Final Report

LinkWISCONSIN Address Point and Parcel Mapping Project

September 2014

CONTENTS

LIST OF FIGURES

LIST OF MAPS

Published September 30, 2014

Printed by UW-Madison Print Services

CONTACT

Public Service Commission of Wisconsin
LinkWISCONSIN Initiative
610 N. Whitney Way
Madison, WI 53707
608-267-1204
PSCStateBroadbandOffice@wisconsin.gov
http://psc.wi.gov/

Wisconsin State Cartographer's Office
384 Science Hall
550 N. Park Street
Madison, WI 53706-1491
608-262-3065
sco@wisc.edu
www.sco.wisc.eu

Wisconsin Department of Administration
Wisconsin Land Information Program
101 East Wilson Street
Madison, WI 53703
608-267-3369
WLIP@wisconsin.gov
www.doa.state.wi.us/WLIP

LinkWISCONSIN Address Point and Parcel Mapping Project

In 2013, the Public Service Commission of Wisconsin (PSCW) approved the LinkWISCONSIN Address Point and Parcel Mapping Project, a collaboration between the State Cartographer's Office and the Wisconsin Land Information Program. This document describes the project, which ran from July 2013 to September 2014 with $168,000 in federal grant funding from the National Telecommunications and Information Administration provided by the PSCW.

Primary Project Objective

Establish statewide address point and parcel GIS map layers by integrating county-level datasets and improve a database of Community Anchor Institutions

This report documents the process to create statewide GIS (geographic information system) map layers in Wisconsin. Built from all known county and municipal data, the final address point layer totaled 2.7 million address points and the final parcel layer totaled 3.7 million parcels, amounting to statewide address point and parcel coverage that is unprecedented within the State of Wisconsin.

PROJECT CHARACTERISTICS

The team looked at case studies from other states, an internal pilot project, formal data standards, and the needs of the PSCW in developing a data model.

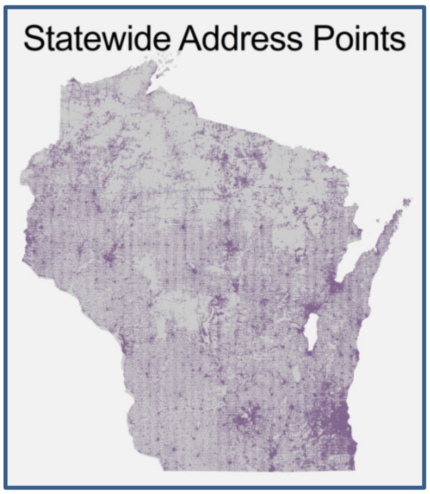
Technical implementation was executed in the Esri ArcGIS environment through a combination of data interoperability tools, ArcGIS tools, and custom Python tools.

DATA METRICS & ANALYSIS

Address Points. Due to variation in attribute information and lack of a parcel identification number in many county address point layers, the team spent a considerable amount of time parsing the data and standardizing each local-level dataset.

Gaps in Local Data. Gaps in county address point and parcel datasets were identified, and municipalities were called upon to fill gaps wherever possible. Parcel centroids were substituted for address points when that was the only option. However, there are some places where neither address point nor parcel data exist in digital form and thus remain as holes in the statewide address point layer.

Community Anchor Institutions. CAIs consist of public schools K-12, libraries, health care providers, public safety entities, higher education institutions, and other community support organizations. This project improved the spatial accuracy of Wisconsin's CAI map layer and expanded it to include correctional facilities, ports, and private schools.



CHALLENGES & LIMITATIONS

Data Sharing. As an example of cooperative data sharing with local governments, the project featured outreach and education measures that resulted in participation by 100% of known local government data stewards. However, five counties requiring license agreements, one a small fee, and the fact that address point datasets are often not stewarded by the county's land information officer prolonged data acquisition.

Technical Hurdles. A variety of technical hurdles were encountered, such as a lack of metadata in local level datasets, glitches in technical tools, and the massive size of the statewide datasets.

BEST PRACTICES & RECOMMENDATIONS

Repeatability. The processes to collect local data and create the statewide layers were documented in detail, including the storage of some technical processes in model form. Though repeatability was a focus, this does not mean a future project would be fully automated, due to variation in local datasets.

The successful experience of this project will benefit the Version 1 Statewide Parcel Map Database Project, another effort by the State Cartographer's Office and Wisconsin Land Information Program that will apply lessons learned from this project in order to create a statewide parcel layer that will be shared with the PSCW and the public.

County Data Reports. Data assessment and observation reports were created for each individual county in order to provide feedback on local data, an example of which appears in Appendix C.

1 INTRODUCTION

1.1 About the LinkWISCONSIN Address Point and Parcel Mapping Project

In 2009, the Public Service Commission of Wisconsin (PSCW) launched a statewide broadband initiative called LinkWISCONSIN. The goals of the initiative included mapping broadband service within the state, developing a vision for broadband in Wisconsin, and working to extend broadband access to underserved areas.

The initiative included the creation of an interactive map to track broadband availability, speed, and providers. In an effort to add capabilities and value to broadband mapping, the PSCW approved the LinkWISCONSIN Address Point and Parcel Mapping Project, a collaboration between the State Cartographer's Office (SCO) and the Wisconsin Land Information Program (WLIP) at the Department of Administration (DOA) to develop statewide address point and parcel layers based on county data, as well as the creation of an improved database of what are known as Community Anchor Institutions (CAIs).

This document describes the LinkWISCONSIN Address Point and Parcel Mapping Project, which ran from July 2013 through September 2014 with funding provided by the PSCW through the American Recovery and Reinvestment Act.

Built from all known existing county and municipal data, the final address point layer totaled 2.7 million address points and the parcel layer totaled 3.7 million parcels. Both layers received 100% participation from county- and municipal-level data stewards.

1.1.1 Objectives

- Establish statewide address point and parcel GIS map layers for the PSCW by integrating county-level and municipal datasets
- Establish a Community Anchor Institution layer that improves spatial accuracy, improves harmonization with the statewide parcel layer and PSCW broadband survey tables, and expands the database of CAIs
- Build on the experience of 2012 Wisconsin Land Information Association parcel mapping demonstration project
- Provide training and technical assistance to counties for statewide geospatial data integration
- Facilitate and refine processes for data sharing between counties and state agencies
- Report on the analysis of county address and parcel datasets, processes for integrating datasets at county boundaries, and lessons learned

1.1.2 What Are Address Points and Parcels?

In the simplest terms, address points give the location of a building or property, while parcels represent property boundary lines and associated ownership information, as illustrated in Fig. 1.

Address point and parcel map layers are two of the most highly demanded geospatial layers within any jurisdiction because of their analytical, planning, and geo-relational potential. Across the Unites States, they are also becoming increasingly important as aggregated statewide layers. According to the National States Geographic Information Council (NSGIC), 15 states had statewide parcel data layers at 96-100% completion as of September 2013.[1] Wisconsin is not listed as one of them.

[1] http://www.nsgic.org/gma-2013/index.php?question_index=77

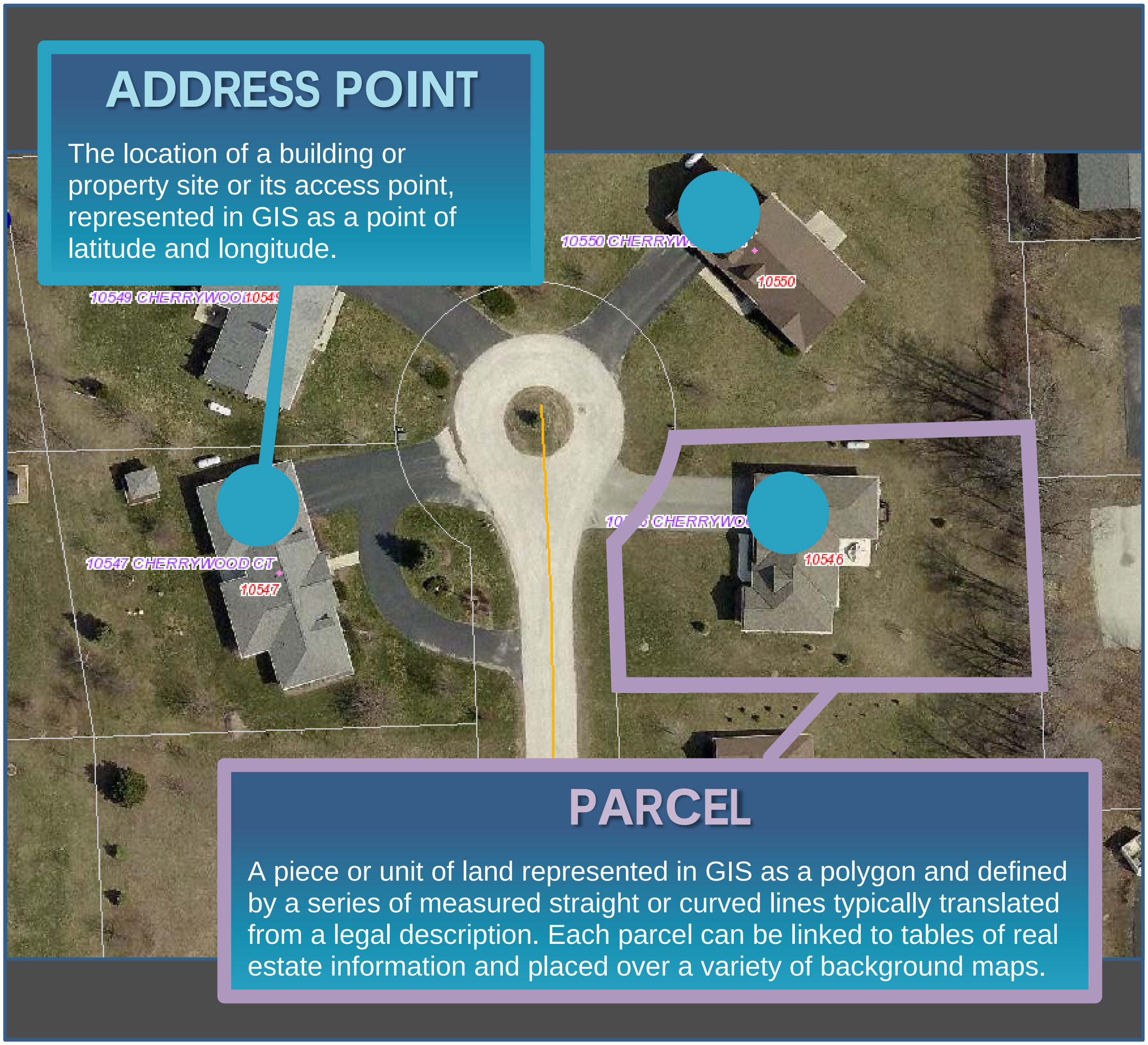


Figure 1. Address Point and Parcel Definitions

Sample Address Point Data	
AddNum	10546
Prefix	W
Street Name	Cherrywood
Street Type	Court
PlaceName	Chippewa Falls
State	WI
ZipCode	54729
Associated Parcel ID	1810005313111W1
Latitude	44.934138
Longitude	-91.400001

Sample Parcel Data	
County	Chippewa
Municipality	Chippewa Falls
Parcel ID	1810005313111W1
Property Address	10546 Cherrywood Court
Acreage	0.804
Assessment Class	Residential
Land Value	$70,000.00
Improvements	$150,000.00
Total Value	$220,000.00
Total Property Tax	$2,256.00

1.1.3 Accountability and Transparency

A Memorandum of Understanding (MOU) between PSCW and DOA defined responsibilities in the project, including the milestones, timeline, and budget. PSCW provided $168,000 in federal grant funding for the project. DOA subcontracted for technical work and some aspects of project coordination with SCO. In addition to quarterly progress reports submitted to the PSCW, the MOU's list of interim deliverables acted as checkpoints between the PSCW and the project team.

1.1.4 Project Team

LinkWISCONSIN Address Point and Parcel Mapping Project	
Project Team	
Howard Veregin, Project Co-Lead	Wisconsin State Cartographer's Office
Peter Herreid, Project Co-Lead	Wisconsin Department of Administration
Codie See, Project Coordinator	Wisconsin State Cartographer's Office
David Vogel, GIS Specialist	Wisconsin State Cartographer's Office
AJ Wortley	Wisconsin State Cartographer's Office
Brenda Hemstead	Wisconsin State Cartographer's Office
Jim Lacy	Wisconsin State Cartographer's Office
Angela Limbach	Wisconsin State Cartographer's Office
Caitlin Wolters	Wisconsin State Cartographer's Office
Patrick Donahue	Wisconsin State Cartographer's Office
Samuel Schumacher	Wisconsin State Cartographer's Office
Davita Veselenak	Wisconsin Department of Administration
Advisory Team	
Justin Conner	Wood County
William Cozzens	Waukesha County
Jeff DuMez	Brown County
Scott Galetka	Bayfield County
Martin Goettl	University of Wisconsin-Eau Claire
Ian Grasshoff	Waupaca County
PSCW Team	
Tithi Chattopadhyay	State Broadband Director
Colter Sikora	Broadband Mapping Coordinator
Matthew Noone	Broadband Mapping Architect

Statewide Address Points and Parcels Background

Some state agencies did work to aggregate Wisconsin's parcel information in the past, but these were disconnected efforts, none of which resulted in a comprehensive statewide layer. Other individual efforts have also lacked analysis of the statewide layer and documentation pertaining to the process used to create them.

In 2012, the Wisconsin Land Information Association (WLIA) Board identified the need to demonstrate the value of a statewide parcel dataset. Four staff members from county land information offices decided to cooperate on a project—Justin Conner, William Cozzens, Jeff DuMez, and Ian Grasshoff. They believed that in order to achieve results they would need to "just do it" and ignore many of common obstacles of a true statewide seamless parcel dataset.

The objective of the demonstration project was to gather as much parcel data as possible and assemble it into a seamless parcel dataset, stored as a single common database schema. The team did not deal with many of common obstacles to statewide layer creation, such as data sharing, license agreements, cost of data acquisition, and edge-matching. Thus, the resulting parcel layer was retired in late 2013. However, the demonstration project was successful as a proof-of-concept. In order to get the full benefit of their experience, those WLIA Parcel Team members were invited to serve on the advisory team for the LinkWISCONSIN Address Point and Parcel Mapping Project.

This project did not utilize the data from the WLIA demonstration project due to its age, data sharing restrictions, and lack of focus on address points. However, the LinkWISCONSIN Address Point and Parcel Mapping Project team built upon the experience in many ways, as illustrated in the table below.

Building From "Just Do It"	
WLIA Parcel Project	**This Project**
Parcels	Address points and parcels
Informal; License agreements and fees not required	License agreements and fees were negotiated
Not formally documented	Comprehensive report as part of project deliverable
Gaps in data unresolved	Action taken to resolve gaps
Rudimentary schema	Schema coordinated to retain attribute information
Attribute data in native form	Parsed, cleaned and standardized attributes
Demonstration project to highlight GIS capabilities	Value added to meet requirements and support PSCW business needs

1.1.5 Project Timeline and Milestones

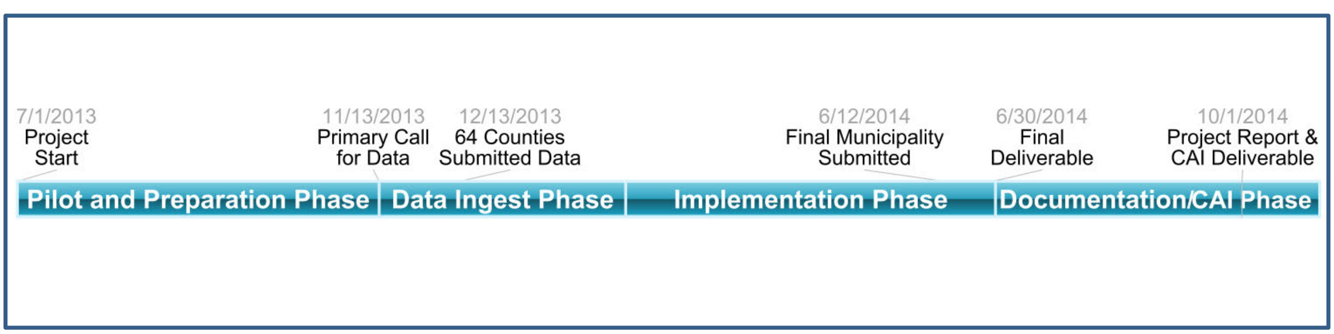


Figure 2. Project Timeline

1.1.6 Utility of Project Deliverables

The goals of the PSCW's LinkWISCONSIN project include mapping broadband service within the state, developing a vision for broadband in Wisconsin, and working to extend broadband access to underserved areas. The LinkWISCONSIN project includes an interactive map (Fig. 3) to track broadband availability, speed, and providers.[2]

