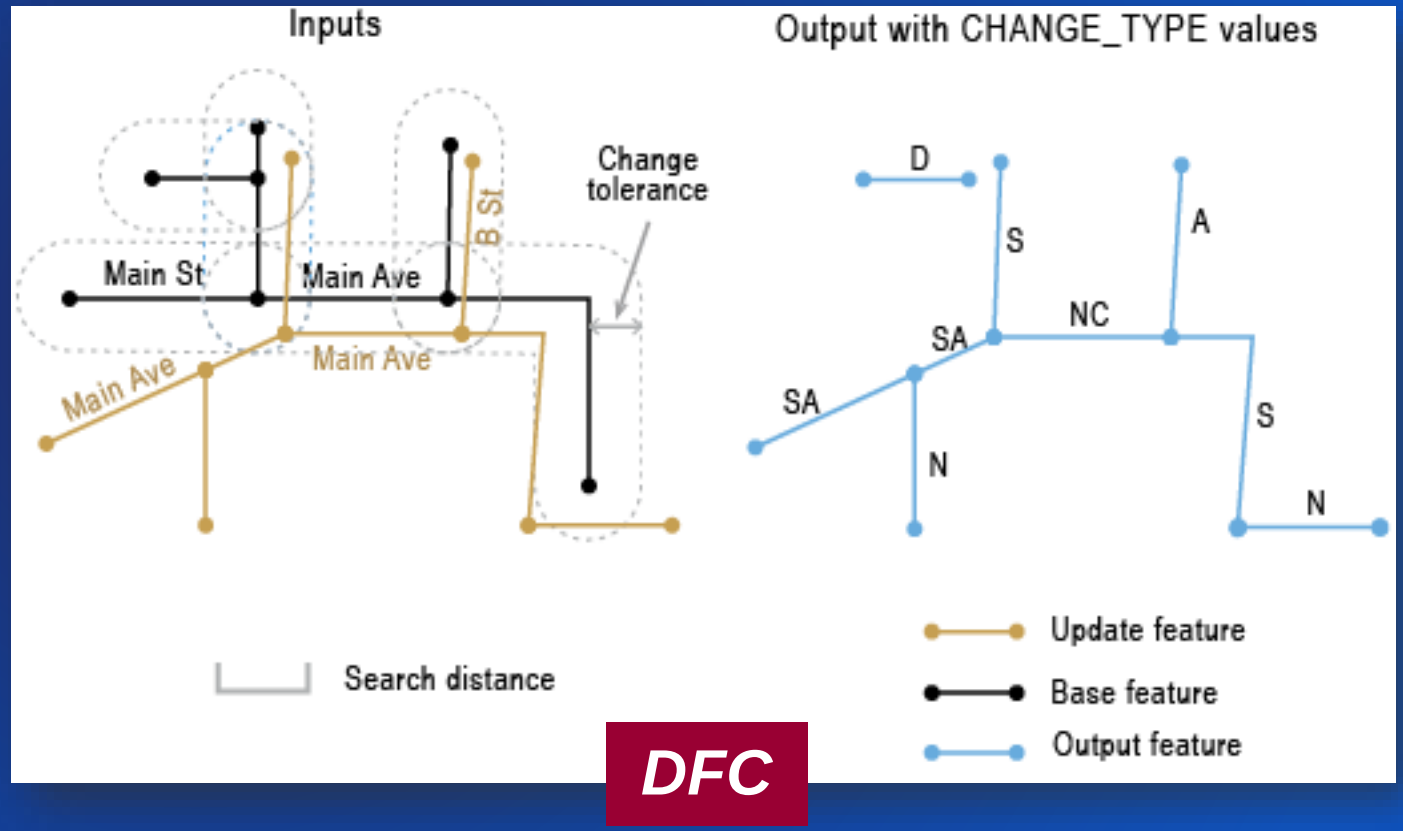
OBJECTID *	SRC_FID	TGT_FID	FM_GRP	FM_MN	FM_CONF
1	1	1	1	2:1	100
2	2	1	1	2:1	100
3	3	2	2	2:2	100
4	3	3	2	2:2	100
5	4	2	2	2:2	100
6	4	3	2	2:2	100
7	5	-1	-1	N/A	0
8	-1	4	-1	N/A	0

FM-based tool #1 - Detect Feature Changes (DFC)

Finds and reports feature differences

Output CHANGE_TYPE

- Spatial change (S)
- Attribute change (A)
- Spatial & attribute change (SA)
- Spatial and line direction change (S_LD)
- Spatial, attribute, and line direction change (SA_LD)
- No change (NC)
- Unmatched update feature (N)
- Unmatched base feature (D)

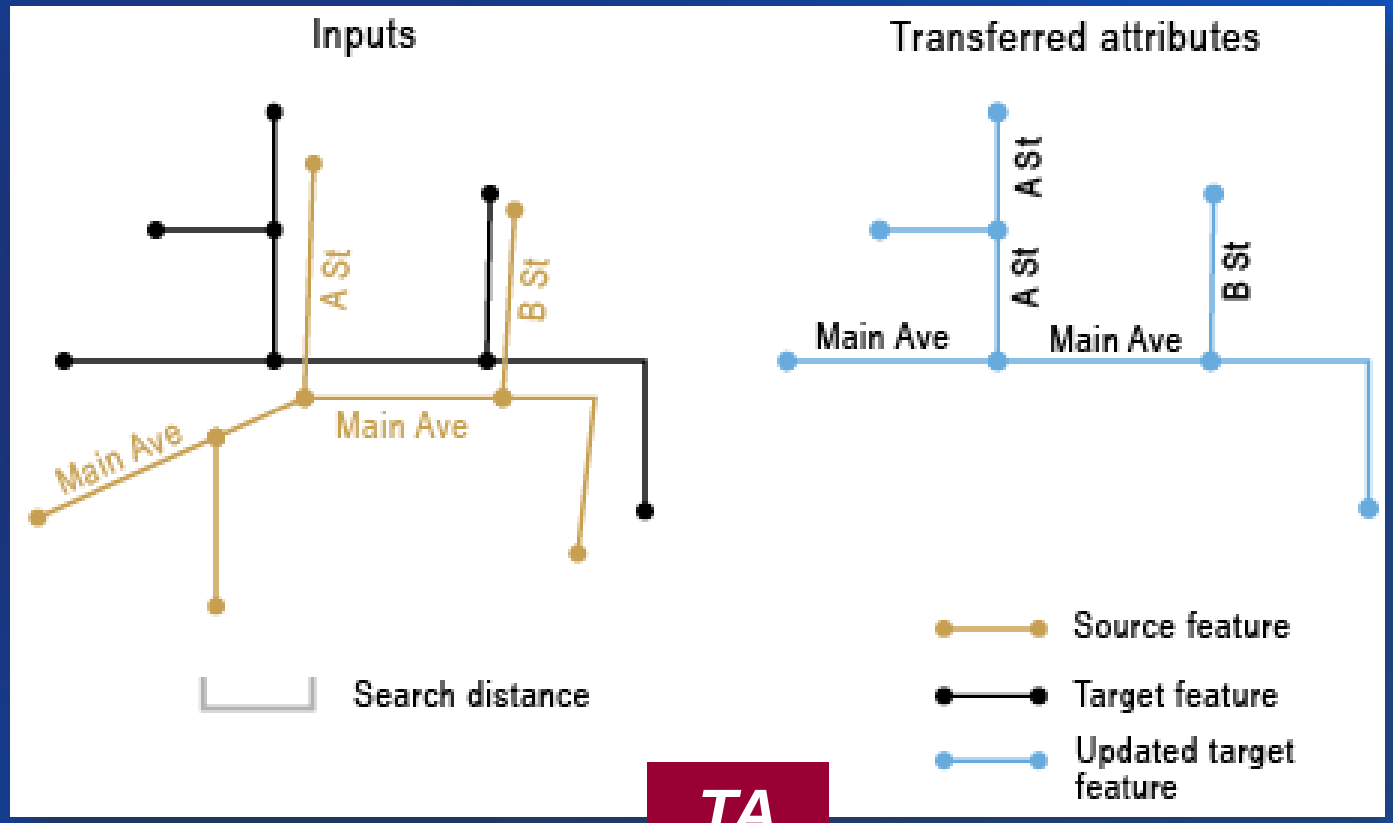


FM-based tool #2 – Transfer Attributes (TA)

Transfers attributes from source to target features

Target features are modified

- Transfer fields (e.g. ROAD_NAME, UniqueID) are added
- Optional transfer rules for m:n matches can be specified



FM-based tool #3 – Generate Rubbersheet Links (GRL)