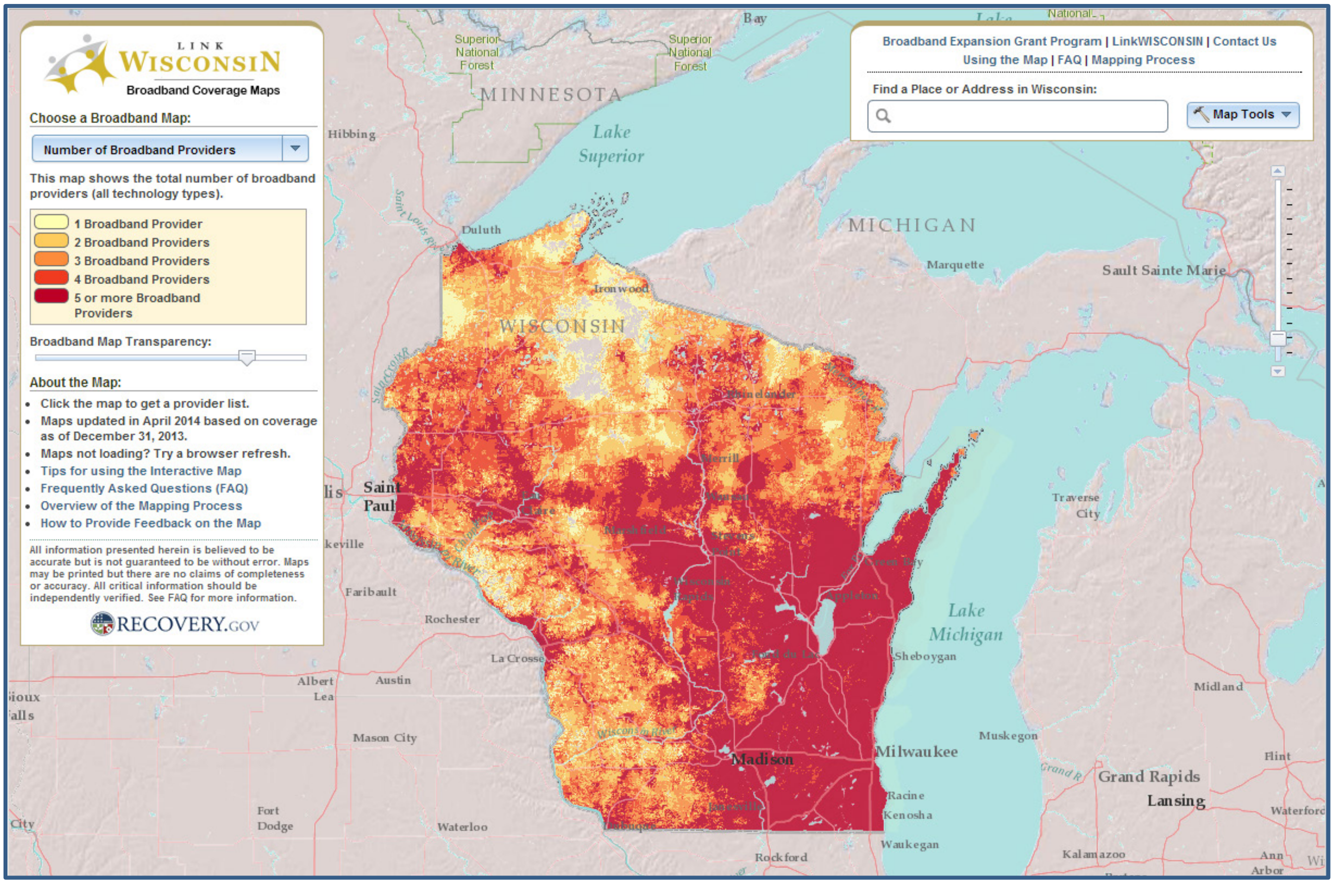


Figure 3. LinkWISCONSIN Broadband Coverage Map

For the LinkWISCONSIN Address Point and Parcel Mapping Project, the objective of developing statewide address and parcel layers based on county address and parcel data was intended to serve the purpose of adding to Wisconsin's broadband build-out resources. The utility of the project deliverables include the ability to use address and parcel data as a planning tool and as a geocoding base for LinkWISCONSIN mapping activities.

While the main focus of this project was on address points, parcel data was included, as it afforded additional capabilities for broadband mapping efforts. First, parcels were used to substitute for address point information where the latter was lacking at the local level. Second, parcels provided a quality check on address points, especially in rural areas. Finally, parcels have the potential to allow for more advanced analyses, such as determination of the spatial relationships between property boundaries and broadband infrastructure.

[2] http://wi.linkamericadata.org/

1.2 Data Collection and Outreach

1.2.1 Call for Data and Data Acquisition

On November 13th, 2013, the project released its primary call for data to each of Wisconsin's 72 county land information offices. The call for data, which was sent via email by the Wisconsin Land Information Program grant administrator, appears below.

Dear Land Information Officer,

On behalf of the Department of Administration, I am writing to request a subset of your GIS data. Data acquired through this request will be used to develop a statewide address point data layer for the LinkWISCONSIN Address Point & Parcel Mapping Project. With your county's participation, this statewide address point layer will play an important role in enhancing the accuracy of the map of broadband service across the state.

What data?

I am requesting the most recent countywide data for these layers at a minimum:

- **Address Points (Structure Points preferred)**
- **Parcels (plus Tax Roll)**

The following reference layers would additionally be useful:

- **Street Centerlines**
- **Right of Ways**
- **Building Footprints**

How to upload?

Submitting data is simple! Our project partner, the State Cartographer's Office (SCO), has created a secure online directory for your files, linked below. Using the upload tool does not require software installation.

- File geodatabase is the preferred format (not required)
- All files must be placed into one zipped (.zip) folder
- Please include the name of your county in the zip file name and in the Description box
- To include tips or caveats about working with your data, submit them as a text file or word document within the main directory of the zipped file
- Please contact me if an alternative method for file upload is needed

http://broadbandmap.sco.wisc.edu/background/upload-data.html

When is it needed?

The deadline to respond is Wednesday, November 27th, but we are ready for your data contributions as soon as possible. For an updated view of county contributions, check out the progress map.

Thank you!

The participation of your county is important to the success of the project and future GIS initiatives. Please feel free to contact me with questions.

Sincerely,

Peter Herreid
Grant Administrator
Wisconsin Land Information Program
608-267-3369

cc: LIO staff

Data Acquisition Process

Sixty-four counties submitted address point and parcel data by December 13, 2013, a month from the original request for data. As the project team identified gaps in address point or parcel layers in the county data submitted, follow-up requests for missing data were made to 11 counties and six municipalities. Yet despite the quick response time from the majority of counties, it took a total of seven months to collect all known county and municipal digital data. The most common reason for extended delays was county license agreements, which are addressed in Chapter 4.

1.2.2 Website and Progress Map

In order to communicate more effectively with stakeholders, the SCO created and administered a Website with information on the project.[3] On the site, readers could learn more about the project, see progress reports and announcements, and learn how and why they should participate. The site included a place where counties could contribute their geospatial data to the project team via a simple upload widget. In addition, it featured a dynamic map that indicated the status of address point data submission by the counties and municipalities.

1.2.3 Education & Training Efforts

Conference presentations, a Webinar, and a workshop were outreach components important to garnering participation in the LinkWISCONSIN Address Point and Parcel Mapping Project. These education and training efforts served to communicate why local data matters to statewide initiatives and provided guidance on how to improve the data and streamline the submittal process.

Conference Presentations	
WLIA Fall Regional Meeting October 2013	Enhanced Broadband Mapping Project
WLIA Annual Conference February 2014	LinkWISCONSIN Address Point and Parcel Mapping: The Story of a Wisconsin Layer
UW Geospatial Summit April 2014	LinkWISCONSIN Address Point and Parcel Mapping: The Story of a Wisconsin Layer
WLIA Spring Regional Meeting June 2014	Panel Discussion on Parcels and Act 20 Requirements with Project Review

Before the initial call for data, the SCO hosted a webinar to provide an overview of the project, the timeline for data collection, and implementation details. In addition, as part of the 2014 Annual WLIA Conference, a workshop was held entitled *Building an Address Repository Using the FGDC Standard: Implementing Data Quality and Data Sharing*. The instructor was Martha Wells, GISP, co-chair of the Urban and Regional Information Systems Association Address Standard (URISA) Working Group.

The URISA-certified workshop helped participants understand the addressing process, practices for managing address data, organizational challenges related to addressing, and the Federal Geographic Data Committee Street Address Data Standard. Twenty-three people attended, including six county land information officers and many local government representatives.

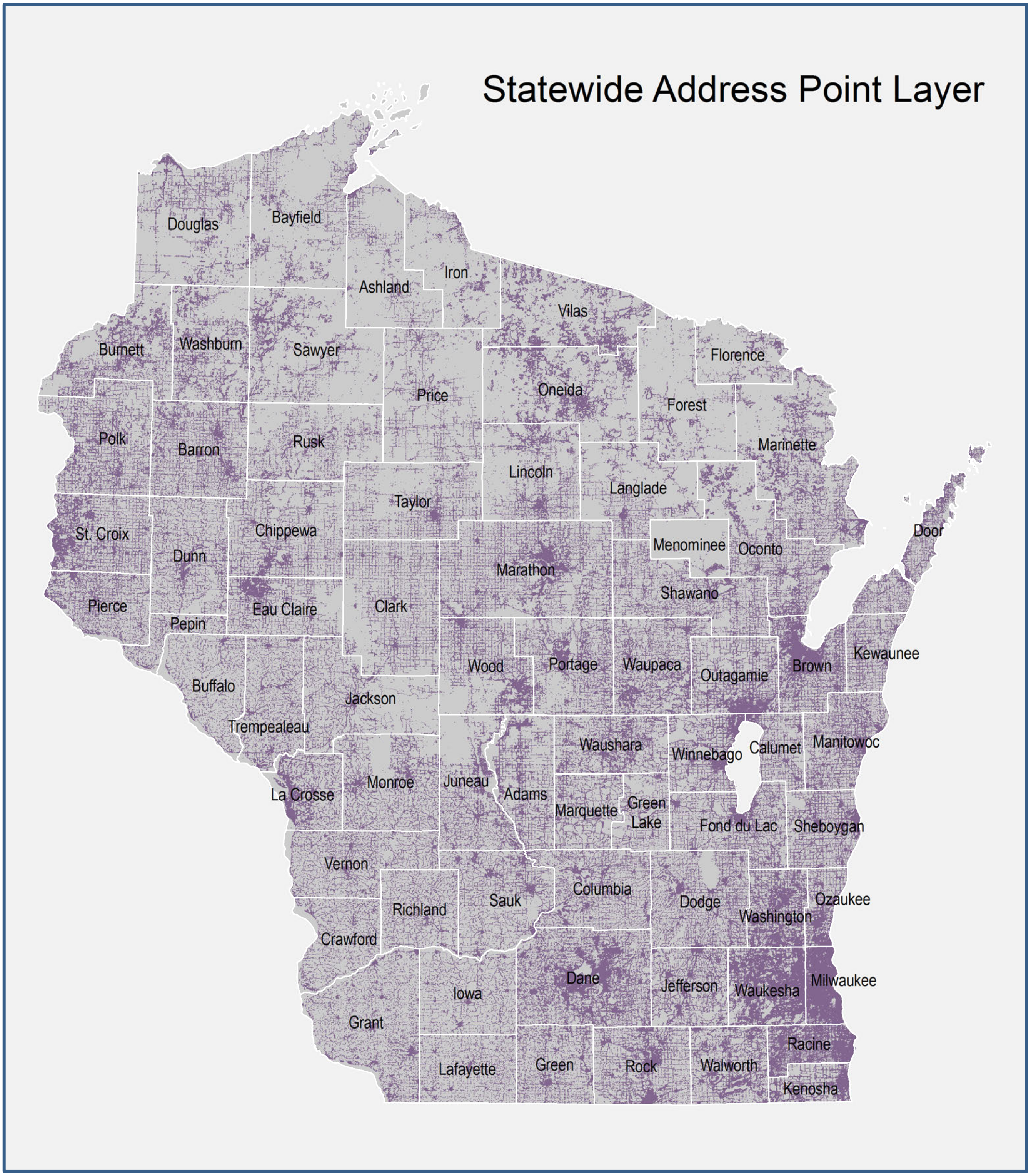
1.3 Final Dataset

The final address point layer totaled 2.7 million address points and the final parcel layer totaled 3.7 million parcels. The list of counties and cities that contributed data to the project appears on the following page.

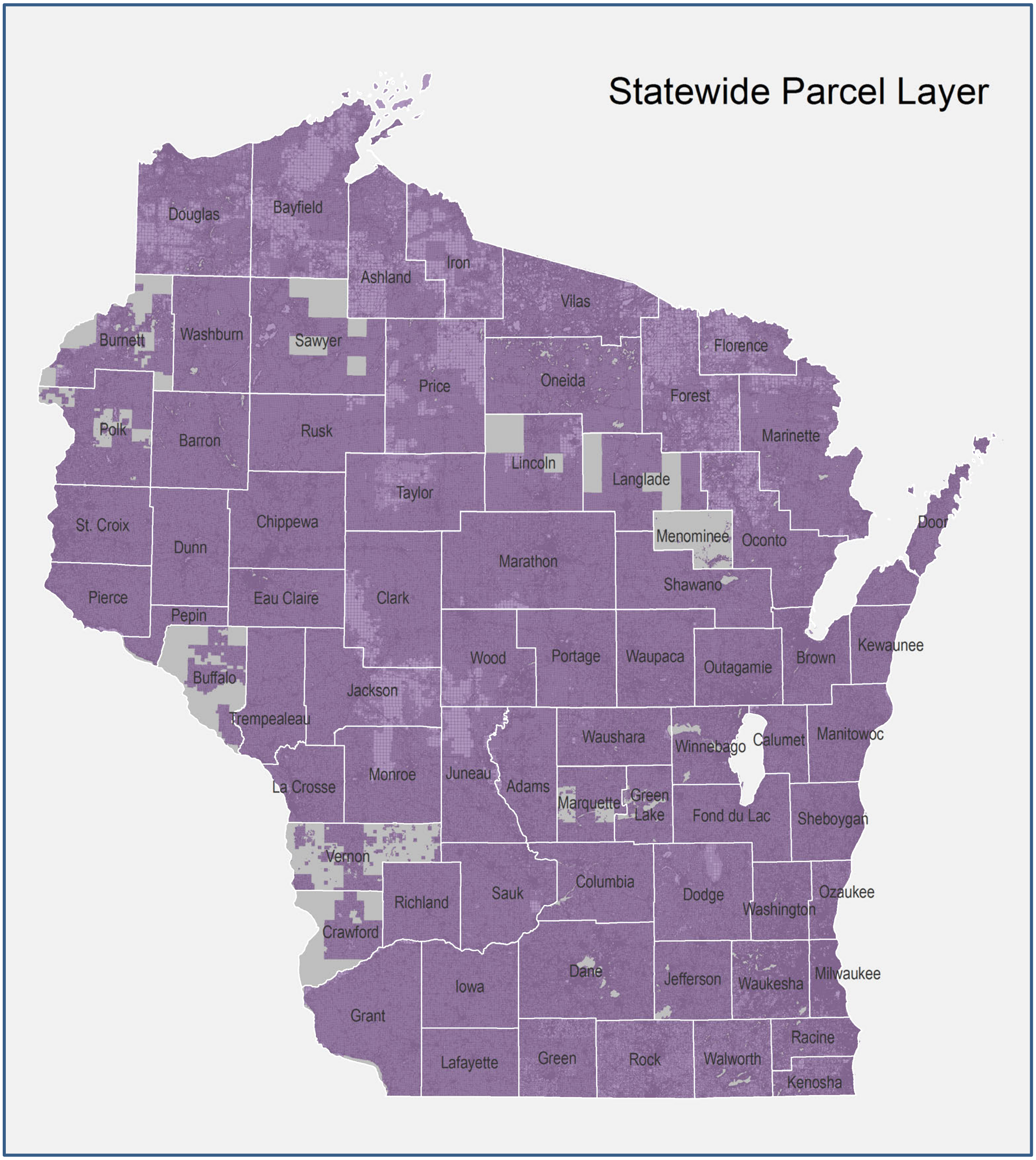
[3] http://broadbandmap.sco.wisc.edu

Address Point and Parcel Data Submittals

County	Address Points	Parcels	Municipal Notes
Adams	Submitted	Submitted	
Ashland	Submitted	Submitted	Building footprints submitted by City of Ashland separately from county
Barron	Submitted	Submitted	
Bayfield	Submitted	Submitted	
Brown	Submitted	Submitted	
Buffalo	Submitted	Submitted	
Burnett	Submitted	Submitted	
Calumet	Submitted	Submitted	
Chippewa	Submitted	Submitted	
Clark	Submitted	Submitted	
Columbia	Submitted	Submitted	
Crawford	Submitted	Submitted	
Dane	Do Not Exist County-wide	Submitted	
Dodge	Submitted	Submitted	
Door	Submitted	Submitted	
Douglas	Submitted	Submitted	
Dunn	Submitted	Submitted	
Eau Claire	Submitted	Submitted	Parcels submitted by City of Eau Claire separately from county
Florence	Submitted	Submitted	
Fond du Lac	Submitted	Submitted	Parcels submitted by City of Fond du Lac separately from county
Forest	Submitted	Submitted	
Grant	Do Not Exist County-wide	Submitted	
Green	Submitted	Submitted	
Green Lake	Submitted	Submitted	
Iowa	Submitted	Submitted	
Iron	Submitted	Submitted	
Jackson	Submitted	Submitted	
Jefferson	Submitted	Submitted	
Juneau	Submitted	Submitted	
Kenosha	Submitted	Submitted	
Kewaunee	Submitted	Submitted	
La Crosse	Submitted	Submitted	
Lafayette	Submitted	Submitted	
Langlade	Submitted	Submitted	Parcels submitted by City of Antigo separately from county
Lincoln	Submitted	Submitted	
Manitowoc	Submitted	Submitted	
Marathon	Submitted	Submitted	
Marinette	Submitted	Submitted	
Marquette	Submitted	Submitted	
Menominee	Submitted	Submitted	
Milwaukee	Submitted	Submitted	
Monroe	Submitted	Submitted	
Oconto	Submitted	Submitted	
Oneida	Submitted	Submitted	
Outagamie	Submitted	Submitted	
Ozaukee	Submitted	Submitted	
Pepin	Submitted	Submitted	
Pierce	Submitted	Submitted	
Polk	Submitted	Submitted	
Portage	Submitted	Submitted	
Price	Submitted	Submitted	
Racine	Do Not Exist County-wide	Submitted	
Richland	Submitted	Submitted	
Rock	Submitted	Submitted	Parcels submitted by cities of Beloit and Janesville separately from county
Rusk	Do Not Exist County-wide	Submitted	
Sauk	Submitted	Submitted	
Sawyer	Submitted	Submitted	
Shawano	Submitted	Submitted	
Sheboygan	Submitted	Submitted	
St. Croix	Submitted	Submitted	
Taylor	Submitted	Submitted	
Trempealeau	Submitted	Submitted	
Vernon	Submitted	Submitted	
Vilas	Submitted	Submitted	
Walworth	Submitted	Submitted	
Washburn	Submitted	Submitted	
Washington	Submitted	Submitted	
Waukesha	Submitted	Submitted	
Waupaca	Submitted	Submitted	
Waushara	Submitted	Submitted	
Winnebago	Submitted	Submitted	
Wood	Submitted	Submitted	



Map 01. Statewide Address Point Layer



Map 02. Statewide Parcel Layer

2 PROJECT CHARACTERISTICS

2.1 Data Standards

2.1.1 Data Standards Development

The LinkWISCONSIN Address Point and Parcel Mapping Project was based on two paramount assumptions that shaped the development of data standards:

- Counties and municipalities within Wisconsin are the authoritative source for address points and parcels. These contributors are understood to be the most accurate, well-attributed, precise, and contemporary sources for their respective jurisdictions.
- Although contributing data is consistently of relative high quality for a given area, errors and inconsistencies naturally exist in the data. It was not within the project scope to actively rectify errors in the contributing datasets.

The International Organization for Standardization (ISO) defines *data standards* as:

> Documented agreements containing technical specifications or other precise criteria to be used consistently as rules, guidelines, or definitions of characteristics to ensure that materials, products, processes, and services are fit for their purpose.[4]

Data standards add essential structure to a dataset by designing and defining the empty framework that will contain the dataset. In defining data standards for this project, research and planning played a key role in achieving a structured and streamlined workflow and successful deliverables.

The project team focused on three aspects of research and development in defining data standards for this project.

Data Standards Development
• Administering an address point and parcel **pilot project**
• Researching success examples through **case study research**, data standards, and best practices in addressing
• Coordinating with the PSCW in **defining the data model** to best fit business needs

Pilot Project

During the early phases of this project, the project's technical team administered a pilot project to assess the nature of address point and parcel data at the county level. The main goal for the pilot was to add context to defining the data model by preemptively assessing data that the team would be calling for in subsequent phases of the project. Documentation from the pilot helped to advise rough drafts of the schema and initial conversations with the PSCW's broadband mapping team. The pilot guided the process in tailoring to the business needs of the PSCW and also granted the team members an overall better understanding of the challenges that the project would face.

The pilot project assessed county-level address point and parcel data held in the Robinson Map Library's Geospatial Archive on the University of Wisconsin-Madison Campus.[5] Datasets were selected according to

---

[4] http://www.iso.org/iso/home/faqs.htm
[5] http://www.geography.wisc.edu/maplib/gisdataWI.php

availability and distribution of population density to provide the best assessment of rural and urban county contributors. The data was skimmed for common attributes and tailored into a schema crosswalk that effectively mapped the elements in one schema to the equivalent elements in another schema.

The pilot project crosswalk identified some of the challenges that the project could expect to face, such as:

- Frequency of one-to-many mappings, indicating that parsing or concatenation would be required
- Missing elements in schemas, indicating that several fields would not be capable of a 100% return across the state
- Lack of equivalency between schemas, indicating that domains would require standardization

Case Study Research

The team looked to geospatial business plans related to addressing and parcels from other states in research geared toward uncovering elements of success. Some examples, like Minnesota's *Business Plan for Statewide Parcel Data Integration*, proved useful as comparisons. Utah's plan for a statewide address point layer, also the product of broadband-driven National Telecommunications and Information Administration (NTIA) grant funding, proved to be a particularly instructive success story. Several reasons made the Utah project a quality success story to build from:

- Relatively well documented, public facing project
- Similar project context and objective
- 1 million address points, compared to Wisconsin's 2.7 million
- Mix of urban and rural residence, similar to that of Wisconsin
- Presence of grid address system, common in southeastern Wisconsin
- Funded by a broadband-related grant with a deliverable to serve similar purposes, i.e., geocoding service[6]
- Thoughtful, forward-looking and transparent schema[7]
- Applied the FGDC Addressing Standard

In tandem with case-study research, the project team also focused on targeting an address data standard that could advise the project's data model. Two addressing data standards were researched, each offering distinct qualities, the National Emergency Number Association (NENA) addressing standard and the Federal Geographic Data Committee (FGDC) addressing standard.

NENA Addressing Standard. The NENA Addressing Standard is focused around E-911 objectives and is a well-adopted standard with a tenured track record. At the time of this project's inception, NENA was in the process of rewriting the standard as part of its alignment with Next-Generation 911 efforts and alignment with U.S. practice for data exchange with European standards.[8]

FGDC Addressing Standard. The FGDC Thoroughfare, Landmark, and Postal Address Data Standard—more colloquially known as the "Street Address Data Standard"—is a standard that targets address data management requirements for local address administration, postal and package delivery, emergency response and navigation, administrative recordkeeping, and address data aggregation. It is a relatively new standard, endorsed by the FGDC as the official data standard for the U.S. in 2011. This standard replaces its predecessor, the "FGDC Address Data Content Standard," which was discontinued in 2005. The FGDC Street Address Data Standard project is the result of over five years of community-based research and development focused on building a forward-looking framework for developing address repositories.[9]

---

[6] http://gis.utah.gov/utah-statewide-address-geocoding-web-service-upgrade/
[7] https://docs.google.com/a/uwalumni.com/document/d/1eTgknNbA0UNXnyMDR5q9gFAm0-XtNYQpLLYPSZtCLTU/edit
[8] http://dev.nena.org/apps/group_public/download.php/2910/20131029%20ECRF-LVF%20NENA%20STA-005%20DRAFT%20PubRvw.pdf
[9] https://www.fgdc.gov/standards/projects/FGDC-standards-projects/street-address/index_html

The project team identified the FGDC addressing standard as the ideal standard to model from, for a few key reasons:

- It is intended to be used as an interchange format and is flexible in how its attributes can be transformed
- It was the most contemporary accredited addressing standard, as NENA's new standard was still in development
- It was followed by the Utah model, which proved to be successful
- It focuses on more than utility-specific objectives (i.e., not exclusively E-911, mailing, or navigation) as addressing for this project was to be leveraged for multiple utilities

Defining the Data Model

The project team utilized the pilot project, case-study research, and FGDC standards to advise the design of a documented data model. The final version of the Address Point Schema Definitions and Parcel Schema Definitions appear in Appendix A and B of this report.

For the purpose of this report it is important to distinguish the terms "Data Model" and "Attribute Schema" (or "Schema"). Both terms are used in context while describing data standards development and throughout the report.

Data Model. Data model refers to the technical configurations for containing geographic objects and their respective attributes as data within a database. For this project, the geographic elements representing address points and parcels were stored in a "vector" data model as collections of points and polygons. The data model also includes the configurations for how attribute data for each respective geographic element is stored as data within the database.

For this project, the attribute data was stored as a table with a one-to-one relationship to its corresponding geographic elements (points or polygons). The data model for this project also describes the relationship of elements within the scope of the project. For example, the RelateID is a field designed for relating address points to the parcel to which they belong as a one-to-many relationship.

Attribute Schema. Attribute schema refers to the tabular structure and organization of the attribute data associated with geographic elements. The attribute schema is important element of the overall data model

The data model and underlying research were presented to the PSCW in the early phases of the project and were beneficial as a planning framework for the project. This schema was altered to fit the PSCW's business needs related to the layers. The research to define the data model proved to be beneficial, as it helped the team to anticipate the issues they would encounter regarding data robustness, quality, and drawbacks of attribute information.

More specifically, the research to define the data model helped address a few central issues:

- Attribute schema the team could expect from local level data contributors
- Where the team could anticipate gaps in the data, in both the tabular or geometrical sense
- Ancillary data the team would need to pursue to meet the PSCW's needs

2.2 Integration Tools

When the LinkWISCONSIN Address Point and Parcel Mapping Project began, the team knew that one of the more challenging factors was going to be the amalgamation of more than 72 different datasets into a single file geodatabase. There were a number of factors considered when deciding which integration tools to use.

Factors Influencing Integration Tool Choice
• Wide variety of data formats
• Need to accommodate a wide variation of incoming attribute schemas
• Ensuring as much repeatability as possible

There were various options for integration tools, some open-source and others proprietary. The team chose to use the Esri ArcGIS Data Interoperability Extension, an add-on extension to the general ArcGIS license. This extension includes a subset of tools from Safe Software FME Desktop product. It has the ability to read in more than 134 file types and write out over 100 file types. It also includes more than 240 prepackaged transformers that allow for a wide variety of data manipulations, transformations, and extractions.

A few of the transformers the team employed include:

- Spatial Relator – Determines spatial relationships between sets of features based on a specified relationship (e.g., touches, overlays, and intersects)
- String Concatenator – Concatenates values of selected attributes and places results in new attribute
- Case Changer – Changes the case of selected text attribute fields

After the schema had been finalized, the team created a template tool containing a writer for both the address points and parcels. Once all the preprocessing was complete for a given county dataset, the team would duplicate the template tool, add the new readers (input datasets), determine what transformers would be necessary, and begin constructing the mappings from the native dataset to the final output schema (writer).

From the beginning of the project, the team desired to make this process as repeatable as possible for future potential iterations. Once county tools were saved, and assuming that the attribute schema from the reader had not changed, the technician would only have to add the new data readers into the workbench, ensure all attribute mappings were maintained, and run the tool again.

The technical team encountered a few issues with the interoperability extension that are addressed in Chapter 4. Fig. 4 depicts an overview of the data interoperability workbench.

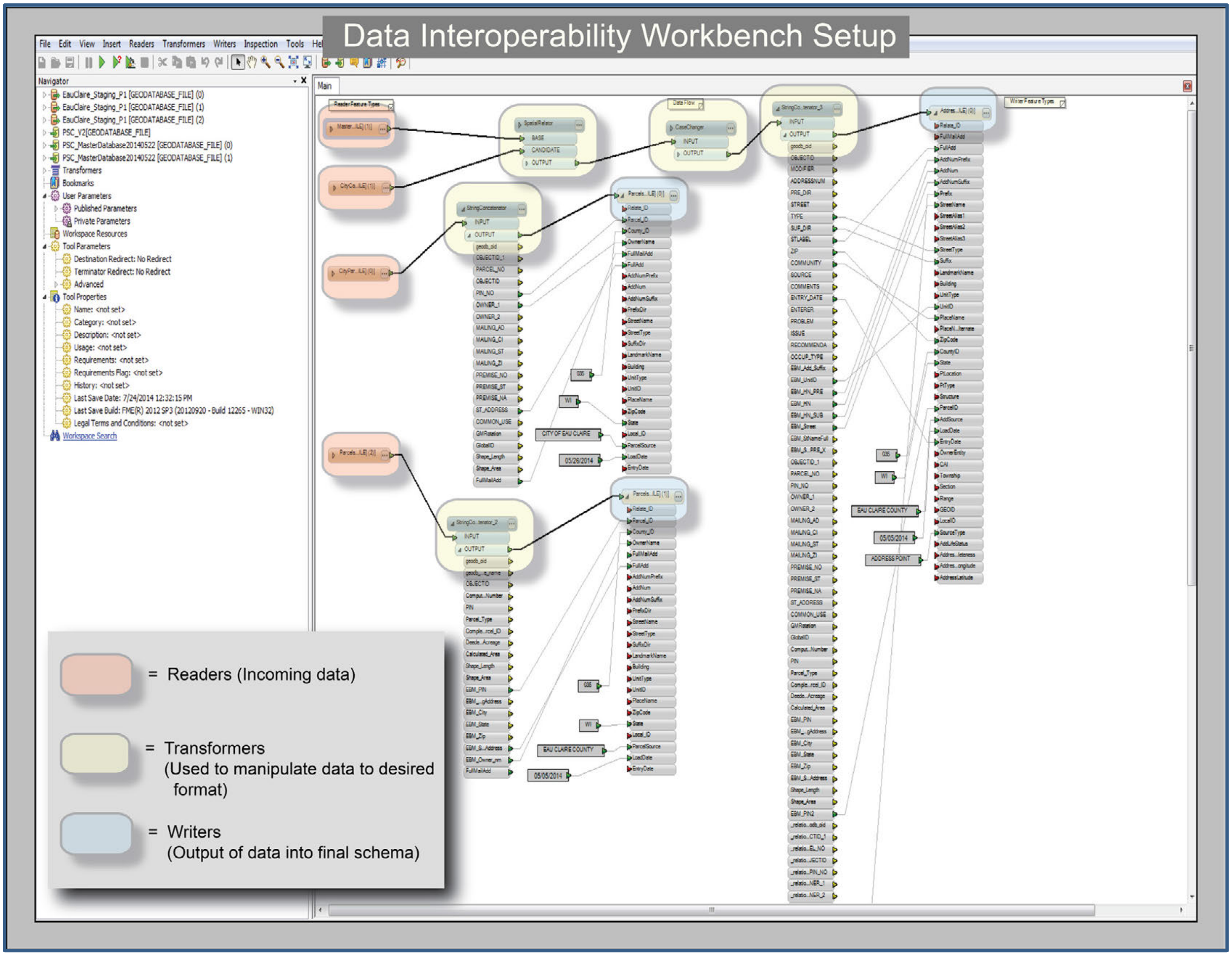


Figure 4. Data Interoperability Workbench Setup

2.3 Workflow

In terms of workflow, the project was broken down into four sequential/temporal phases: 1) Data Submission, 2) Ingest & Staging, 3) ETL (extract, transform, load), and 4) QA/QC.

A secondary breakdown of the project was marked by two logical phases in workflow: 1) Local-level Logic, and 2) State-level Logic. Both breakdowns of the technical workflow are depicted below (Figs. 4 and 5).

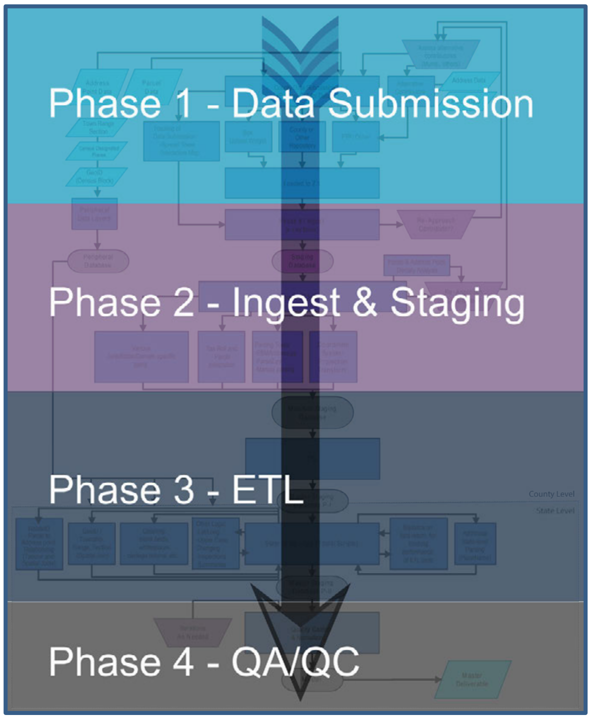


Figure 5. Technical Workflow – Chronological Breakdown

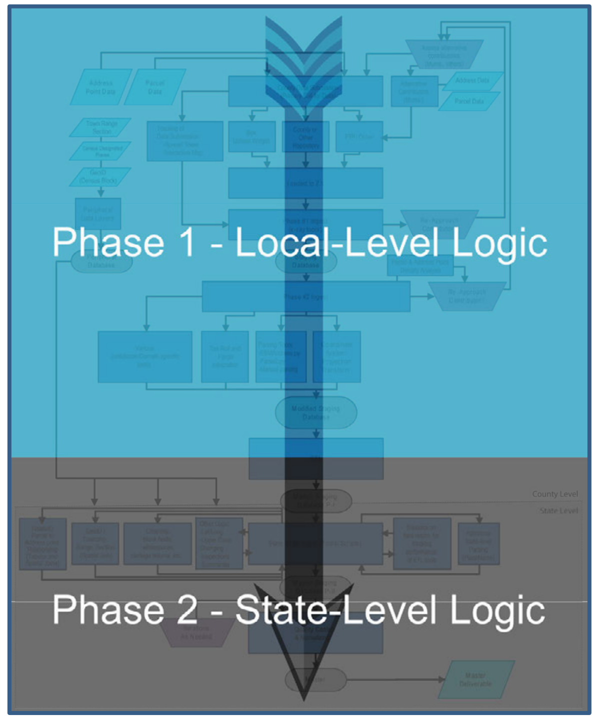


Figure 6. Technical Workflow – Logical Breakdown

Phase 3 of the chronological breakdown was the main technical focus of the project, where the team built logic around the ETL tools. This was the point where data was aggregated from local-level to state-level.

A number of processes were performed over datasets en route to the final deliverable. Each process was identified to be most appropriately applied to the participating jurisdiction through “local-level” logic or to the state as a whole through “state-level” logic. This segmentation of processes was a natural determinant in the workflow of the project, as state-level logic could only be performed after the ETL process and local-level logic could only be properly applied before the ETL process. In other words, in Figure 3 above, the state-level and county-level phases are divided by the ETL process.

The detailed workflow identifies the general logic that the data flowed over in the process of achieving the final deliverable from local-level data submission to final statewide database. The detailed technical workflow appears in Fig. 7.