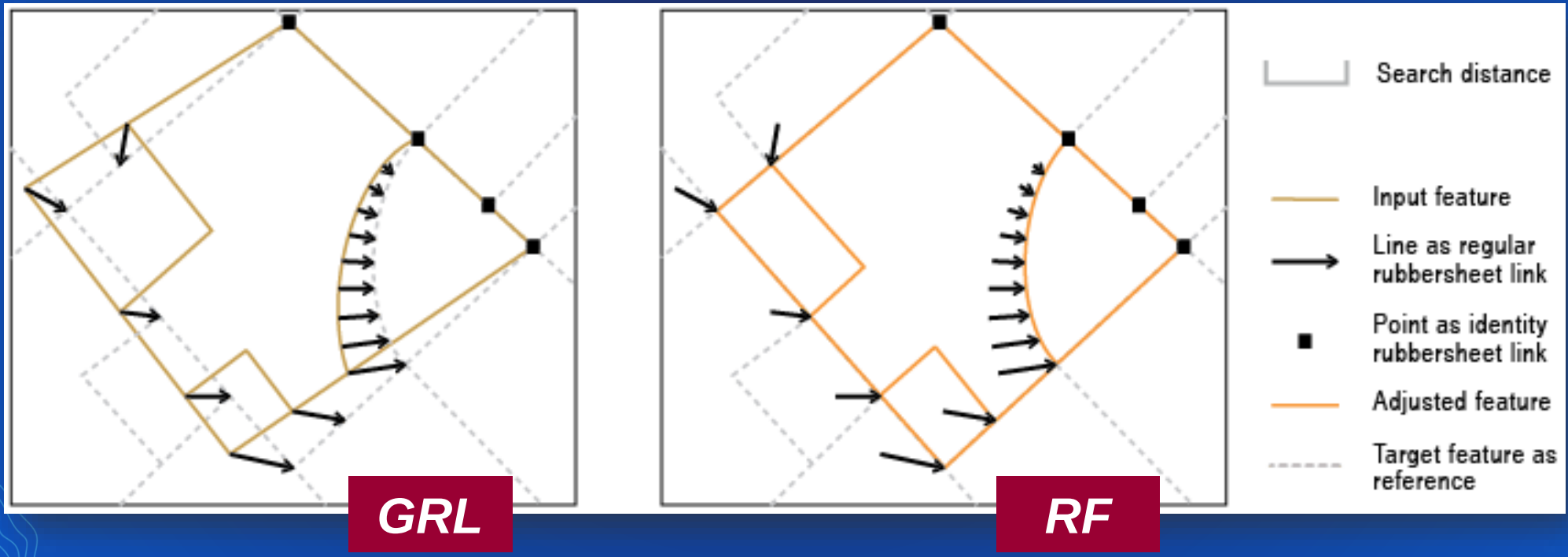
Derives links from source to corresponding target locations for rubbersheeting adjustment

Generate Rubbersheet Links (GRL)

- Regular links and identity links

Followed by Rubbersheet Features (RF)

- Adjusting input features to target locations



Conflation Workflows

Typical processes

- Consistent projection
- Data validation
- Selection of relevant features

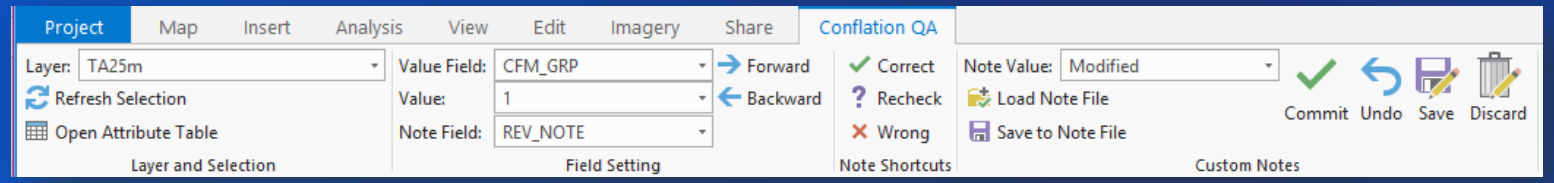
- Conflation tools
- Workflow tools

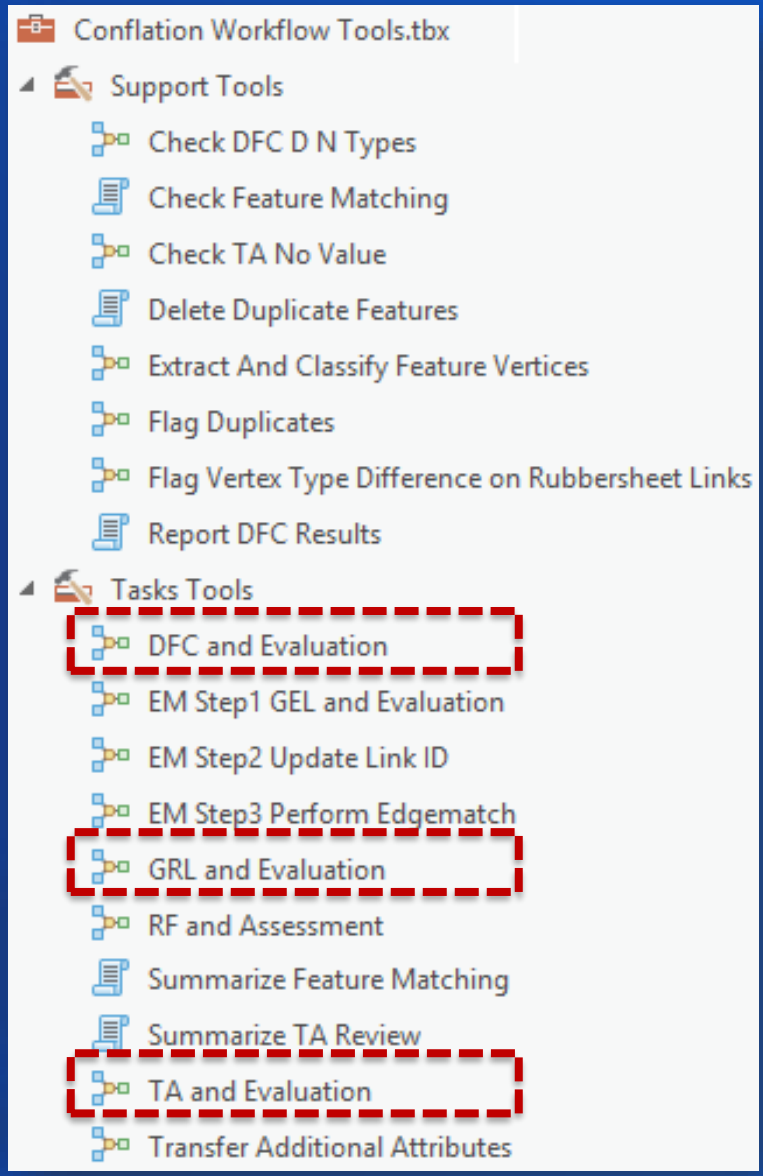
- Interactive review and edit

Supplemental Workflow Tools

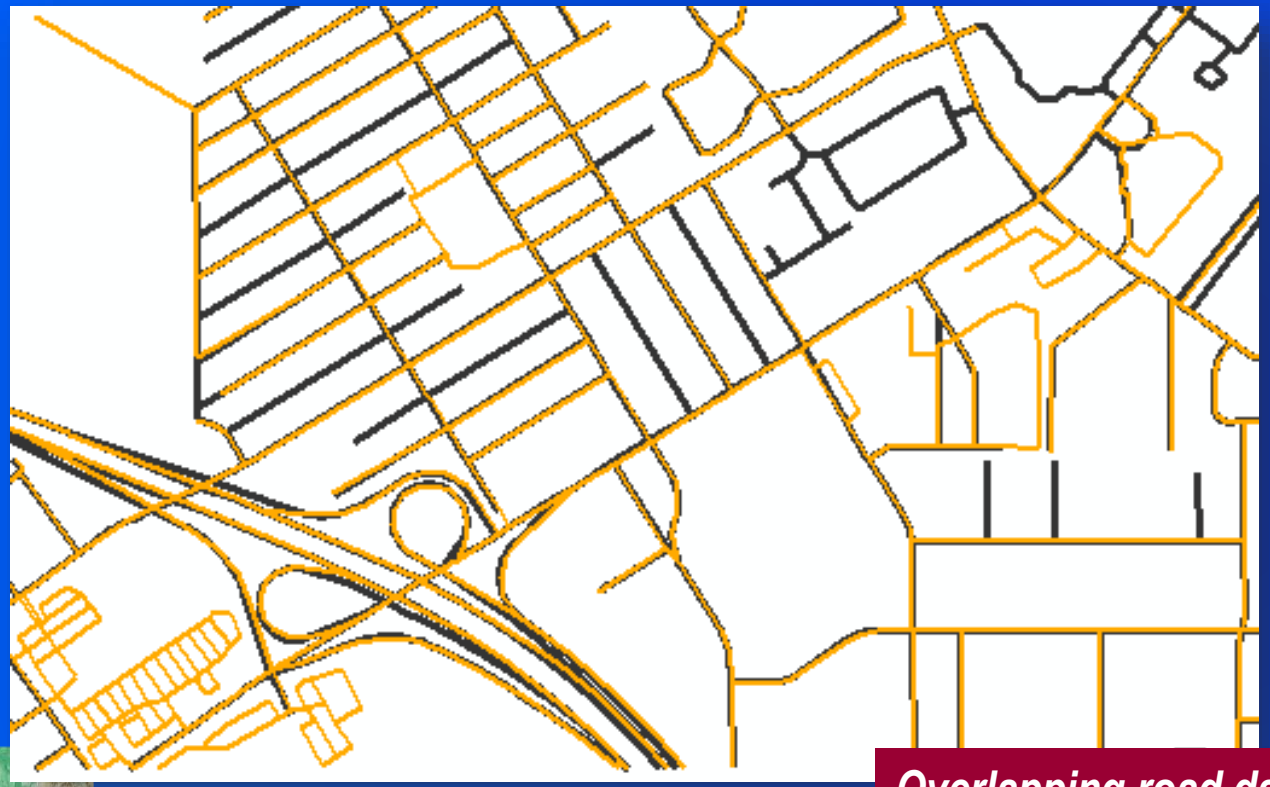
Focusing on popular tasks:

- Each task involves system tool + Evaluation
- Potential issues are flagged in the output
- Review is done interactively using Conflation QA tools (SDK add-in for Pro; python add-in for ArcMap)





Let's take TA workflow as an example …



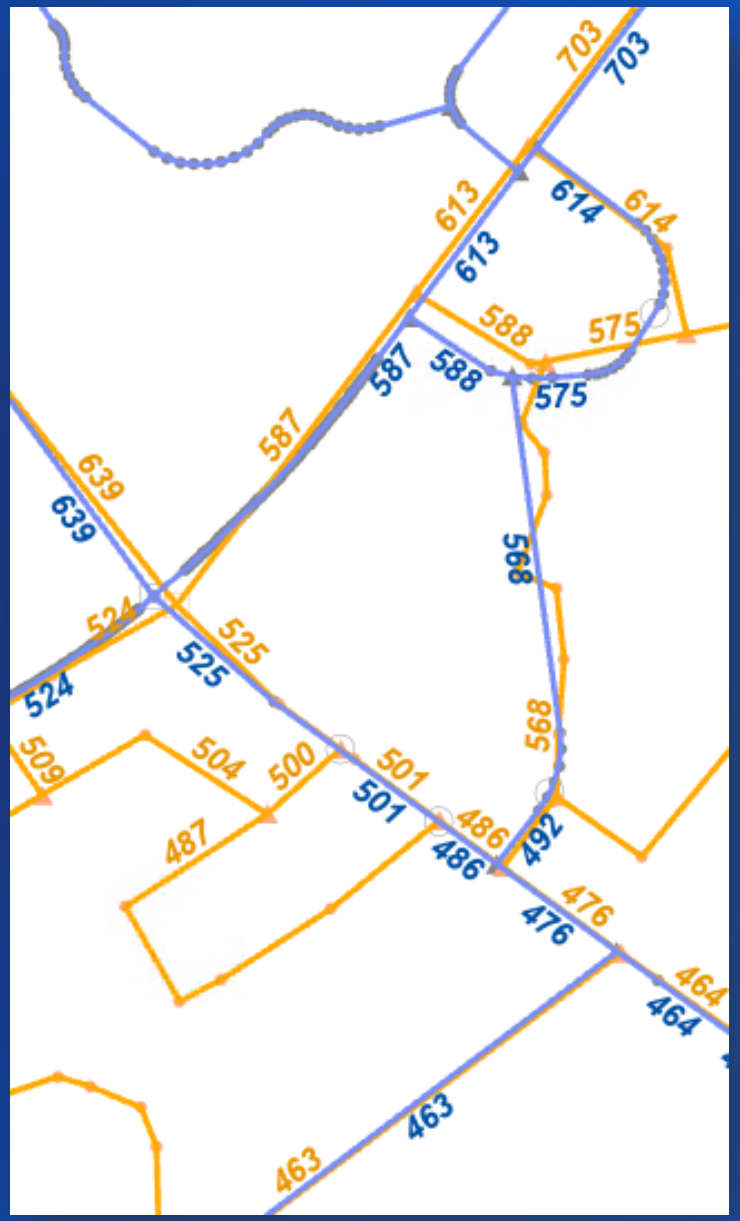
Overlapping road datasets from two sources

Transfer Attributes workflow (demo)

- Using linkage strategy

Linkage

- Common field values for corresponding source and target features
- Key for transferring additional attributes
- Can be established using Transfer Attributes workflow



Establishing linkage for matching features

Two inputs

Source

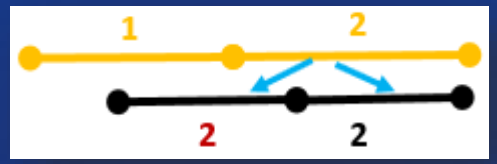
Existing fields…	TA_uID
	1
	2

Target

Existing fields…

Added field with calculated unique values

Automated TA result: showing a 1:2 match



Target received TA_uID

Existing fields…	TA_uID
	2
	2

Interactive review and edit: favoring two 1:1 matches

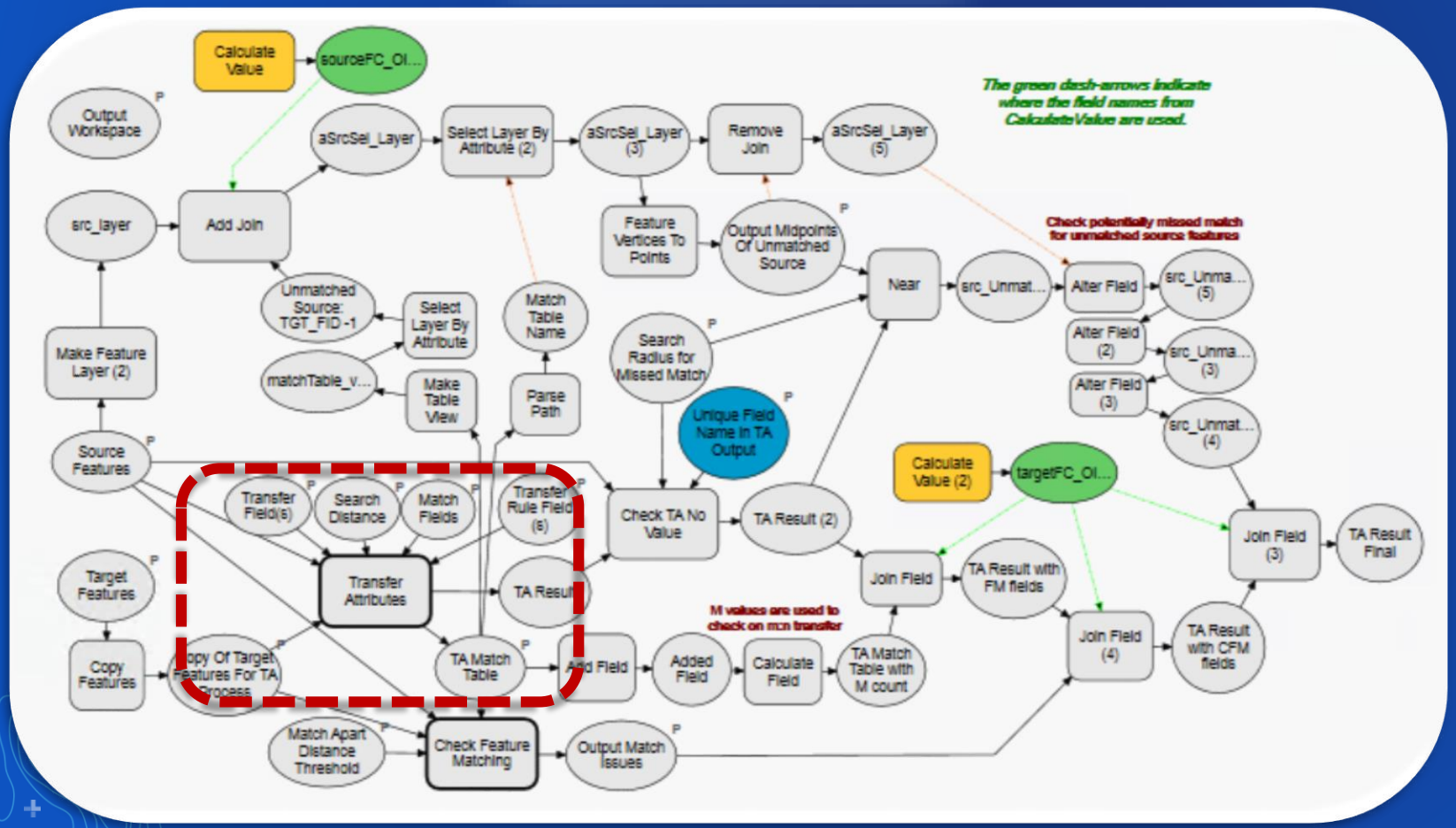
Target after interactive review and edit