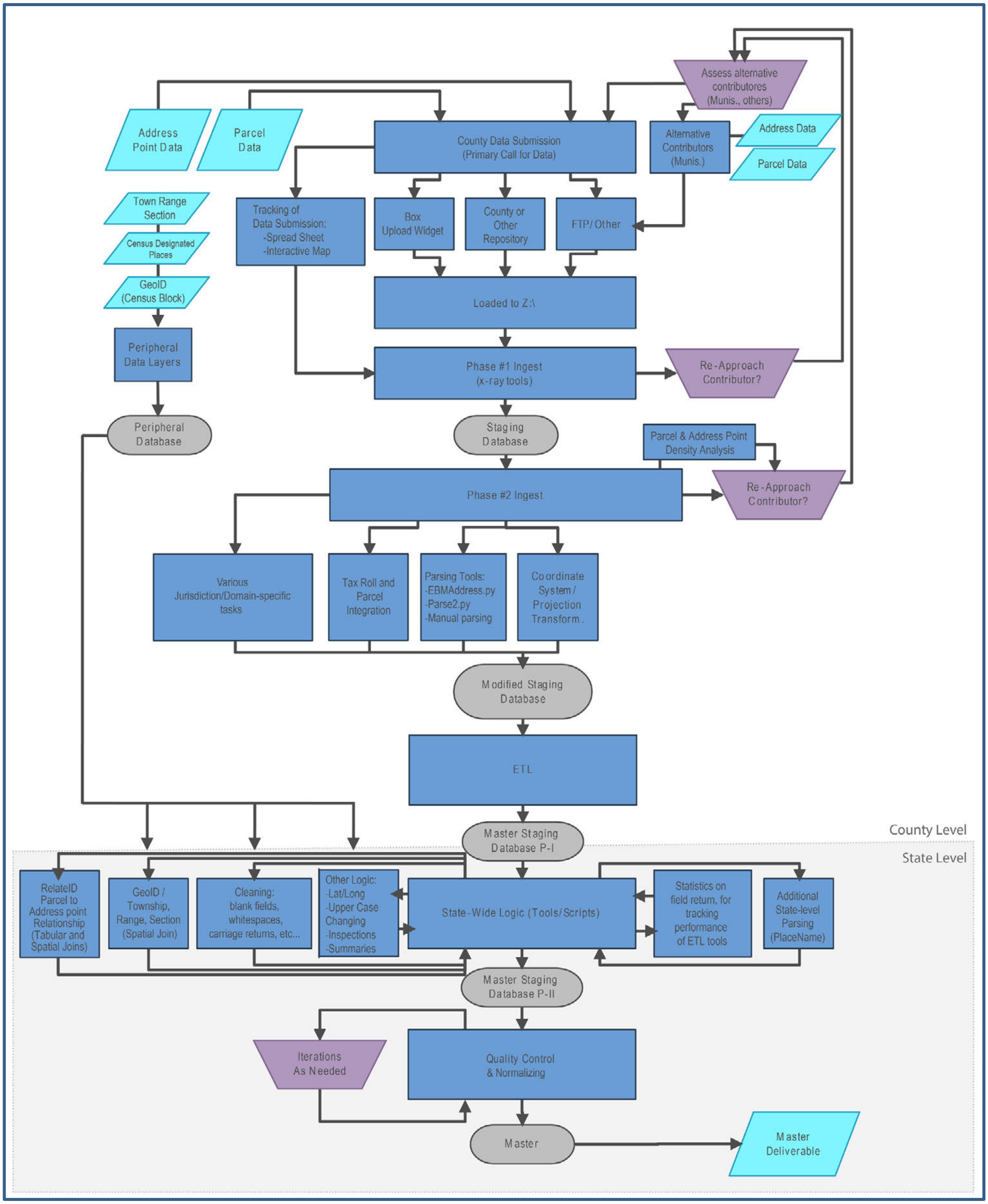


Figure 7. Technical Workflow – Detailed View

2.4 Milestones and Interim Deliverables

The project MOU set forth a set of interim deliverables, which proved to be valuable in continually assessing the data model and how it could best fit the business needs of the PSCW.

Project Deliverables	
Deliverable	
Version 1 Database Delivery December 30, 2013	• Included Wisconsin's earliest contributing counties that best fit the project's attribute schema in the deliverable (20 counties total) • Did not include quality control or standardization
Version 2 Database Delivery March 31, 2014	• Included Wisconsin's "early-adopter" and "mid-adopter" counties in the deliverable (55 counties total) • Included limited implementation of quality control and standardization
Version 2.1 Database Delivery June 10, 2014	• Included all of Wisconsin's contributing datasets in the deliverable (all participating counties and municipalities) • Included limited implementation of quality control and standardization
Final Database June 30, 2014	• Included all of Wisconsin's contributing datasets in the deliverable (all participating counties and municipalities) • Included all formal deliverable items for address points and parcels with the exception of those items related to the CAI project • Included comprehensive quality control and standardization measures
Final CAI Database September 30, 2014	• Improved the spatial accuracy of features in PSCW's previous CAI database • Expanded the current classification of CAIs to include correctional facilities, ports, and private schools

3 DATA METRICS & ANALYSIS

3.1 Parsing and Concatenation

Due to the various configurations of attribute schemas for address points and parcels at local levels across the state, significant parsing and concatenation actions were required to fit local schemas.

3.1.1 Parsing Objectives

One of the most essential tasks to properly map components to the project schema was to parse all required elements out of their various schema configurations. The parsing tools targeted would need to be flexible in handling the various and unique parsing needs of each data contributor.

There are several out-of-the-box commercial parsing options available through cloud services or desktop applications that are effective for general address parsing. These parsing options are often part of smaller components to a geocoding workflow, as parsing address elements is often a necessary step taken for a geocoder to digest and locate an address. While these services are well-designed, intuitive, and mostly cost effective, the project team identified custom parsing options to be the most appropriate approach for this project. With parsing and address standardization amongst the largest challenges that the project would face, the decision to use custom parsing tools was based on several factors, elaborated below.

Utility of Custom Parsing Tools

Efficiency. There were 2.7 million address records in the final deliverable and it was out of the project scope to validate every record manually. This concept increases the importance of fully understanding and having control of the logic behind the tool since the team could not personally validate every record.

Preservation of Authoritative Data. Commercial address parsers utilize auxiliary or underlying data sources to drive logic, validate results, and serve as surrogate for missing data. Pursuing such logic would conflict with the project's concept that data contributors are the authoritative source for their jurisdictions. Making this dataset conform to the likes of third-party datasets would undermine the objective of creating a statewide layer from authoritative data. Avoiding auxiliary datasets would help maintain the data integrity intended by the counties and municipalities.

Platform Continuity. The majority of this project's logic was implemented in the ArcGIS environment, which accommodates Python scripting. Python is a good language to use for writing parsing code due to its support for regular expressions, ease of use, and broad community support.

Less Cumbersome Workflow. Commercial/third-party parsers typically require processing of CSV (comma-separated values) or other non-spatial files. These types of files are a bit more cumbersome when working with geospatial data because tables need to be joined back to their geometries after the parse is complete. Keeping all logic within the same GIS environment significantly improved workflow time and reduced the risk of errors.

Flexibility. Commercial software is also generally packaged in a way that does not offer a high degree of flexibility in the type of components being parsed. The team wanted to be able to implement the same or similar tools across all native datasets, despite the variation in elements parsed and varying inputs across the datasets.

Local Adaptability. Custom logic is necessary for Wisconsin-specific address styles, such as grid addresses. The more conventional linear address appears in most parts of Wisconsin, with the important exception of the southeastern part of the state where grid addresses are commonly found. The image on the next page illustrates a grid address in comparison to a more conventional linear address, and thus the need for a custom parsing tool to handle such cases.

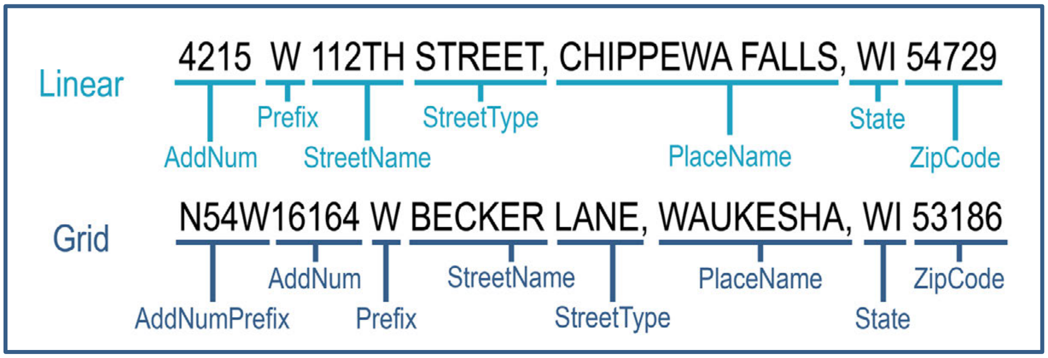


Figure 8. Linear and Grid Address Type Comparison

3.1.2 General Parsing Challenges

Each contributing county or municipality has unique arrangements of attribute information, as detailed in the address points crosswalk, available as a Digital Appendix to this report from the SCO website. One of the biggest challenges for data aggregation projects such as this is that no two attribute schemas are 100% equivalent. Any given schema may have a field that does not exist in another schema, or it may have a field that is split into two different fields in another schema. In many cases, this is why data is often lost or integrity is compromised when mapping from a complex schema to another similar one. Thus the LinkWISCONSIN Address Point and Parcel Mapping Project entailed a fair amount of parsing and concatenation in order to get local schemas to fit the project schema.

Common Parsing & Concatenation Challenges

Challenge	Example	Parsed or Concatenated Solution
Schema with an element requiring segmentation to multiple other elements matching the target schema, requiring a "one-to-many" mapping	2334 E. Johnson St.	→ 2334 | East | Johnson | Street
Differing domain formats within a single dataset	County Highway, County Trunk, Cty Tk, County Road	→ CTH
Schemas include extraneous information	Unit #4 (Dairy Barn)	→ Unit | #4
Full address does not exist in dataset, except as individual elements	Elements 2334 | East | Johnson | Street	→ 2334 E. Johnson St.

3.1.3 Project Address Points Schema

The elements of the project's address schema are pictured in Fig. 9, with special attention to address elements.

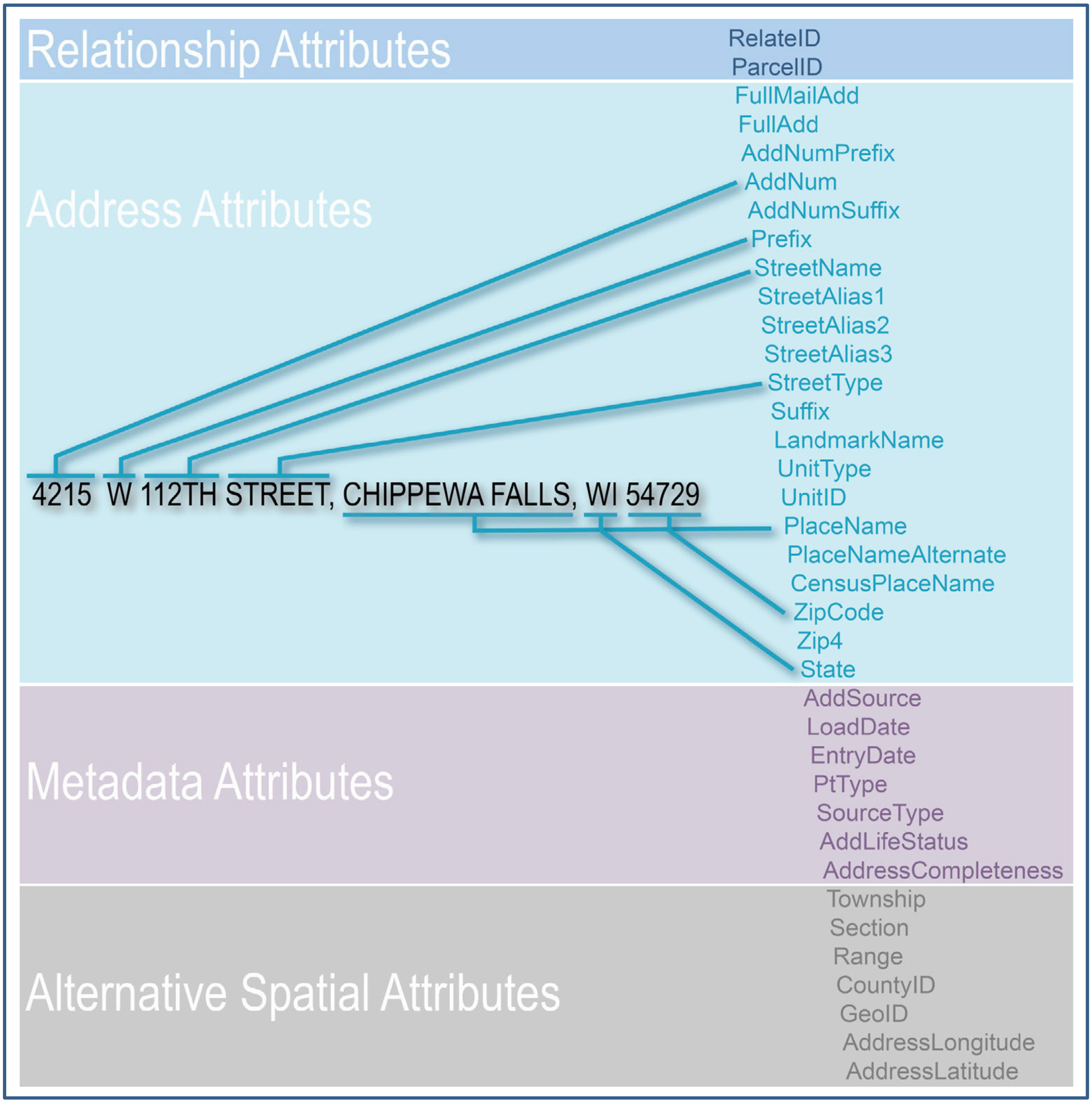


Figure 9. Address Elements of Project Schema

Some common address scenarios are demonstrated in Fig. 10 against the schema elements designed to contain them.

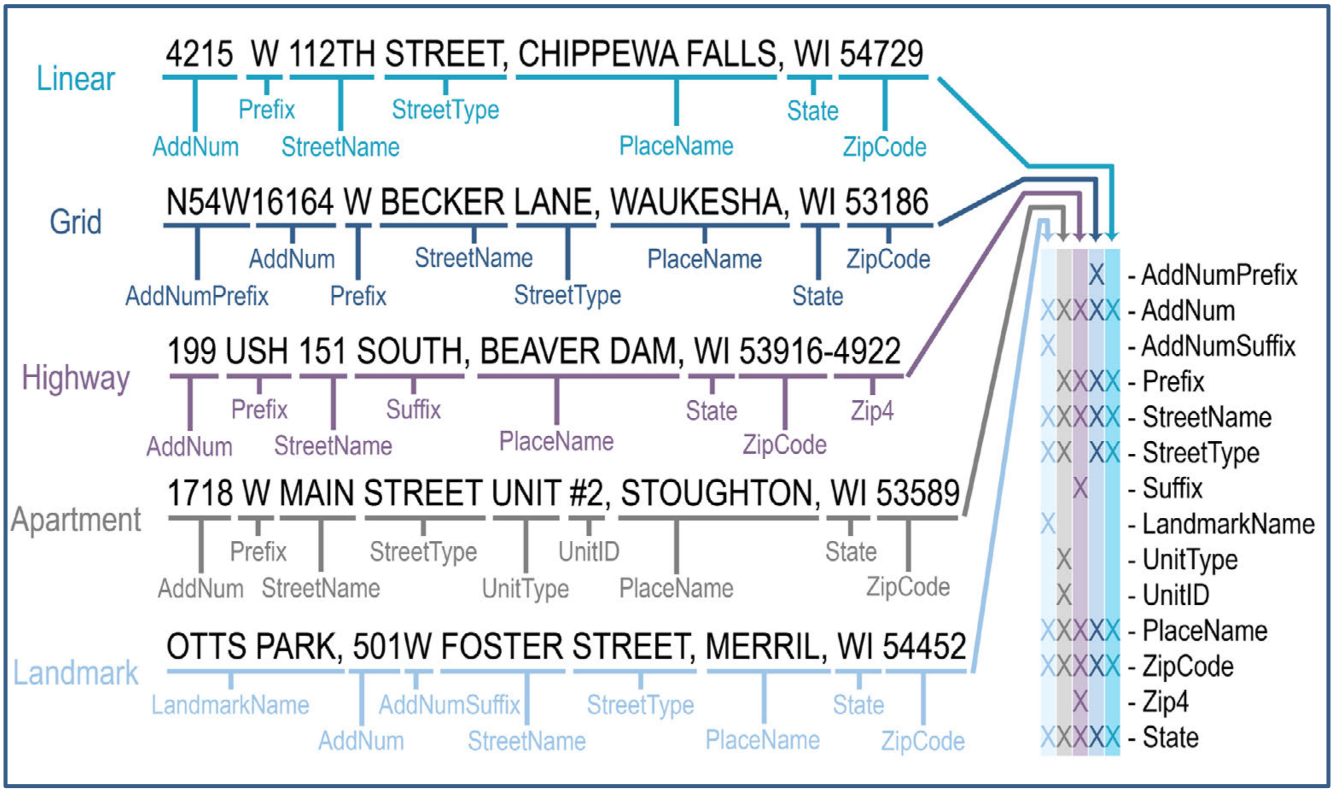


Figure 10. Address Scenarios and Schema Attributes

3.1.4 Building the Tool

In preparation for the ETL process the project team used ArcGIS software to format each contributor's dataset into Esri file geodatabase format, in what were referred to as staging databases. To maintain continuity with the ArcGIS platform, the team leveraged the ArcPy module for writing custom parsing scripts. ArcPy provides an interface to which developers can read and write geospatial elements through Python code but also gives developers access to Esri-specific geoprocessing tools for use in development. ArcPy scripts can be packaged as a "script tool" for ease of use. These tools can also be integrated as part of a sequence of processes, such as ModelBuilder or other scripts. Leveraging the ArcPy module allowed the project team to create a highly streamlined processing sequence.

The project team did not have to start from scratch in writing the source code for the parsing tool. The tool leveraged Python source-code from GitHub called Pyaddress[10] that offered a good initial framework. This source code enabled the team to rearrange how certain parsing logic was handled so that it best fit the desired objectives. Pyaddress was originally developed to parse CSV files, so the project team altered the code so it was able to read and write GIS formats such as the file geodatabase. After these adjustments, Pyaddress could be used as part of an ArcPy tool capable of reading an address or address component, parsing and writing it to the appropriate address fields directly within the GIS.

[10] https://github.com/SwoopSearch/pyaddress

Project Tool Capabilities	
Degree of Parsing	
Full address parse	Address number prefix through zip code
Partial address parse	All elements that exist between address number prefix through suffix and any configuration of those elements in between
Split on various characters parse	Multi-purpose tool used for a variety of needs, e.g., multiple elements existing in one field

3.1.5 Post-Parsing Observations

Early stages of project coordination identified a common attribute schema for the contributing datasets to meet, as depicted in the images of the schema above. This attribute schema focused on quality addressing practices by segmenting and isolating all recognized address elements, a concept driven by the FGDC Thoroughfare, Landmark, and Postal Address Data Standard. Among other assets, a schema with segmented and isolated address elements is a precise schema to model for organizing, field-mapping, standardizing, and implementing quality control to addresses within a database.

Segmentation and isolation would also be required to meet the needs of less precise attribute schemas. Segmentation is generally required to divide all address components before QA/QC (Quality Assurance/Quality Control), standardization, or proper concatenation can be performed.

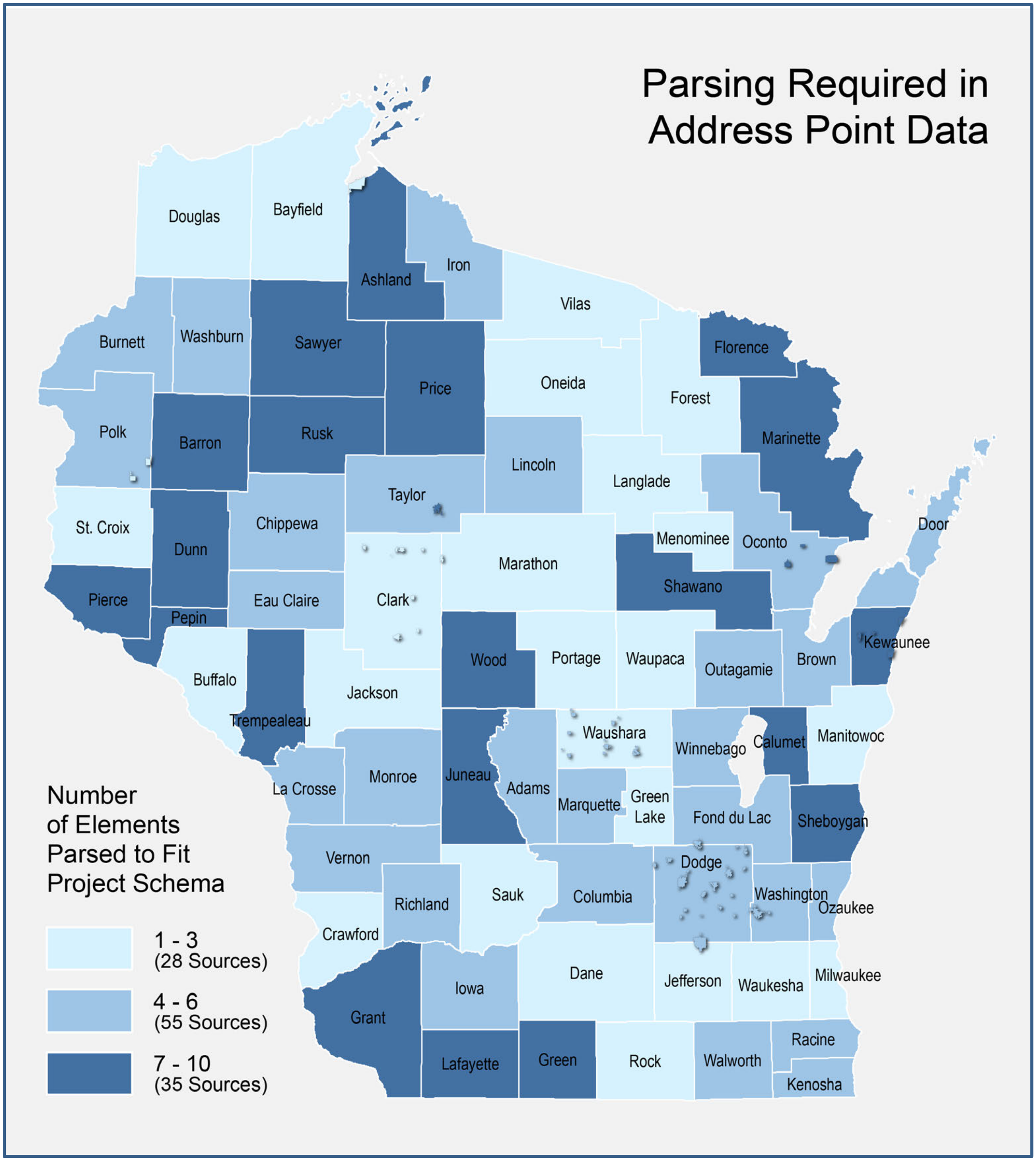
In meeting the common attribute schema, the project team encountered a wide variety of parsing needs, with variation from contributor to contributor. The need for parsing also varied within each of the contributor's datasets. These parsing needs ranged from splitting small address components into individual fields to splitting, identifying, and organizing full addresses into multiple address components. Each data contributor required a different parsing strategy to generate the appropriate address elements fitting the project's attribute schema.

Parsing Challenges

The amount of parsing required varied greatly from county to county. In some cases a full address parse was required to segment out all of the desired address elements. Most often, a partial parse of varied degrees was required. Examples of partial parses included:

- Address Number Prefix & Address Number & Address Number Suffix
- Prefix & Street Name
- Street Name & Street Type

The map on the next page identifies the data contributors with the most parsing required in dark blue. The amount of parsing required amongst individual datasets ranged from one to ten elements. In total, there were 389 elements segmented across all data contributors amounting to an average of five elements requiring parsing per contributing dataset.



Map 03. Parsing Required in Address Point Data

Extraneous Attribute Data. The team encountered a few cases where extraneous data had to be removed prior to running the parsing script. In most cases, these instances were dealt with during the pre-processing and cleaning phase. This cleaning was essential to ensuring that the parsing script would perform correctly and place the appropriate address elements in to their appropriate element field. A few examples of extraneous data uncovered were:

- County FIPS attached to PlaceName
- Building descriptor attached to Full Address
- Alt Street Name attached to Full Address
- PlaceName abbreviation attached to end of Full Address

Domain-Specific Tweaks to Parsing Approach. There is no simple one-tool-fits-all approach to parsing every dataset in a project of this size. For example, different counties had different ways of abbreviating State Hwy, County Hwy, and US Hwy. The parsing tool had an external CSV file, Replacements.csv, which was referenced when parsing address elements to further tailor the tool. Prior to the execution of the parsing script, time was taken to examine the native data and identify any unique abbreviations and unique street names. Once identified, the native spelling and desired output spelling were added to the Replacements.csv file. The process of altering the Replacements.csv file prior to executing the script helped greatly reduce the amount of standardization and cleaning that was required during the state-level processing stage.

Common Scenarios Requiring Address Parsing

Description	Example
Street Name and Street Prefix consistently included in one field	East Johnson • East / Johnson North Main • North / Main
Highway type (Highway, Interstate, County Highway, State Highway) and the highway route number or letter included in one field	County Highway AA • CTH / AA Interstate 94 • INT / 94
Address numbers with grid components	Required parsing to remove and categorize grid address components (addNumPrefix), unit numbers, unit types or addressNumSuffix

Figure 11 displays the number of elements that required parsing to meet the the project's attribute schema by element. Street Name and Prefix rank the highest among these elements.

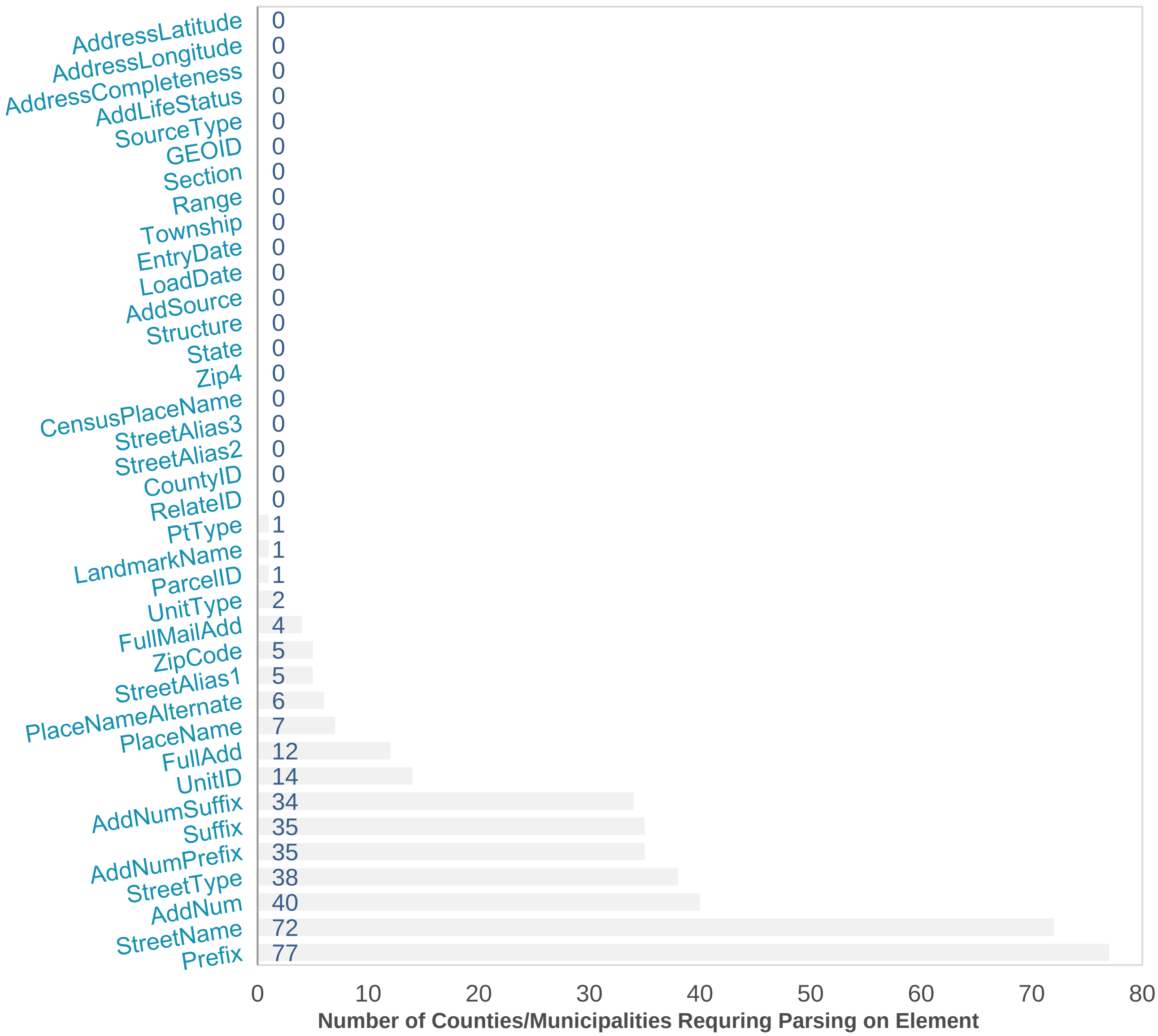


Figure 11. Parsing by Element

3.2 Cleaning and Standardizing

From the pilot project and research, the team expected the cleaning and standardization of the final database to be a large task. Due to the historic nature of cadastral data, error, and discrepancies within a single dataset are likely to exist. When considered alongside the integration of more than 72 different datasets into one final database, the potential for discrepancies and error is significant. The team wanted to ensure that the final product closely followed the FGDC standards of the attribute schema, and thus took careful cleaning and standardizing measures.

The focus was on cleaning and standardizing on a number of attributes, including Prefix, Street Type, Suffix, Unit Types, and Zip Codes. Some examples of standardized attributes are listed in the table below, with the full list in the schema definitions, Appendix A and B.

Examples of Standardized Attributes		
Attribute Type	**Domain-Specific Examples**	**Standardized Attribute**
Prefix	CTY TK, CTK, County Road, CY TK, CTH, etc.	CTH
Street Type	CE, CI, CIR, CR, CRL, etc.	CIRCLE
Street Type	BLV, BLVD, BV, etc.	BOULEVARD

Put simply, the idea behind standardizing attributes is to ensure that the final product has one spelling of each entity that can exist in a given field. To help automate this cleaning and standardization, the team created some ArcPy script tools. They first produced summary tables of fields they wanted to standardize, and added a new column that contained the correct spelling or abbreviation for a given entity. The result was a lookup table that could map variations in a domain to one common domain. Using the tool, they joined the lookup table back to the master dataset and changed the targeted field to the new cleaned and standardized name.

Where it was identifiable, the team also made corrections to certain parsing errors that existed in the Street Name field as well. In a number of cases, the native data did not have the Street Name and the Street Type separated. These cases were identified and corrections were made using the standardization tool.

The team also encountered a few special cases that required some manual cleaning done via a variety of different methods. Examples of attributes requiring manual cleaning appear in the table below.

Attributes Requiring Manual Cleaning	
Challenge	**Examples**
Duplicate information attached to the end of the full street address	301 W 1st St. CTH H 7665 S Main St. **State Road 53**
Extraneous information attached to the end of a full address	110 Johnson St. Cabin 3442 State Highway 66 **Dairy Barn**

These special cases were handled using a combination of Microsoft Excel find and replace methods, and a split script the team developed that divided an attribute field based on a delimiter designated by the data processor. These instances of cleaning often took place prior to any standardizing or parsing, due to the problems this additional information could cause when running native data through the parsing scripts.

3.3 Local Model to Project Model

3.3.1 Local Model to Project Model Trends

During the secondary inspection of the data, the team examined how a county's native attributes would fit into the final project data model. The amount of time and effort necessary to determine how these native fields mapped to the final data model attributes was significant. In some cases the attribute mapping was straightforward. Other times, a more in-depth investigation of particular attributes was necessary. This was also the stage where the team began to determine parsing requirements for a given county. The Digital Appendix for this report features a full crosswalk of local to project data model mapping for address points.

Consistency within a spatial database is one of the key elements to spatial data quality assessed in relationship to the database itself. Like other elements of data quality, there are three dimensions to consider—space, time, and theme.

Space: Topological Consistency

To achieve topological consistency, the data must conform to topological rules and have a consistently applied topology. It was beyond the scope of this project to formally assess or rectify the topology of datasets participating in this project individually or as a statewide layer. However, several observations on topological consistency were made that can advise future projects:

- There are counties with internal topological inconsistencies such as duplicate polygons or points (for example, Dane County has one polygon per address. This is not an error, but is inconsistent with the practices of most other counties). Additionally, topological errors known as rogue polygons and sliver polygons were encountered in various counties.
- There are topological errors that result from data amalgamation such as parcels overlapping or underlapping along county boundaries. This was observed at a magnitude of up to 100 feet in some places.
- Address points do not consistently adhere to the same topological rules with relation to parcels, such as address points existing in right-of-way in many places in the state.

Time: Temporal Consistency

Temporal consistency requires the data represent only one event occurring at one place and one time. Address point and parcel datasets are a representation and product of geographic entities that change over time. If the dataset represents two timeframes at once, temporal inconsistencies will exist. It was beyond the scope of this project to actively assess the temporal consistency of this address point data, but some observations were gathered:

- There are retired addresses that exist in some contributing datasets. In many cases, these are attributed as "retired", which is advisable by the FGDC standard. The majority of contributors do not track retirement status of addresses and *if* retired addresses exist in the dataset that are unidentified as retired, this would be an error.

Theme: Consistency in Attributes

Thematic consistency, or consistency in attributes, is characterized by lack of contradictions and redundancy. Consistency in attribute information is one of the single largest flaws in the typical address or parcel dataset. Some of the most common thematic inconsistencies uncovered in this project included:

- Domain inconsistencies – e.g. State Highway entered in a multitude of ways in the same dataset, such as STH, ST HY, State Road, etc.
- Attributes inconsistently attributed to appropriate fields – e.g. Unit Numbers existing as part of the Address Number field for some records while existing in a Unit ID field for others

3.3.2 Join Errors and Unresolved Domains

The data for this project was received in a variety of different formats, including but not limited to character-delimited text files, CSV, Excel, and database files. Unfortunately, in some cases, the ability to join certain files within ArcMap tabularly is not straightforward and seamless. Join errors, post-join missing data, and automatic formatting of spreadsheet attributes are just a few of the issues encountered when working with external tabular information.

A few of the workaround solutions used to resolve these join issues included legacy saving Excel spreadsheets to a 97-2003 Excel Workbook, converting tables to stand alone database files, cleaning and removing extraneous attribute fields within Excel, and shrinking down the size of the file joined to the spatial data.

The team encountered a number of instances where native domains provided by contributors were unresolvable. In a few cases they were able to resolve these domains through their research (e.g., identifying a point's place name from a three character code provided in the municipality field). In other cases, this information had to be omitted from the final database since the domains were unresolvable and they did not want to introduce error by guessing at what a given domain value was representing.