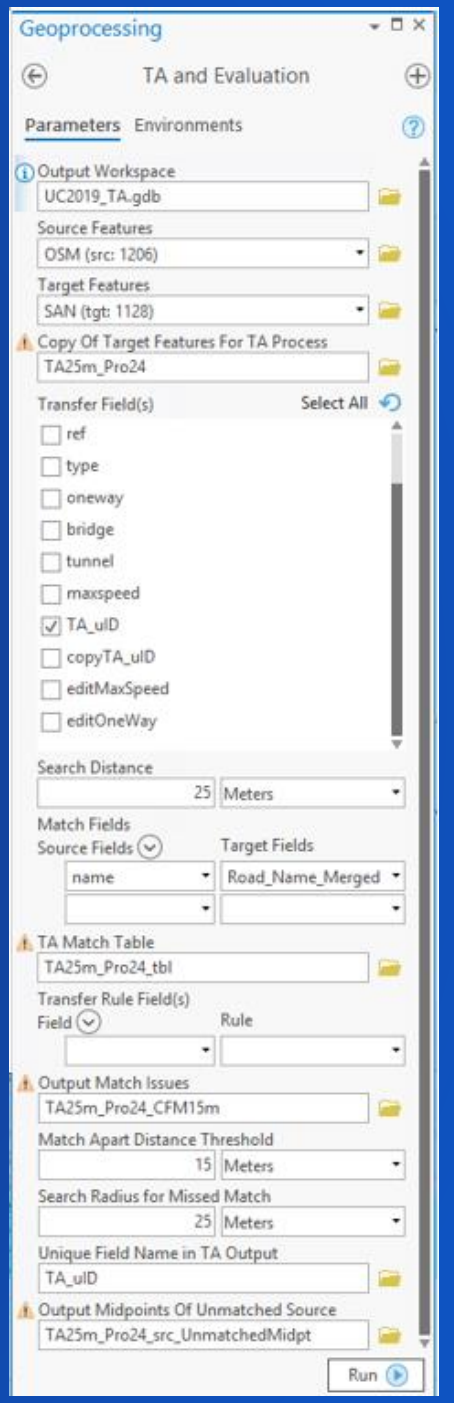
Existing fields…	TA_uID	rev_ TA_uID	final_ TA_uID
	2	1	1
	2		2

Linkage

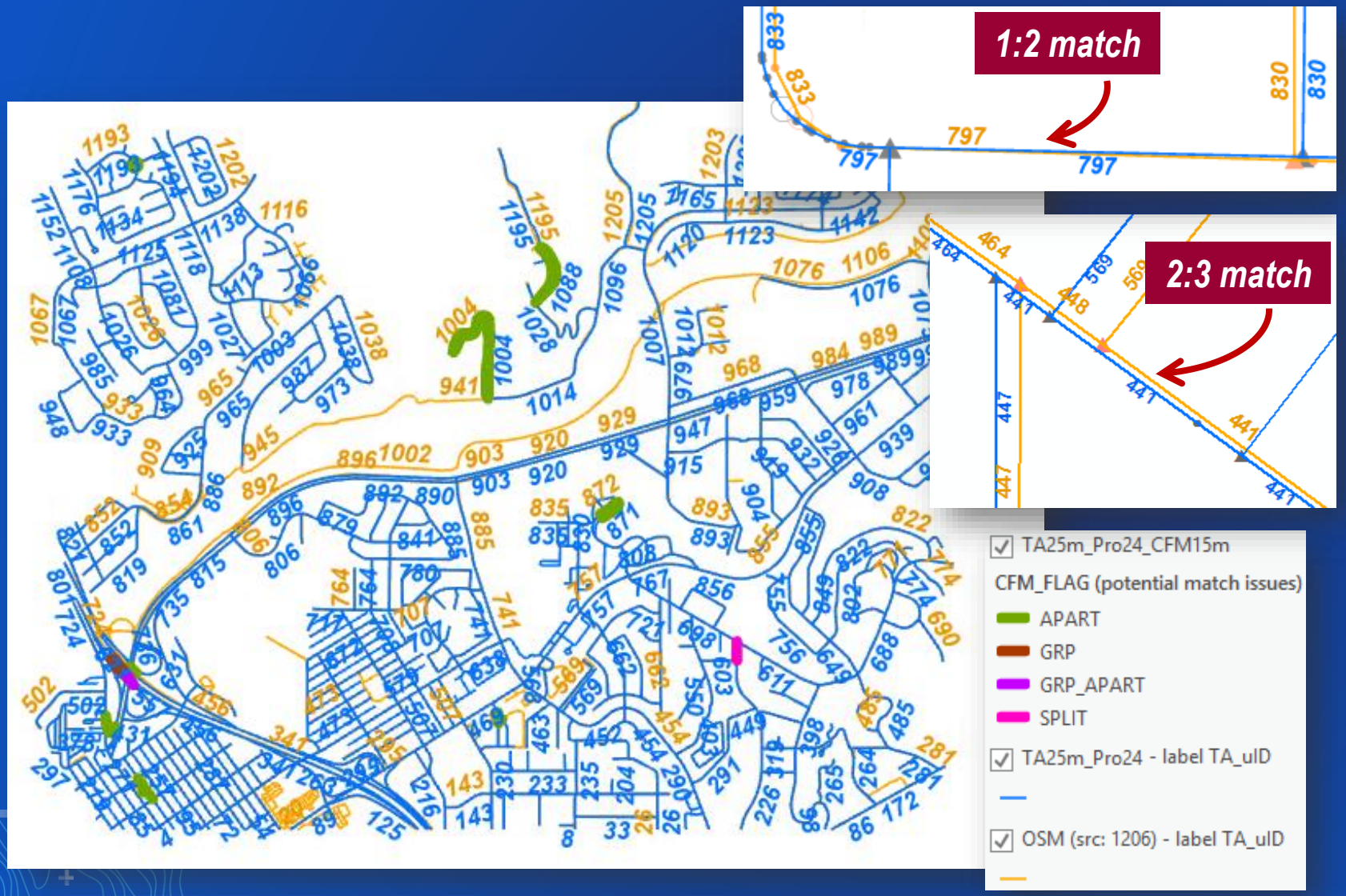
Automated TA and Evaluation

TA_uID is transferred; potential issues are flagged for review





TA result and flags on potential issues



To review

- **Potential wrong transfers**
- **Potential missed transfers**
- ***The m:n transfers noted in srcM_inMN (optional)***

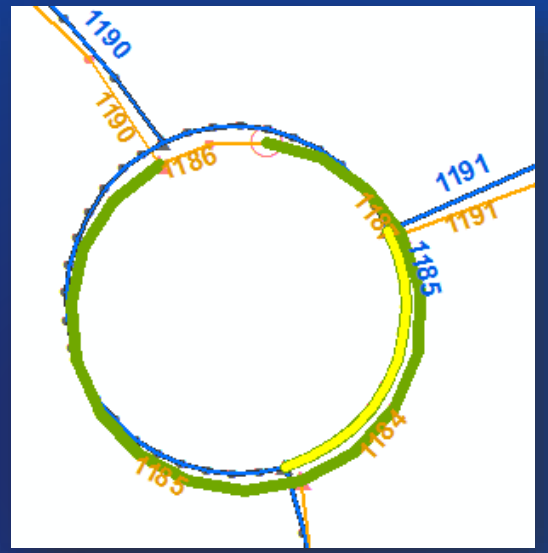
Let's see the review process…

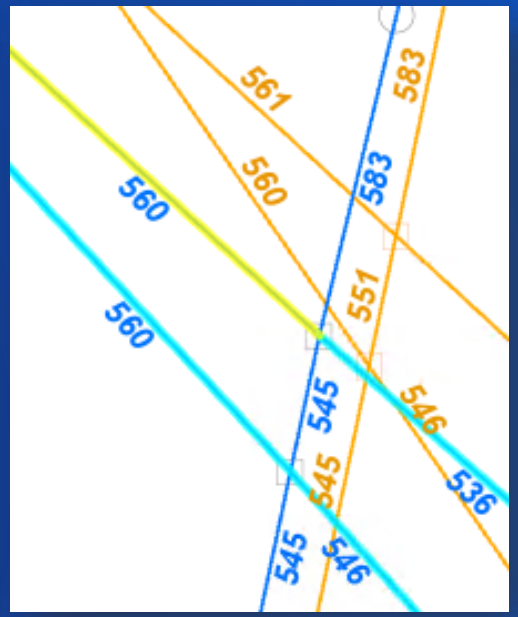
Review transfers with potential match issues

CFM_GRP >=0

19 records were reviewed:

- **5 TA_uID values were corrected**





TA25m_Pro24

OBJECTID	Shape_Length	TA_uID	NEAR_FID	NEAR_DIST	FM_GRP	FM_MN	srcM_inMN	CFM_GRP	CFM_FLAG	CFM_DIST	srcORIGFID	srcNearFID	srcNearDist	REV_NOT	rev_TA_uID	final_TA_uID
147	24.825103	536	<Null>	<Null>	462	2:1	2	3	GRP_APART	21.839679	552	147	6.838031	Wrong	552	<Null>
402	90.700884	627	<Null>	<Null>	460	1:1	1	1	GRP	-1	<Null>	<Null>	<Null>	Wrong	561	<Null>
403	25.757369	560	<Null>	<Null>	461	1:1	1	2	GRP	-1	561	403	7.169561	Wrong	561	<Null>
515	90.268456	627	<Null>	<Null>	455	1:1	1	1	GRP	-1	<Null>	<Null>	<Null>	Wrong	560	<Null>
664	32.981175	1185	<Null>	<Null>	884	3:1	3	14	APART	17.71337	<Null>	<Null>	<Null>	Wrong	1184	<Null>
148	21.253369	536	<Null>	<Null>	458	1:2	1	3	GRP	-1	<Null>	<Null>	<Null>	Correct	<Null>	<Null>
149	10.856186	536	<Null>	<Null>	458	1:2	1	3	GRP	-1	<Null>	<Null>	<Null>	Correct	<Null>	<Null>
275	23.797976	585	<Null>	<Null>	234	2:1	2	5	APART	18.11008	<Null>	<Null>	<Null>	Correct	<Null>	<Null>

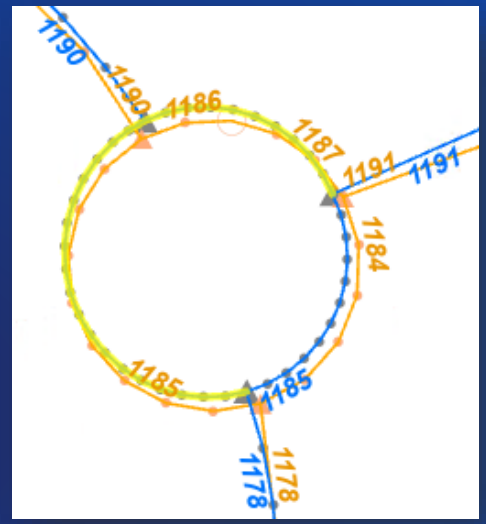
19 of 1,128 selected Filters: 78 %

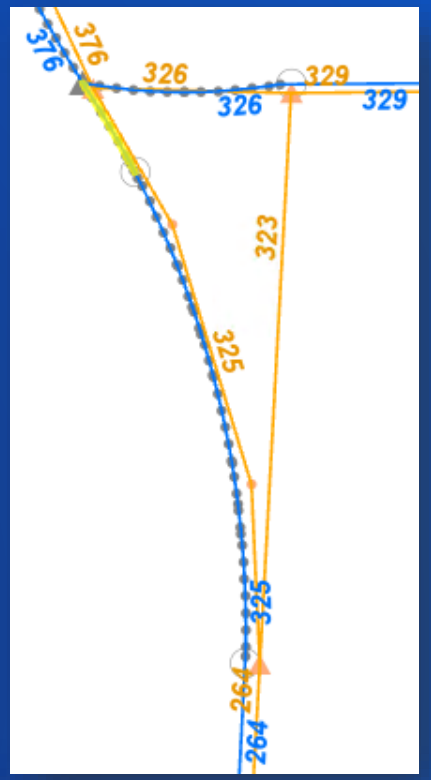
Review potential missed transfers

TA_uID IS NULL AND (NEAR_FID >0 OR srcNearFID >0)

48 records were reviewed:

- **7 transferred TA_uID values were incorrect**





TA25m_Pro24

OBJECTID	Shape_Length	TA_uID	NEAR_FID	NEAR_DIST	FM_GRP	FM_MN	srcM_inMN	CFM_GRP	CFM_FLAG	CFM_DIST	srcORIGFID	srcNearFID	srcNearDist	REV_NOT	rev_TA_uID	final_TA_uID
142	42.646716	<Null>	599	16.108438	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	599	<Null>
401	40.03363	<Null>	731	9.641416	-1	N/A	-1	<Null>	<Null>	<Null>	731	401	9.630519	Wrong	731	<Null>
421	57.363602	<Null>	1185	0.214332	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	1185	<Null>
625	11.066227	<Null>	325	0.525794	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	325	<Null>
816	30.400859	<Null>	1187	1.463928	-1	N/A	-1	<Null>	<Null>	<Null>	1186	816	1.876831	Wrong	1187	<Null>
956	43.071085	<Null>	726	5.306041	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	726	<Null>
1018	0.069553	<Null>	850	0.02129	-1	N/A	-1	<Null>	<Null>	<Null>	850	1018	3.947928	Wrong	861	<Null>
1	10.550205	<Null>	415	23.914806	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Correct	<Null>	<Null>
58	120.644704	<Null>	-1	-1	-1	N/A	-1	<Null>	<Null>	<Null>	559	58	23.366778	Correct	<Null>	<Null>

48 of 1,128 selected Filters: 78 %

Attribute transfer accuracy estimates

- **Total target features: 1128**
- **Incorrect: 12**
- **Correct: 1116**

Accuracy = 1116 / 1128 => 98.9%

- **Remaining no transfer features (TA_uID IS NULL AND REV_FLAG IS NULL): 80**

All correct

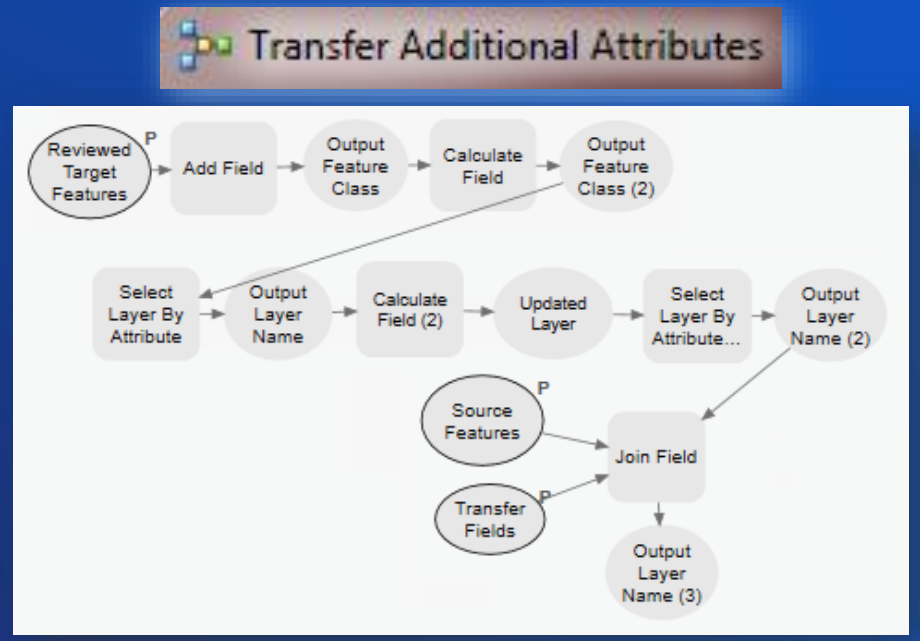
Post transfer if necessary

Merge TA result with edits

- Add a field final_TA_uID and Calculate final_TA_uID = TA_uID
- Select REV_NOTE = ‘Wrong’ and calculate final_TA_uID = rev_TA_uID