3.3.3 Related and Generated Fields

The preparation phases of the project enabled the project team to look toward alternate sources for fields that were not adequately populated through local-level data contributions. A handful of attributes were populated at a statewide level as one of the many processes that participated in the state-level logic (see Fig. 12).

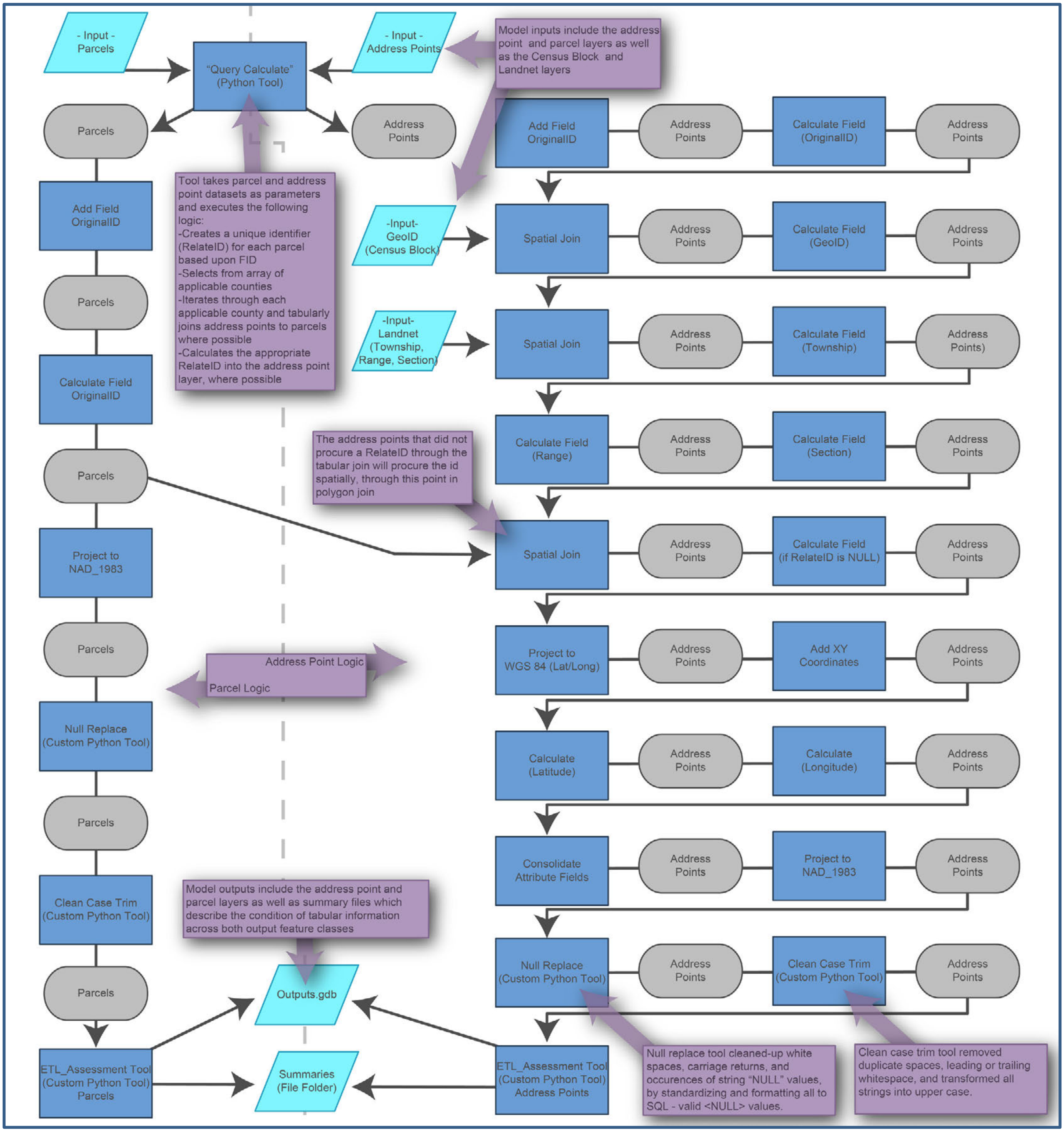


Figure 12. State-Level Logic

In contrast to the county-level logic, which was naturally segmented and entailed dataset-specific logic and components, the state-level logic was implemented as a consolidated sequence of processes. These processes all participated in one model that was developed separately but in tandem with processing and preparation of data for the ETL procedure. This model was designed so that the address point and parcel outputs from the ETL procedure could be ingested and processed over a weekend, directly following the completion of ETL.

As displayed in the diagram illustrating statewide logic on the previous page, this model involved a number of processes, some of them quite processor-intense, especially when being performed over millions of features. Development of the model required testing and scrutiny against a small sample of data which allowed for quick test processing times and a controlled test subject. The model utilized concepts for managing processing described in Chapter 4 to help keep processing under control.

Statewide Logic Components

RelateID. RelateID is the unique identifier that properly ties an individual address point to the parcel to which it resides. This element was constructed through the state-level logic through two logical phases. The first phase utilized a parcel ID in the address point (if it was available) to tabularly join the address point to its respective parcel. Achieving a tabular join on locally provided parcel IDs was identified as more accurate than using point-in-polygon logic.

The tabular process yielded a join on 1,225,252 features of the 2,613,727 features that procured a RelateID (46.9%). The features that procured a RelateID (2,613,727 features) accounted for 95.3% of the points included in the final deliverable. A total of 127,740 points did not gain a RelateID. They were non-relatable for the following key reasons:

- A tabular join was not available, *and* address points fell within the right-of-way of the parcel dataset, thus, the topology of point and polygon datasets would not allow for a spatial join. Eau Claire and Green Lake Counties as well as the City of Fond Du Lac were contributors where this was a common occurrence. This reason accounts for the vast majority of non-relatable points and was found to a small degree across all data contributors.
- Cases where Highway Mile was included in the address point layer. By design, these points have no corresponding parcels and thus could not be related or obtain a RelateID (common in Waukesha, Milwaukee, and Walworth Counties).
- Cases where there were no digital parcels available for a given area from the dataset due to incomplete digital parcel datasets. Thus, no relationship with an address point would be possible. This was observed within Buffalo, Burnett, Crawford, Langlade, Lincoln, Marquette, Menominee, Polk, Sawyer, and Vernon Counties, and accounts for approximately 31% (39,060 points) of the 127,740 non-relatable points.

AddressCompleteness. AddressCompleteness is a measure of essential address elements attributed within each element of the address point deliverable. The contents of this field were generated at the end of the QA/QC phase with a Python tool written by the project team.

AddressLatitude/AddressLongitude. AddressLatitude/AddressLongitude were generated with an out-of-the-box Esri tool within the state-level logic to consistently apply a latitude and longitude measure to each point. This is a practice that is recommended by the FGDC standard and makes the point file more exchange-ready and platform independent.

Census Place Name. Census Place Name was attributed to each point using point-in-polygon logic as a means to augment the contributor provided Place Name data with one logically-consistent Place Name field, as identified in the 2013 U.S. Census Bureau's Boundary and Annexation Survey.

TRS. TRS, known as township, range, and section were attributed as individual fields to each point using point-in-polygon logic. Wisconsin's "Landnet"[11] (PLSS quarter-sections) was utilized as the polygon feature class to spatially join each point to.

[11] http://dnr.wi.gov/maps/gis/documents/plss_quarter_sections.pdf

GeoID. GeoID distinguishes the well-known GeoID of the Census block that the address point resides within. The GeoID is a 15-digit code created and maintained by the Census. It is comprised of two FIPS ID components, a tract ID and a block ID. The appropriate GeoID was attached to each address point through a spatial join during the state-level logic.

Tabular Joins and Spatial Relator

In some cases, there was information provided in the local parcel layer that the team wanted to include with the address points. There were two different approaches taken to transfer this data to the address point features.

Tabular Joins. The first and most desirable approach was to join the data via a tabular join. The primary field the team attempted to do this tabular join on was the Parcel ID field (or a unique identifier like a parcel ID). There was a success rate of approximately 35% when it came to accomplishing a tabular join.

Spatial Relator Transformer. The second approach that was taken when a tabular join was not possible was through the use of the Spatial Relator Transformer within the data interoperability workbench. The primary attribute value that was obtained with this method was the Parcel ID. This method was required for roughly 65% of the counties in the state.

In areas where parcels were not available in digital format, it was not possible to obtain any information spatially. Some of the additional non-spatial attributes added to the address point features by way of a tabular join or the spatial relator included:

- Placename
- Placename Alternate
- Full Address
- Property Type

The tabular join was the preferable method for joining the data because it helped limit the amount of possible error. As one might expect, a spatial join can only be as successful as the accuracy of the address point spatial location. For example, in cases where address points fell within the right-of-way, no additional spatial information was added to those features.

3.3.4 Notes on Parcel IDs

A striking observation the team made was the number of counties that do *not* include a parcel ID in their address point layer. Roughly 35% of counties included associated parcel IDs in their address points. Around 65% of parcel IDs were obtained whenever possible using the Spatial Relator transformer in the data interoperability workbench setup.

There was also a large amount of variation amongst the attribute fields that the team presumed contained the parcel ID from county to county. In some cases it was very easy to identify the field containing this attribute, but in other cases they had to make an educated guess, selecting the field that contained a unique identifying number.

Local Variation in Parcel IDs
Attribute Fields Parcel IDs Appear In
• GISPIN
• TxParID
• TAG
• ONCPIN
• LANDCOMP
• COMPUTERNO
• Miscellaneous others

The team also observed an extremely large variation in the types of parcel IDs that existed throughout the state. The length of parcel IDs varied, ranging anywhere from 3 to 33 characters. A wide array of characters also existed in the parcel IDs ranging from numeric values, alphabetic characters, and numerous special characters like spaces, dashes, commas, periods, and pound signs. Understanding that each parcel managing entity has their own parcel ID requirements, it was not surprising to find such large variations in parcel IDs across the state.

3.4 Commonalities in Local GIS Datasets

3.4.1 Delivery Commonalities

One of the most common themes of addresses and parcel data across the state is that of data and delivery format. There are a few possible options for formats, including file geodatabase, shapefile, personal database, and DWG.

Formats

File Geodatabase. Performs as a native GIS database although it is comprised of a collection of various types of GIS data files contained within a file system folder named with a .gdb extension. This is a relatively new GIS file format that was first supported by Esri products in the ArcGIS 9.2 release in November of 2006. As this Wisconsin project attests, the file geodatabase has been widely accepted as an industry standard and is the recommended native data format for ArcGIS desktop use. The file geodatabase offers significant advantages over the personal geodatabase or the shapefile due to structural, performance, and data management improvements.

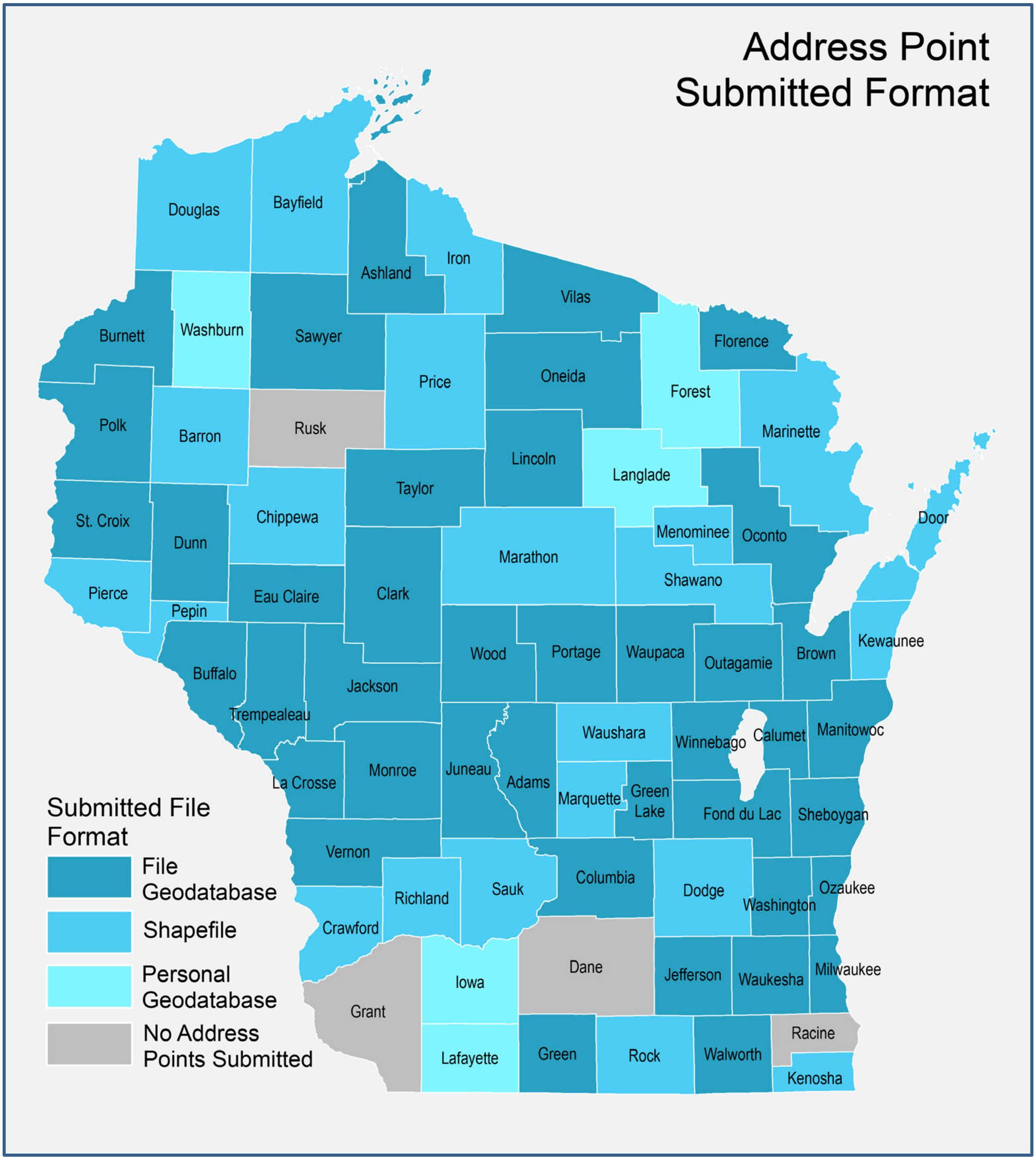
Shapefile. A commonly-used GIS data format developed and stewarded by Esri which was introduced in the early 1990s. It has an open data standard, which makes it more interoperable with other software in contrast to other formats listed here. Also, in contrast to a file geodatabase or a personal geodatabase, the shapefile does not have the ability to store topological information, a common asset to properly maintaining parcel datasets. The shapefile is limited to 2 gigabytes in file size, which can be problematic.

Personal Geodatabase. A Microsoft Access database that is enabled to store geospatial data. Like the file geodatabase, the personal geodatabase has the ability to store topological information but, like the shapefile, is limited to 2 gigabytes in size. Prior to the introduction of the file geodatabase, the personal geodatabase was the industry standard and recommended data format for working with parcel fabrics. Although the personal geodatabase is still supported by Esri, it is no longer recommended as other formats such as the ArcSDE or file geodatabase offer better structure, performance, and data management.

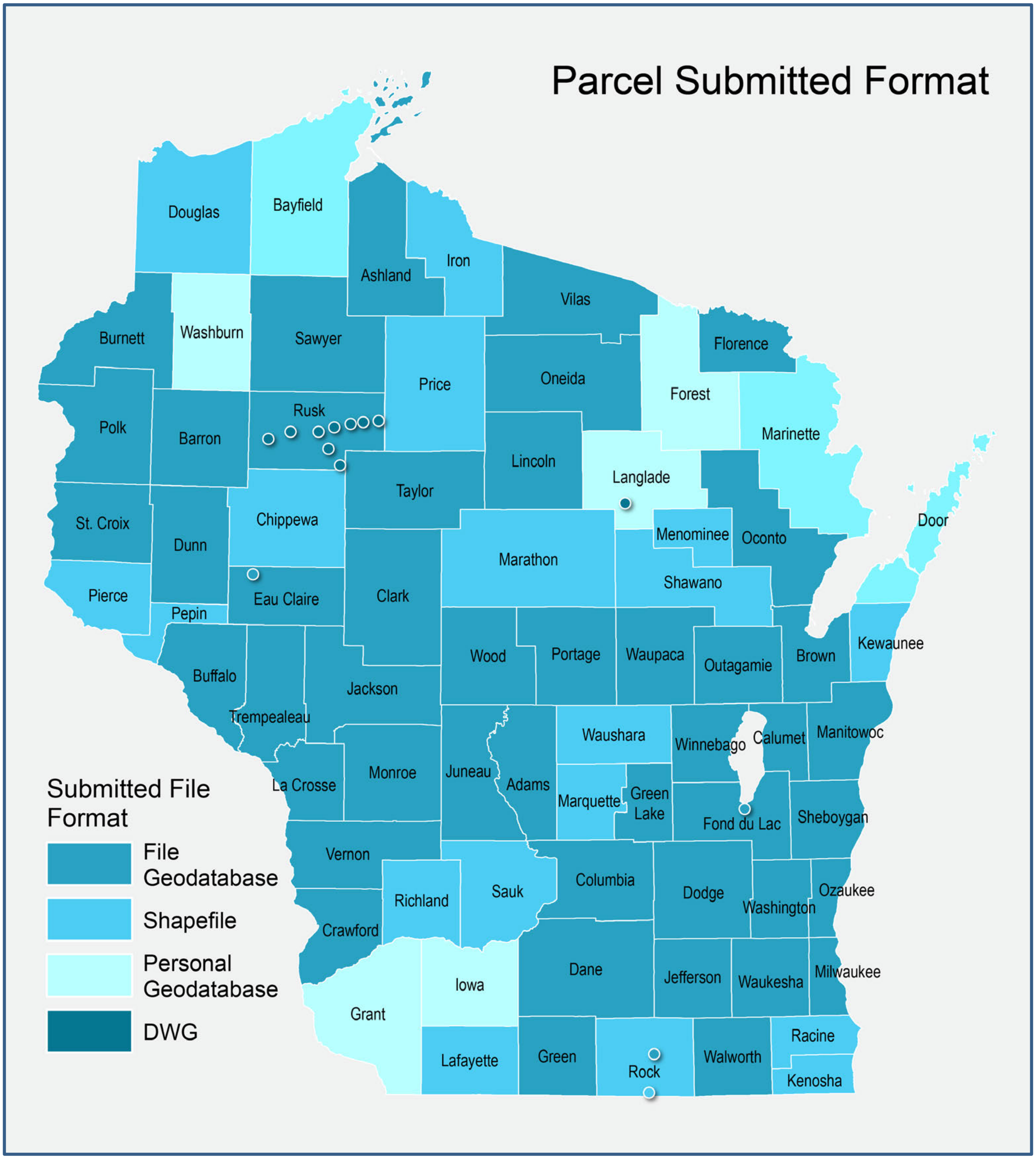
DWG. Short for “drawing” file, DWG is one of the most common CAD (computer-aided design) formats. The format is native to AutoCAD software, but is widely interoperable across other CAD software such as MicroStation. DWG files need to undergo complex interoperability procedures for integration into GIS environments like ArcGIS, while at the same time maintaining attribute and geo-reference information correctly. ArcGIS provides an interoperability extension that is useful in this procedure, but this extension does not solve attribute interoperability objectives well

Although individual and internal configurations within the data format widely varied across data contributors, Esri format was clearly the leading trend. As the maps below illustrate, file geodatabase was the leading submission format, as requested in the call for data.