Transfer additional attributes via Join Field

- The common fields are: source TA_uID and target final_TA_uID



OBJECTID	TA_uID	NEAR_FID	NEAR_DIST	FM_GRP	FM_MN	srcM_inMN	CFM_GRP	CFM_FLAG	CFM_DIST	srcORIGFID	srcNearFID	srcNearDist	REV_NOTE	rev_TA_uID	final_TA_uID	name
816	<Null>	1187	1.463928	-1	N/A	-1	<Null>	<Null>	<Null>	1186	816	1.876831	Wrong	1187	1187	
421	<Null>	1185	0.214332	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	1185	1185	
664	1185	<Null>	<Null>	884	3:1	3	14	APART	17.71337	<Null>	<Null>	<Null>	Wrong	1184	1184	
1018	<Null>	850	0.02129	-1	N/A	-1	<Null>	<Null>	<Null>	850	1018	3.947928	Wrong	861	861	Capistrano Drive
401	<Null>	731	9.641416	-1	N/A	-1	<Null>	<Null>	<Null>	731	401	9.630519	Wrong	731	731	San Diego Freeway
956	<Null>	726	5.306041	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	726	726	San Diego Freeway
142	<Null>	599	16.108438	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	599	599	San Diego Freeway
402	627	<Null>	<Null>	460	1:1	1	1	GRP	-1	<Null>	<Null>	<Null>	Wrong	561	561	San Diego Freeway
403	560	<Null>	<Null>	461	1:1	1	2	GRP	-1	561	403	7.169561	Wrong	561	561	San Diego Freeway
515	627	<Null>	<Null>	455	1:1	1	1	GRP	-1	<Null>	<Null>	<Null>	Wrong	560	560	
147	536	<Null>	<Null>	462	2:1	2	3	GRP_APART	21.839679	552	147	6.838031	Wrong	552	552	San Diego Freeway
625	<Null>	325	0.525794	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Wrong	325	325	Mac Donald Street
1	<Null>	415	23.914806	-1	N/A	-1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	Correct	<Null>	<Null>	<Null>
2	543	<Null>	<Null>	8	1:1	1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	543	San Diego Freeway
3	534	<Null>	<Null>	7	1:2	1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	534	San Diego Freeway
4	534	<Null>	<Null>	7	1:2	1	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	<Null>	534	San Diego Freeway

12 of 1,128 selected Filters: 76 %

User Stories and Future Work

Thanks to:

- Department of Public Works (DPW), Los Angeles County, USA.
- Institut Cartogràfic i Geològic de Catalunya (ICGC), Barcelona, Spain.
- Kevin Hunt, New York State Department of Transportation, USA.
- Richard Fairhurst, Riverside County Transportation and Land Management) CA, USA RCTLMA,
- National Institute for Water and Atmospheric Research (NIWA) and Land Information New Zealand (LINZ) - Crown Copyright Reserved.
- Resource Management Service, LLC, Birmingham, AL, USA.
- All others who supported us along the way.

User story 1: Enhancing county roads by spatially more accurate city roads

County centerline attributes and direction must be retained.

Original_county_centerlines
Temecula_city_centerlines

Updated_county_centerlines
Original_county_centerlines

Updated_county_centerlines
Temecula_city_centerlines

- Use DFC to find matching features and line direction differences
  - For 1:1 matches, flip city centerlines of opposite direction (Flip Line)
  - For m:n matches, merge/split city or county centerlines to get 1:1 matching segments, recalculate address ranges for county roads as needed, and flip city centerlines of opposite direction (tools + scripts)
- Transfer city centerline geometry to county centerlines (script)

~ 98%+ accuracy

Data/information source: RCTLMA (Riverside County Transportation and Land Management) CA, USA

Acknowledgement: Thanks to Richard Fairhurst, for providing the information and screenshots.

User story 2: Combining electoral roads and topographic roads

There is no “most accurate” dataset.

Electoral roads

Topographic roads

Conflation workflow

Land Information New Zealand
Toitū te whenua

Feature matching
- many to many (ArcGIS tool)

~ 90% accuracy

Segmentation
(script)

Feature matching
& Spatial Differencing
- one to one (ArcGIS tool)

~ 99% accuracy

Select best spatial alignments against imagery
(custom plugin)

Merge datasets
- snap, rubbersheet, merge attributes (script)

Information source: Land Information New Zealand (LINZ)

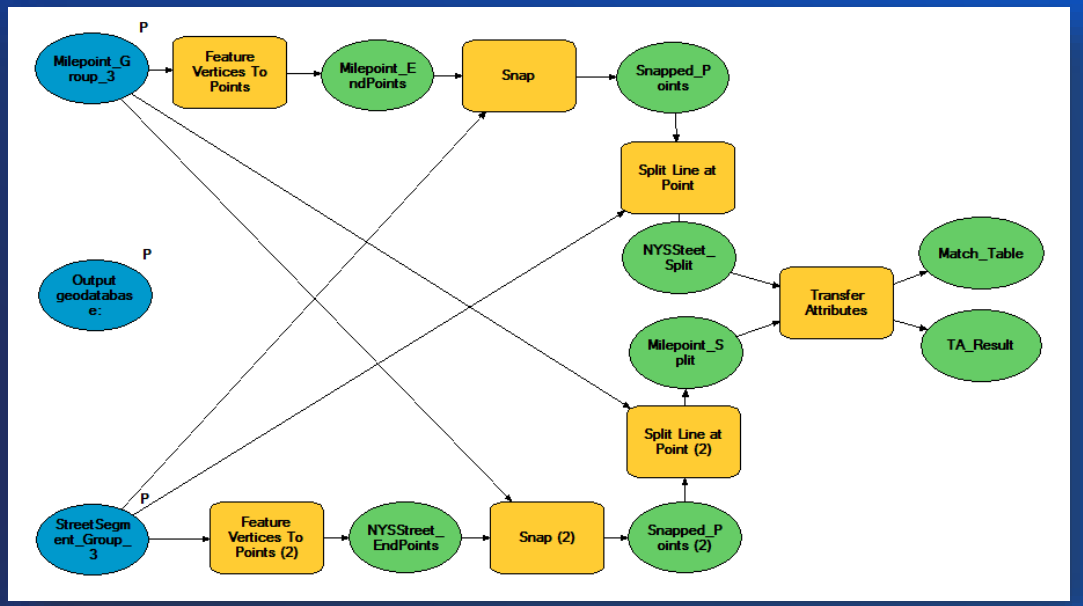
Acknowledgement: Thanks to Douglas Kwan, LINZ, for providing the information.

User story 3: Transferring attributes from State routes to Street segments

Segmentation for the datasets was different

State routes

Street Segments



- **State Routes and Street segments were split by end points to provide a more similar segmentation between the two datasets.**

~99.5% matching rate

Data/information source: NYSDOT, USA

Acknowledgement: Thanks to Kevin Hunt, for giving us the opportunity to work with him and share his data.

Conflation is a necessity and goes a long way

It can be done more efficiently now

- Customize workflows for your scenarios

It improves data quality and usability

- It adds value to your data; time is worth spending

You can better collaborate with the communities

- Sharing up-to-date and consistent data
- Performing reliable analysis and quality mapping

Please send us your feedback and share your stories … dlee@esri..com ☺

Future work

New tools and enhancements

- Improve pattern recognition and feature matching
- Further enhance outputs
- Develop new tools to support more use cases and other feature types

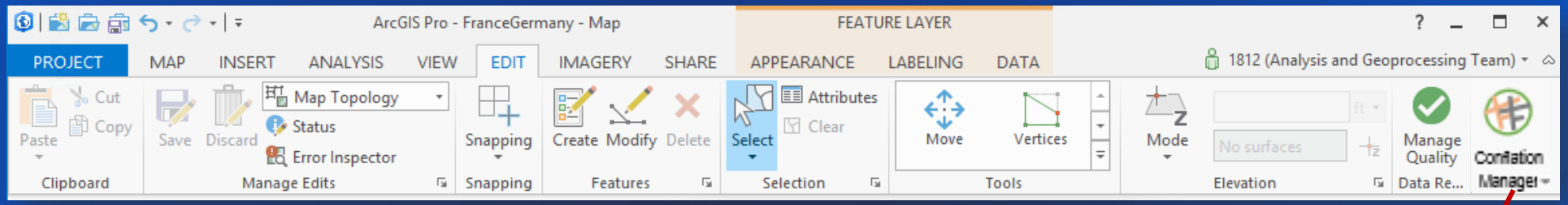
Integrated processing and inspection system

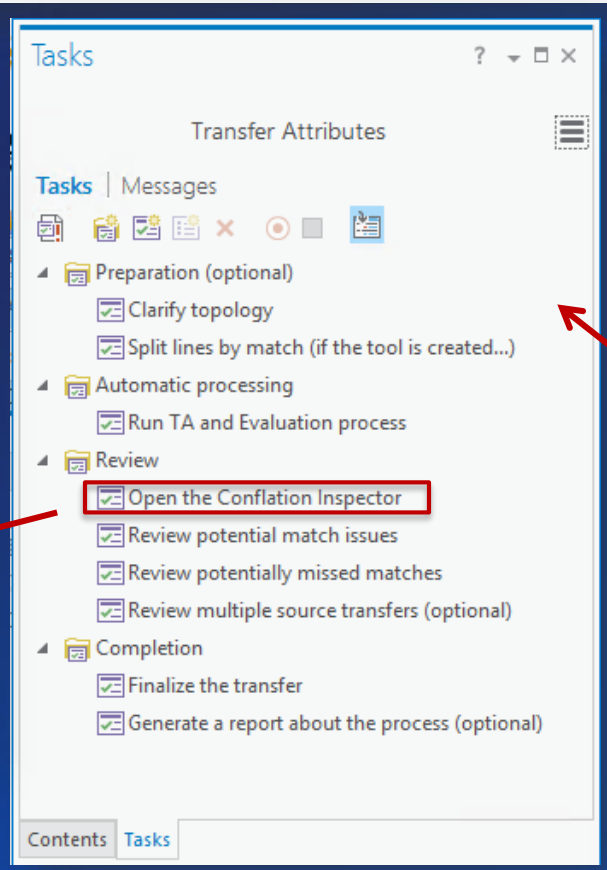
- Tasks based Conflation Manager for ArcGIS Pro

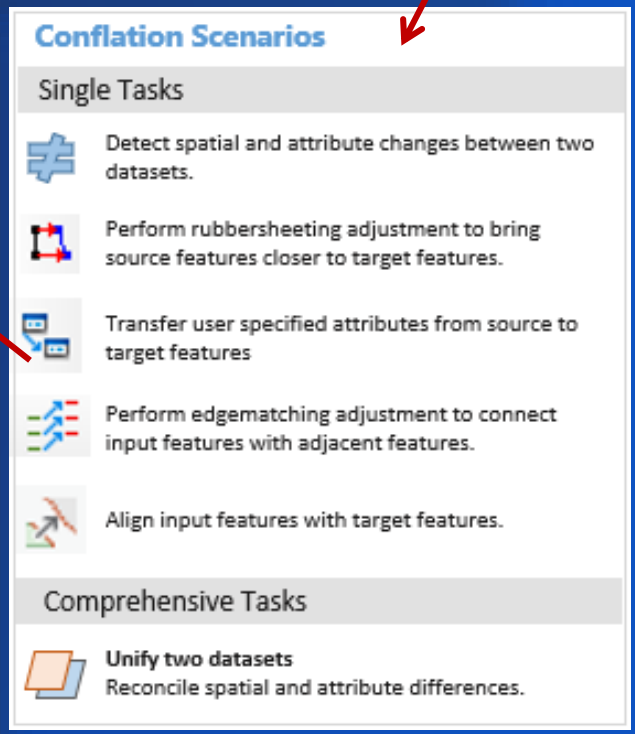
Formalization of workflows

- Focus on common needs (e.g. multi-scale data updating and linking features)
- Incorporate other data sources (imagery, lidar, GPS)
- Research on contextual conflation (spatially related features)

Conflation Manager (ConfMgr)

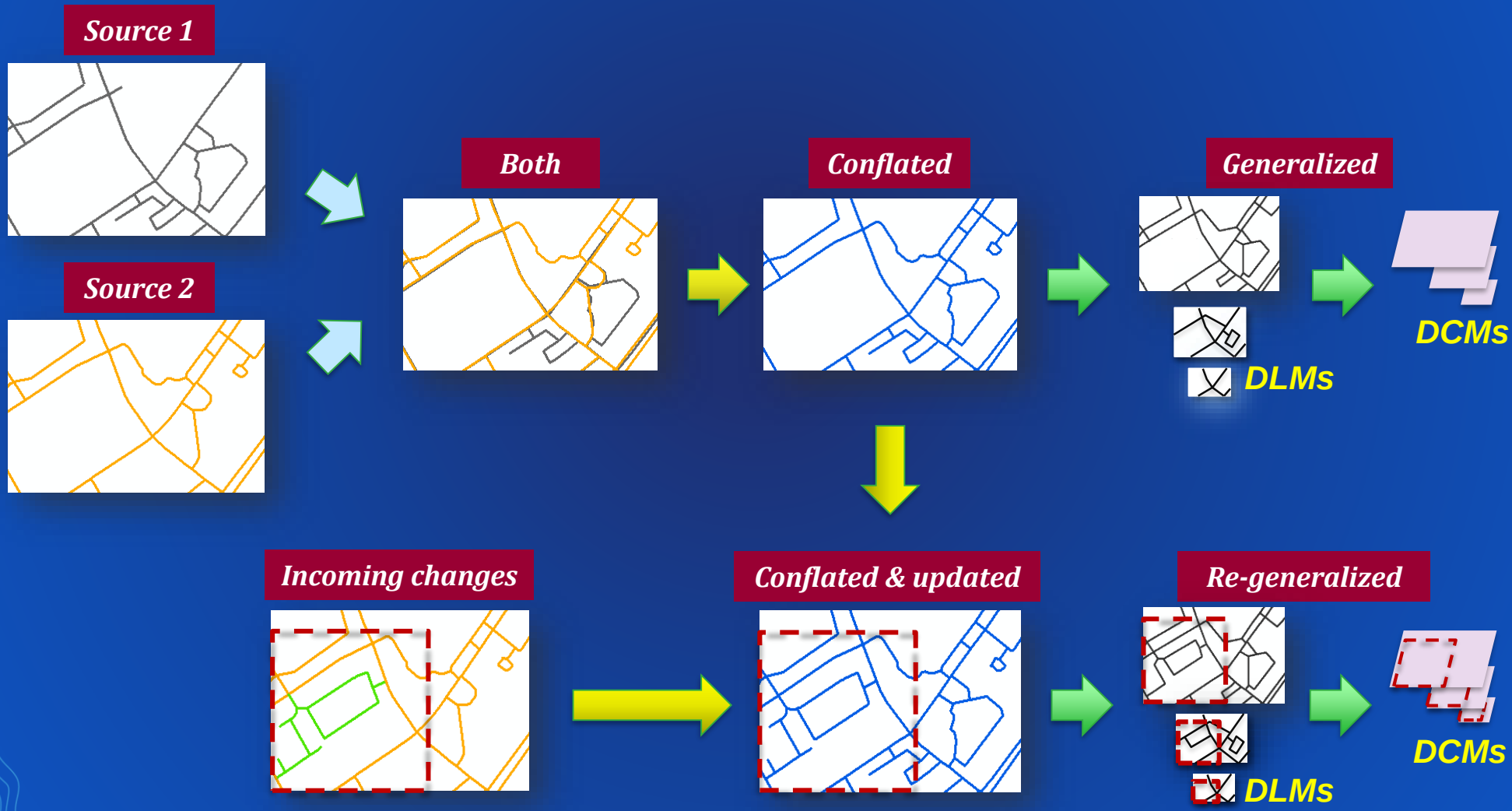




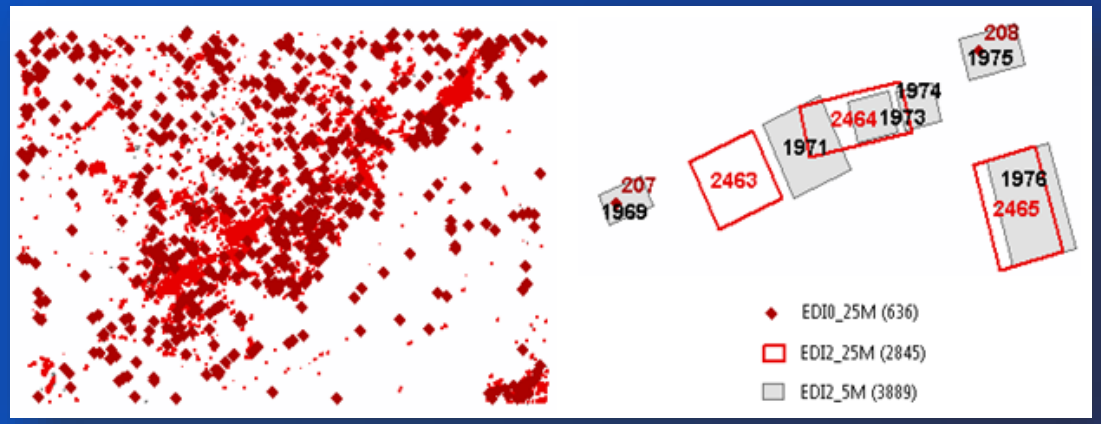


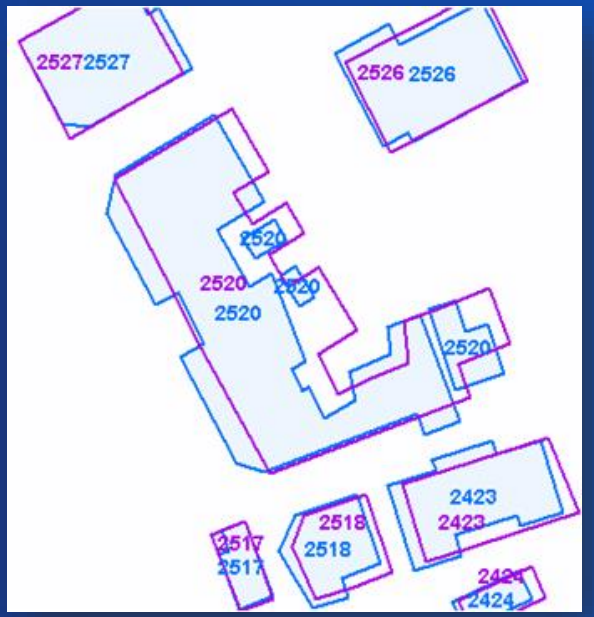
Conflation in multi-scale data updating and mapping