Shapefile and personal geodatabase were second and third most common, respectively. The only non-Esri format submitted were DWG files, which were submitted for several municipalities within Langlade and Rusk Counties.



Map 04. Address Point Submitted Format



Map 05. Parcel Submitted Format

3.4.2 DWG File Interoperability Workflow

Fig. 13 outlines the workflow used by the project team in translating DWG datasets into GIS format.

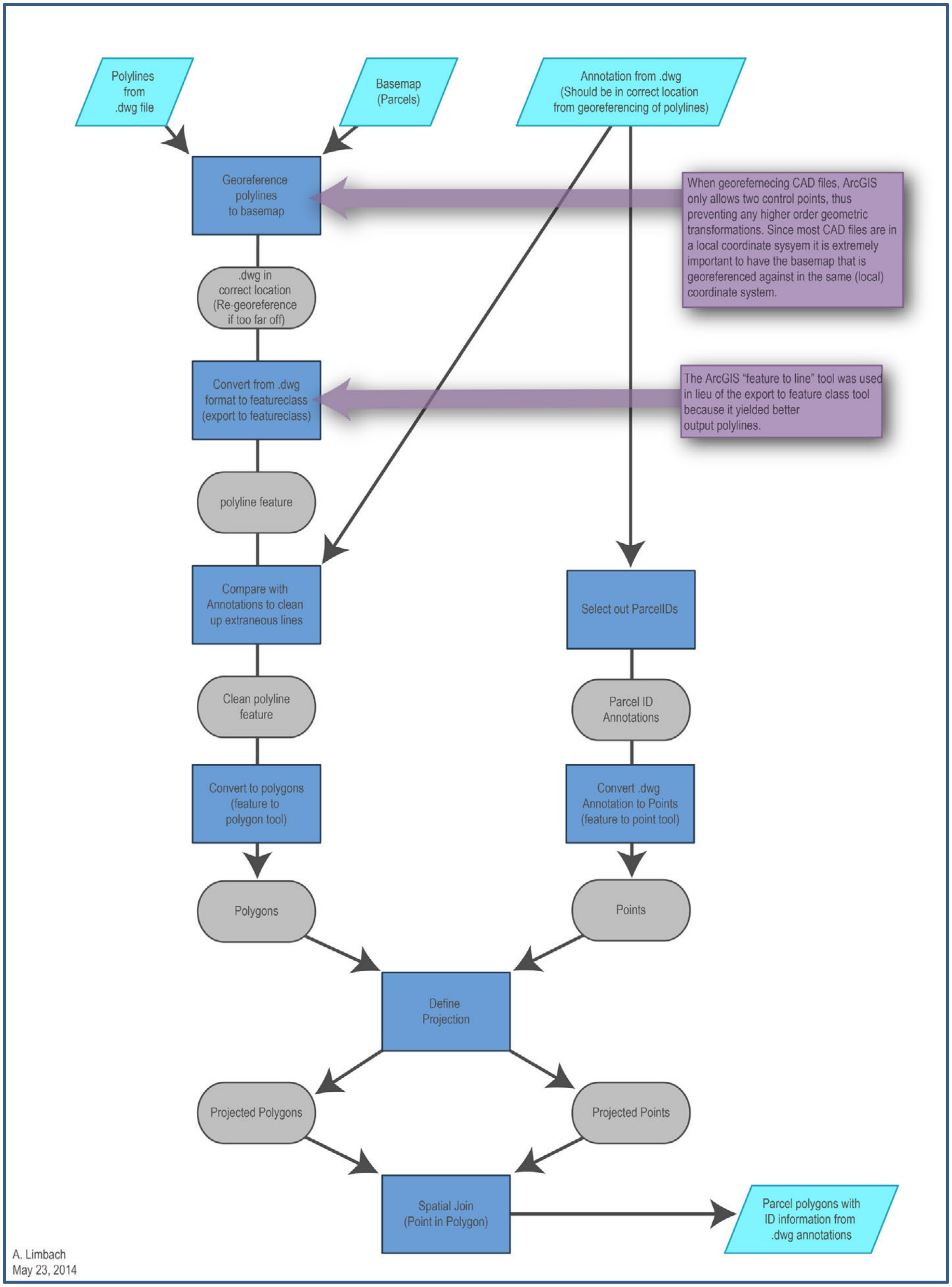


Figure 13. DWG Conversion to Feature Mode

3.4.3 Address Completeness

Address Completeness Scenarios

Upon the completion of the cleaning and standardization efforts, the team wanted a way to assess the completeness of a given address. To achieve this, they developed an address completeness assessment that looked at a number of key address components. The completeness scale was designed to aid the PSCW in detection of incomplete addresses.

Three essential address components were identified that must be present to represent a complete street address. There were two address scenarios containing these three elements representing a complete street address:

Scenarios for Complete Street Address		
Components		
Address Number	**+ Street Name**	**+ Street Type**
Address Number	**+ Prefix**	**+ Street Name**

The first scenario is common in urban and rural areas while the second scenario occurs most often with addresses along county and state highways and generally in more rural settings. The table below shows the different address scenarios and their values on the completeness scale.

Address Scenarios & Completeness Scale

Address Elements	Example	No. of Address Components	Completeness Scale Value
Address Number + Street Name + Street Type	301 Main St.	3 components	**3** - Complete address
Address Number + Prefix + Street Name	301 CTH K	3 components	**3** - Complete address
Address Number + Prefix	301 CTH	2 components	**2**
Address Number	301	1 component	**1**
Street Type	BLV, BLVD, BV, etc.	1 component	**1**
No address data	–	0 components	**0**

Address Completeness Tool

Assigning each point feature a value from 0-3 that would differentiate the less-complete addresses required the development of a Python tool to implement these values.

Python Tool & Address Completeness Scale		
Essential Element		**Value**
If point is a highway address	If point contains an address number	→ +1
	If point contains Prefix	→ +1
	If point contains Street Name	→ +1
	Complete Address	= **3**
If point is a street address	If point contains an address number	→ +1
	If point contains Street Name	→ +1
	If point contains Street Type	→ +1
	Complete Address	= **3**

By following this logic, the high score for an address point would be a value of three, meaning that the address point contains all three essential address elements. Values of one and two are degrees of incompleteness and a value of zero indicated that the address is incomplete altogether. The table below displays the percentage of points containing each possible measurement of address quality.

Final Address Point Completion Percentage	
Value	**Return**
0	0.16 %
1	0.12 %
2	1.04 %
3	**98.68** %

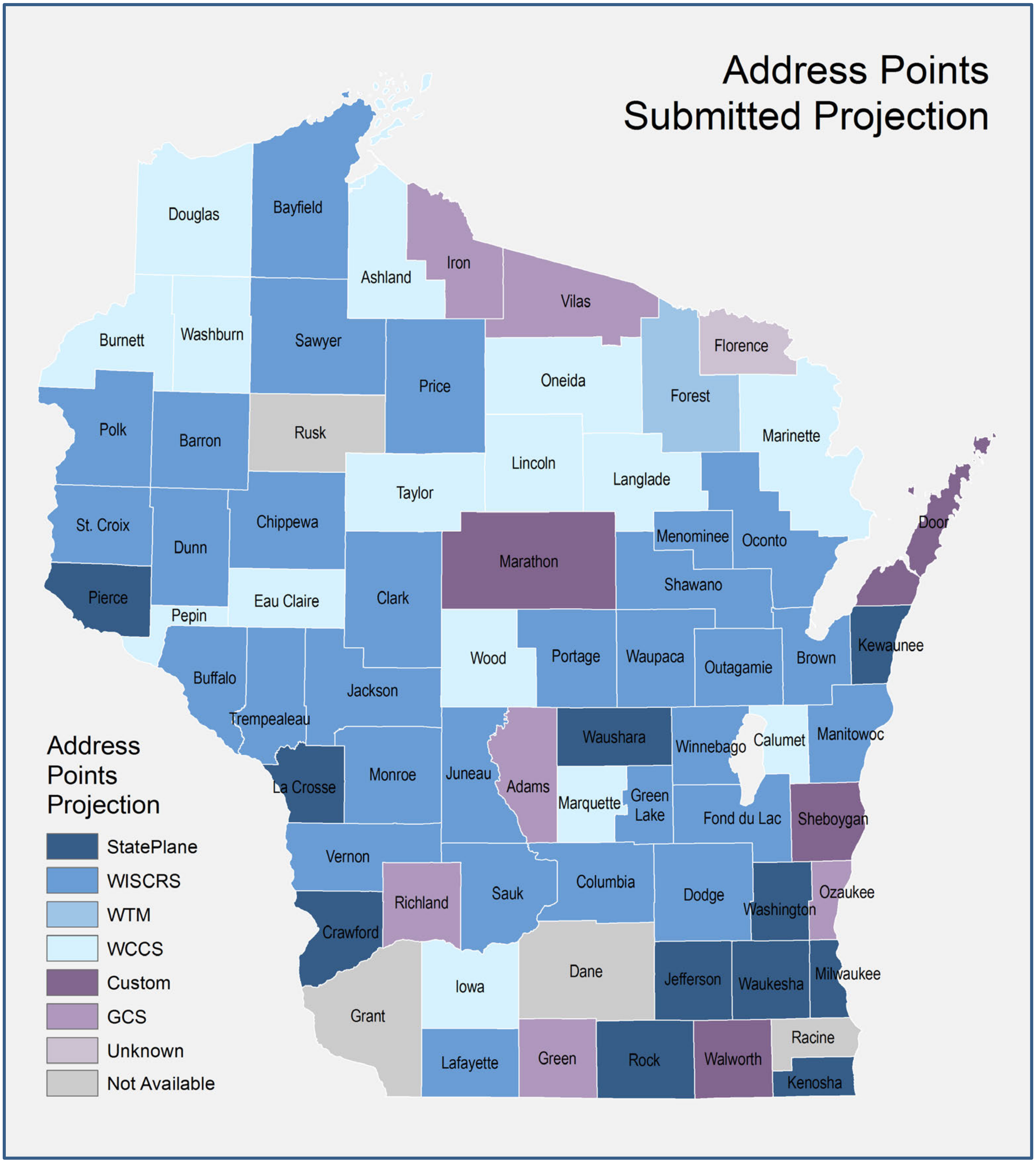
This statistic can be interpreted as a successful result, as 98.68% (2,705,261 points) of the final address points layer obtained a perfect score on the Address Completeness Scale.

3.5 Commonalities in Coordinate Systems

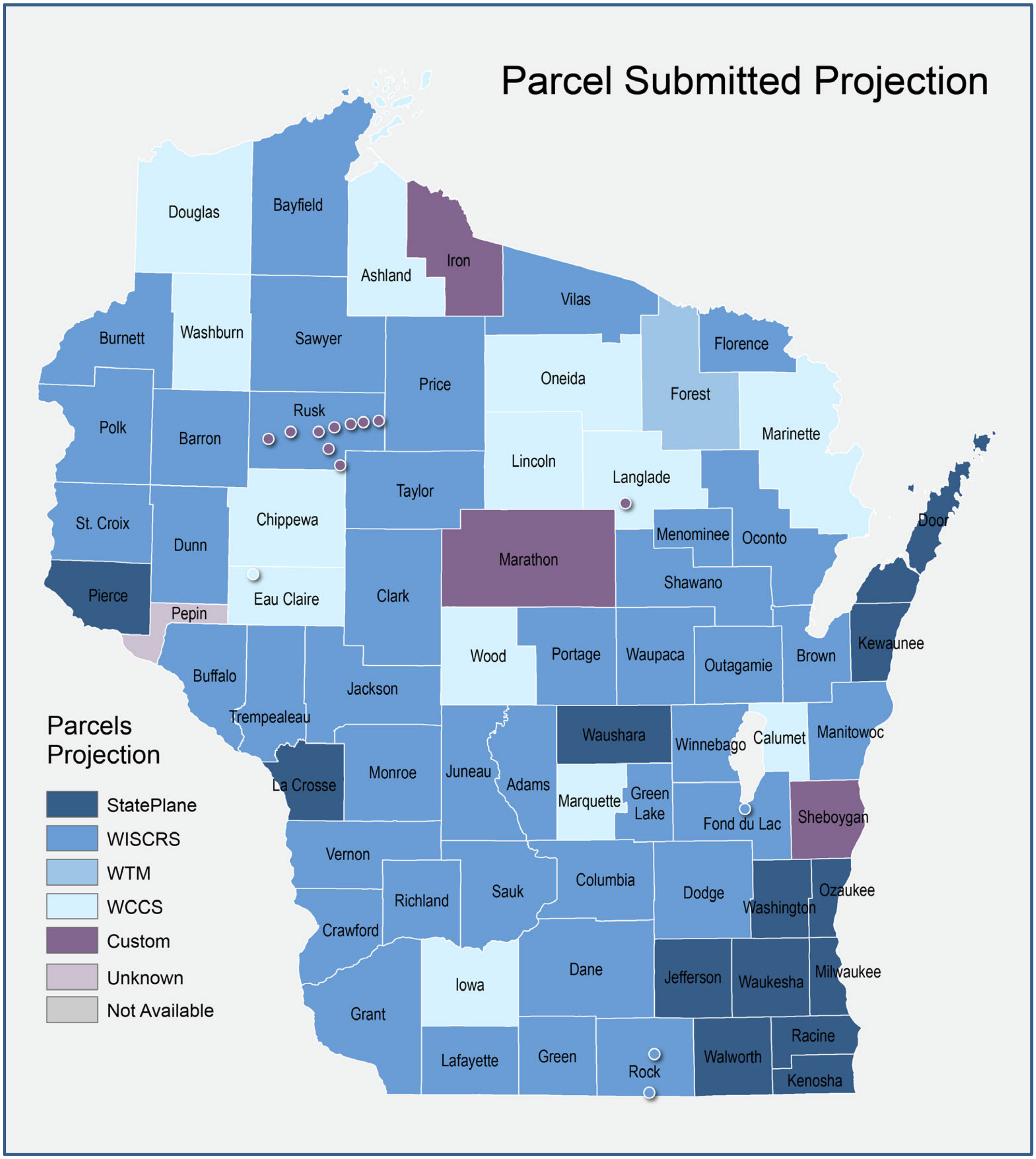
There are a variety of coordinate systems and projections used in Wisconsin, including:

- StatePlane – State Plane Coordinate System
- WISCRS – Wisconsin Coordinate Ref. System
- WTM – Wisconsin Transverse Mercator
- WCCS – Wisconsin County Coordinate System
- Custom
- GCS – Geographic Coordinate System
- Unknown coordinate system or projection

Maps 6 and 7 display the submitted projections for address points and parcels.



Map 06. Address Points Submitted Projection



Map 07. Parcel Submitted Projection

3.6 Community Anchor Institutions

In addition to the final address and parcel GIS layers, the LinkWISCONSIN Address Point and Parcel Mapping Project team worked on an additional task for PSCW to improve the spatial database for Community Anchor Institutions (CAIs).

The CAI database plays a fundamental role in assessing broadband connectivity of public institutions throughout the state. The CAI portion of the project was intended to improve on CAI data, focusing on four main objectives:

- Improve the spatial accuracy of features in the current CAI database
- Expand the current classification of CAIs from public schools K-12, libraries, health care facilities, public safety, higher education, government and non-government institutions to include correctional facilities, ports, and private schools
- Focus on authoritative sources to drive the CAI categories
- Build-out and update current CAI data to include missing CAIs and new CAIs

The task of incorporating the four major objectives listed above resulted in a workflow that employed numerous joins, checks, and validations. With completeness, precision, accuracy, and consistency a key focus, the team targeted methods for assessment.

3.6.1 CAI Workflow

One of the challenges associated with providing consistency and completeness in a deliverable is determining the metrics of inclusiveness for a given data layer. The team decided to adhere to authoritative sources (government or other officially-authorized sources) in governing inclusiveness, and thus aspects of completeness and consistency.

Another challenge of the project involved scrutinizing the spatial precision and accuracy of the CAI deliverable. The team focused on automating this process as much as possible, through sequences of validation logic. Nonetheless, a degree of manual research and validation was required (through cross-referencing) for portions of the deliverable.