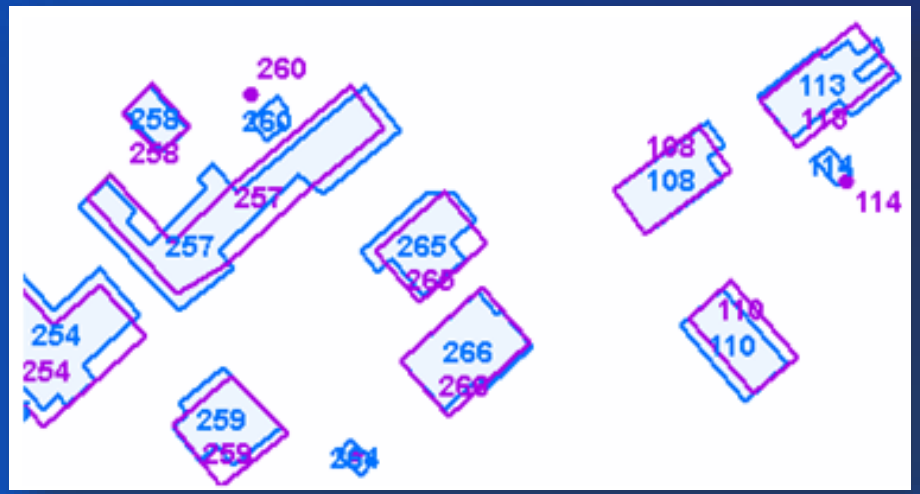
DLM – digital landscape model; DCM – digital cartographic model

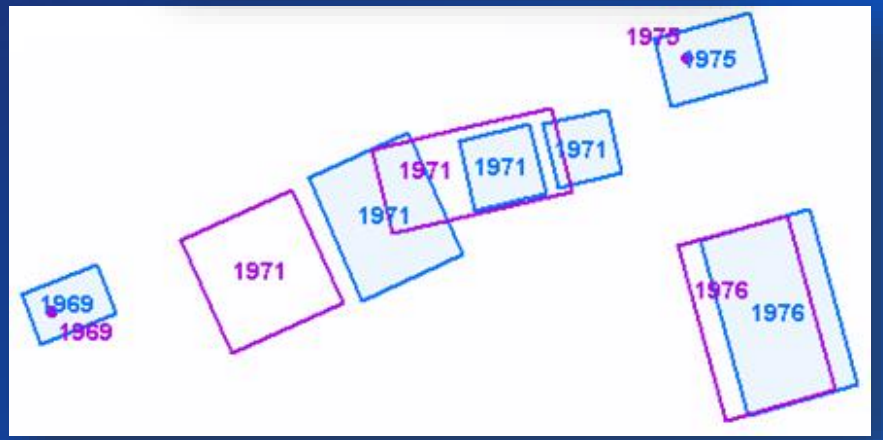


Example of linking features: ICGC 1:5k and 1:25k buildings







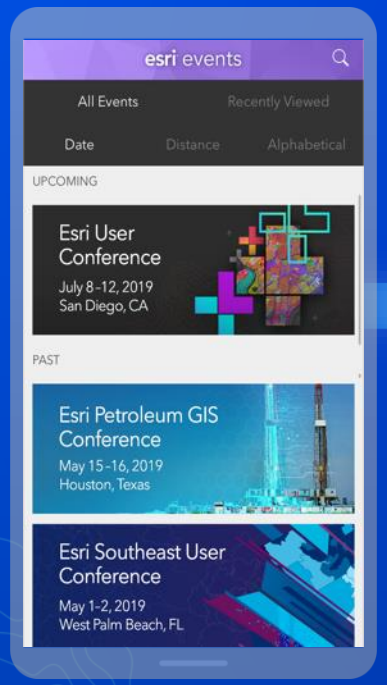


This session will repeat at
4:00pm – 5:00pm, Wednesday
Room 08

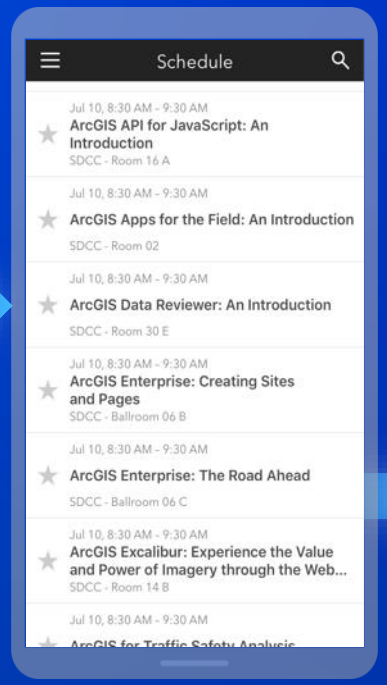
Conflation: Edgematching Tools and Workflows
12:15pm – 1:00pm, Tuesday
Demo Theater 10

Please take our survey and share your feedback on the Esri Events App!

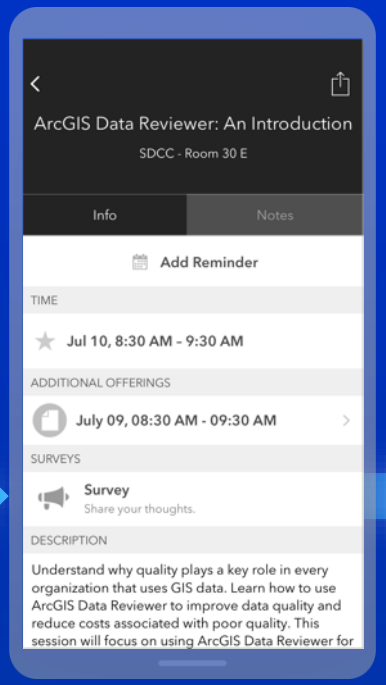
Download the Esri Events app and find your event



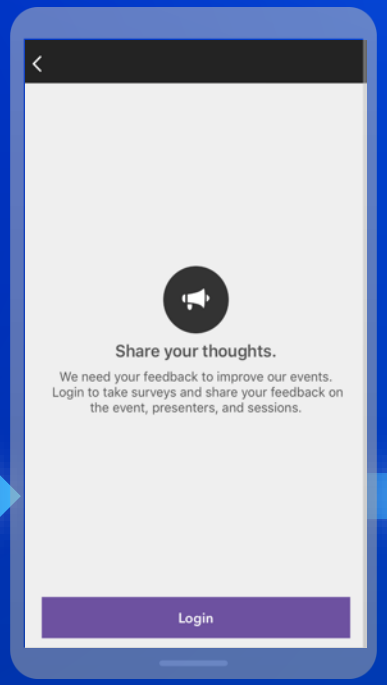
Select the session you attended



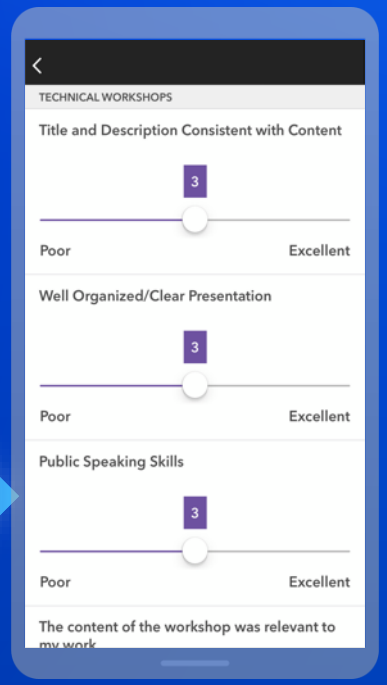
Scroll down to "Survey"



Log in to access the survey



Complete the survey and select "Submit"



Thank you for attending!

Questions & Answers

esri | THE SCIENCE OF WHERE