Locating Authoritative Sources

The first step in the CAI process was locating and acquiring authoritative sources for those CAI categories to update and build out. Because address format often differs from one authoritative source to the next, the team had to be creative in how they tabularly joined these datasets to our spatial data. Joins were completed using a variety of different join fields including address, CAI name, and phone number, among others. Address standardizations were a key strategy to increasing join success.

Spatial Accuracy

Given the importance of the spatial accuracy associated with CAIs, the team next focused on ensuring they had attached the most accurate information. They lumped these levels of spatial accuracy into 3 different groups and assigned 3 different verification numbers to help focus efforts during the validation phase.

The first attempt was to join each category with the Homeland Security Infrastructure Program (HSIP) layer,[12] where HSIP was applicable. HSIP is an infrastructure geospatial data inventory. According to the Department of Homeland Security, HSIP compiles geospatial data from federal agencies, commercial vendors, state, and local partners for common use by Homeland Security; Homeland Defense; and Emergency Preparedness, Response, and Recovery communities. These datasets allow for nationwide infrastructure information access to assist decision makers in analyzing threats (whether natural or manmade) and modeling for emergencies and other missions. The HSIP datasets were acquired from the Wisconsin Department of Military Affairs.

[12] http://www.dhs.gov/infrastructure-information-partnerships#2

CAI Spatial Accuracy Measures	
Value	**Type of Address**
1 – Highest Degree of Spatial Accuracy	HSIP Spatial Location
2 – 2nd Degree	Project Address Point Spatial Location
3 – 3rd Degree	Geocoded Address Spatial Location

Those successful records were given the HSIP spatial location and a verification number of 1 and moved to the pre-spatial relator bucket.

The second attempt was joining to the project Address Point database. Successfully joined records were given the project Address Point spatial location as well as a verification number of 2.

The records that we were unable to get to join with either the HSIP or project address point database had their full address geocoded, spatial locations attached, and were given a verification number of 3.

Fig. 14 indicates the three degrees of spatial accuracy in light blue.

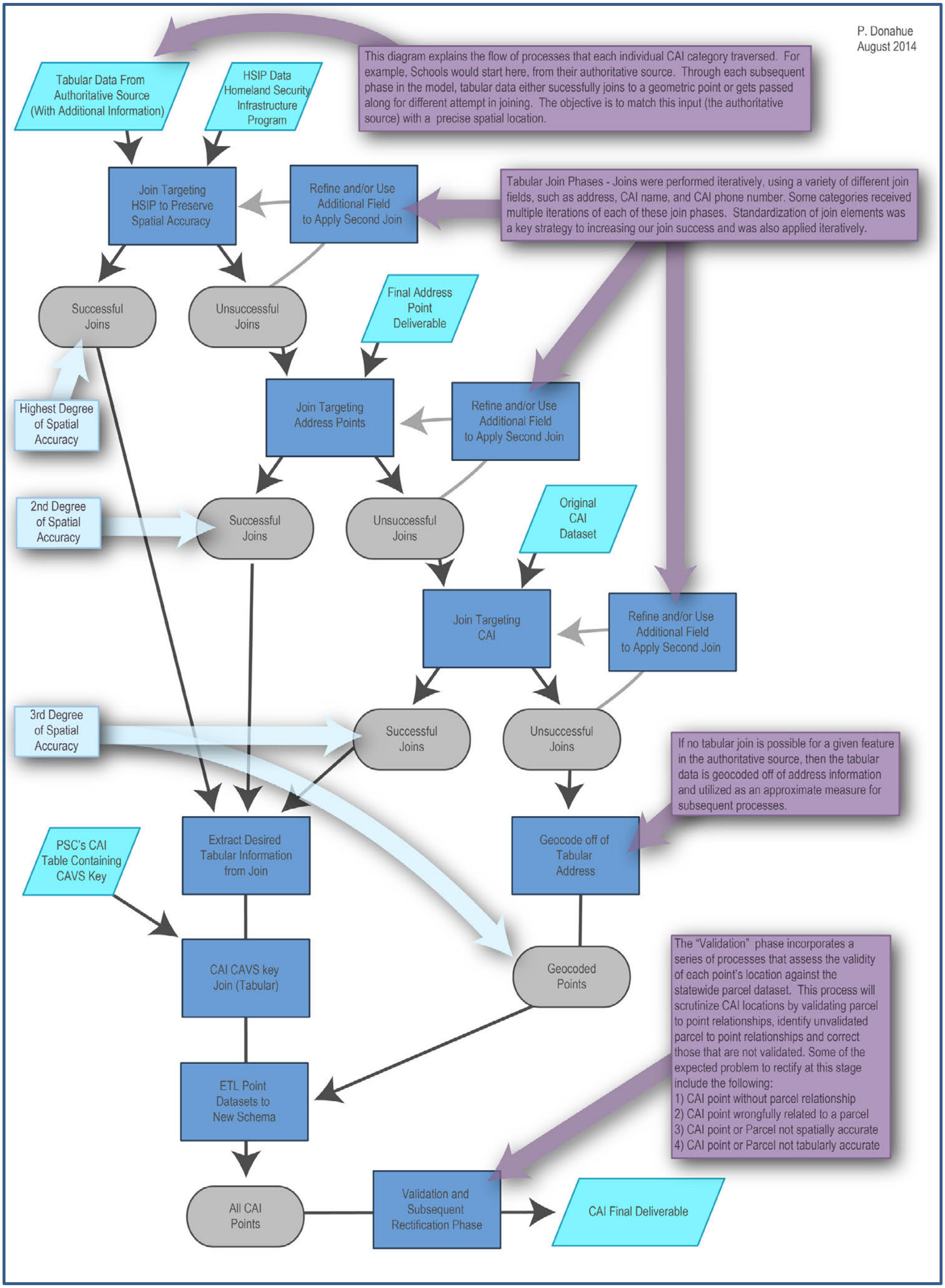


Figure 14. CAI Workflow

Upon assigning spatial locations and verification numbers to each point, the team focused on joining the data back to the PSCW's CAI table with PSCW's unique CAI identifier field, known as CAVSKey. The CAVSKey spreadsheet to obtain a given record's CAVSKey (which is referred to as "UID" after it has been attached).

After achieving this join and adding the CAVSKey, the dataset was ran through the ArcGIS Data Interoperability Extension work bench for two reasons:

- To map native data fields to new CAI schema
- To acquire Parcel ID and Relate ID associated with each point

CAI Categories

The final CAI database consisted of ten categories, each category includes the institutions defined by its respective authoritative source or sources.

CAI Categories and Authoritative Sources

Category Name	CAI Number	Authoritative Source(s) Used	Completeness Scale Value
Schools K-12	1	HSIP, Department of Public Instruction	Alternative, Charter, CESA, Corrections, Office, Special Education, Traditional
Libraries	2	Department of Public Instruction	Library, Branch, System
Health Care	3	Department of Health Services	Hospice, Hospital, OutPatient Rehabilitation, Nursing Home, Rural Clinic
Public Safety	4	HSIP, WI Law Enforcement Directory, Department of Safety and Professional Services	Fire Dept, Career, Mostly Career, Volunteer, Mostly Volunteer, Emergency Communications Center, Police Dept, Sheriff's Office, State Law Enforcement, Tribal Law Enforcement, University Police
Higher Education	5	US Department of Education	2 Year Private, 2 Year Public, 4 Year Private, 4 Year Public, Technical Private, Technical Public
Government	6	HSIP, WI Circuit Court Directory	Air Force Reserve Center, Airport, Army Reserve Center, Service Center, State Office, Tribal Office, Courts
Non-Government	7	N/A (Category Contains Miscellaneous or Retired CAI Points)	Museum, Park, Tourism Bureau, Historic Site, Heritage Center
Department of Corrections	8	Department of Corrections	Adult Institution, Correctional Center, Juvenile Facility, Women's Correctional System
Ports	9	HSIP	Air, Marine
Private Schools	10	HSIP, Department of Public Instruction	N/A

3.6.2 CAI Validation

After successfully aggregating all of the data into one master CAI layer, focus was shifted to validating the spatial location of the points. It was during this phase that the verification number associated with each point played an important role. Points with a verification number of 3 received the most validation scrutiny while points with a verification number of 1 were considered to have a fairly accurate spatial location.

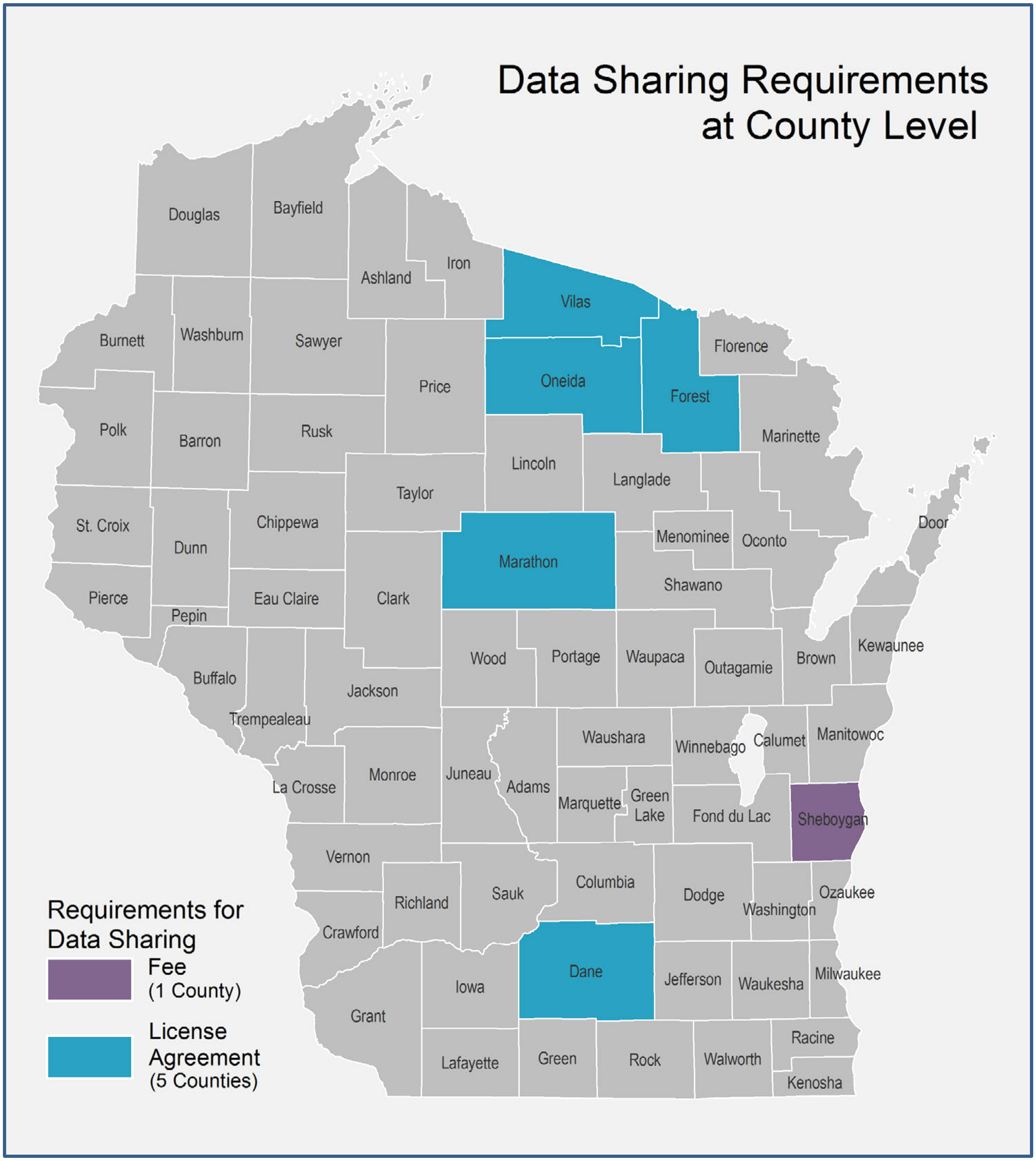
A number of different methods were involved with validating CAI point location and ensuring that points reside in the appropriate parcel.

CAI Validation Methods	
Comparison of CAI Name with the OwnerName of the parcel it resides in	If a match was received, that respective point was marked as validated.
Locus search comparison involving a point's CAI Name with the parcels within a certain radius of the point	Helped determine if a parcel in the area surrounding the CAI point included an OwnerName that would yield a match. If so, the parcel was flagged and the point was manually moved into its respective parcel.
Attempted address match of the CAI point to any project full address in a given surrounding radius	Did not directly involve comparing an associated parcel, but helped to narrow down the area of which a given CAI point is supposed to reside. It helped speed up the manual process of correcting a CAI point's spatial location.
Manual validation of a CAI point's location through the use of aerial imagery for visual location validation	Most time intensive method, but necessary when the above methods were exhausted and automated validation was not possible.

4 CHALLENGES & LIMITATIONS

4.1 Data Sharing

Data sharing restrictions on geospatial data within and amongst government entities were one of a few obstacles encountered in the LinkWISCONSIN Address Point and Parcel Mapping Project.



Map 08. Data Sharing Requirements at County Level

4.1.1 License Agreements and Fees

The five license agreements DOA received were reviewed by both a DOA attorney and UW-Madison legal. Alterations had to be made to the liability clause for DOA to be able to sign two of the agreements. This process was informed by the experience of the Robinson Map Library and other state agencies, where staff had experience dealing with GIS data sharing license agreements. The table below summarizes the license agreement modifications made for this project.

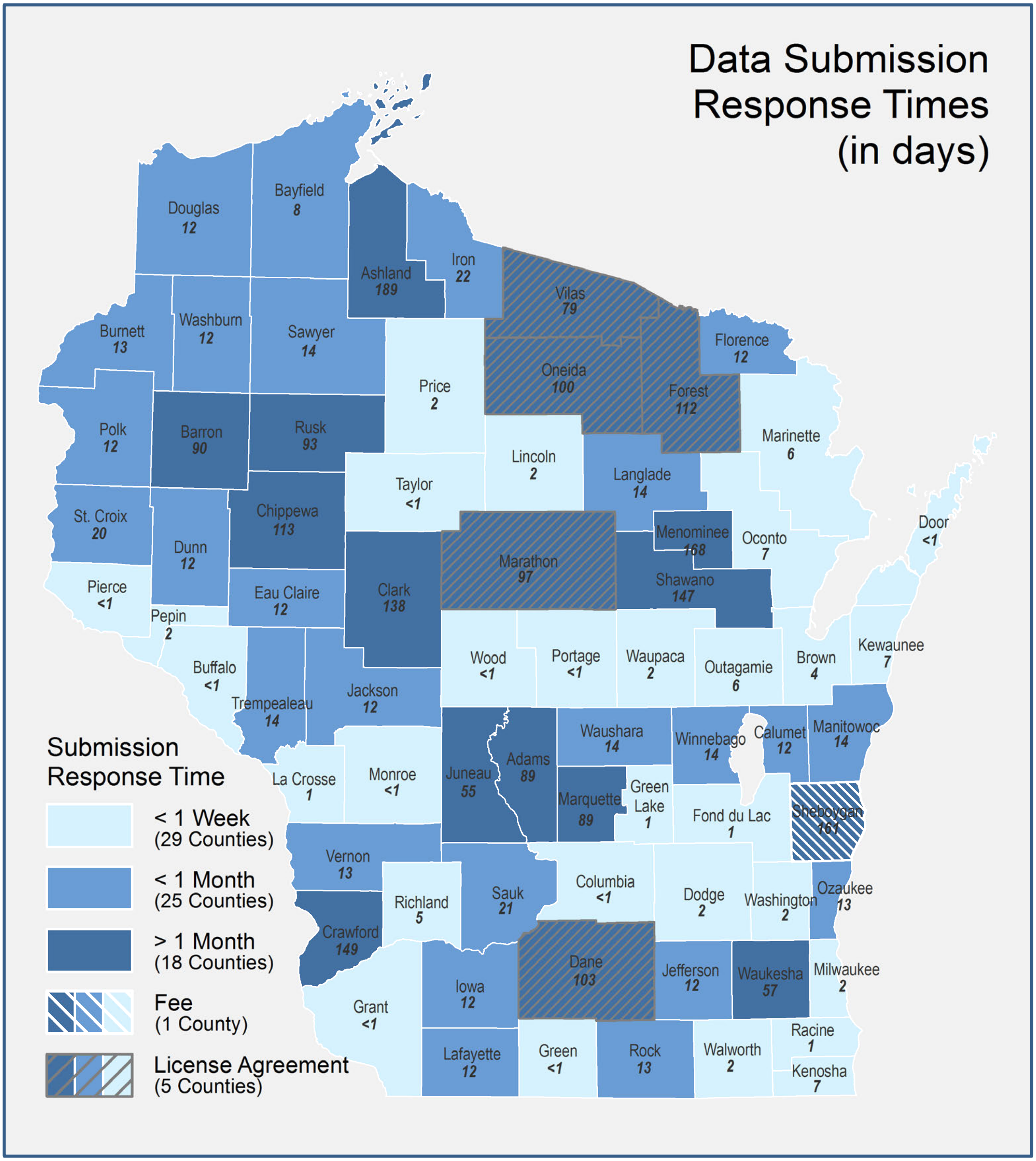
License Agreement Summary

County	Notes
Dane	Signed as-is
Forest	Modified the original agreement by adding a paragraph on prior written consent, expressly authorizing DOA to share data with the project partners at SCO and PSCW; Struck fee
Marathon	Altered indemnification clause: • Struck "indemnify" and "and defend" from "Licensee agrees to ~~indemnify~~, hold harmless ~~and defend~~ __ county." • Appended at end of liability clause: ", where such liability is founded upon or grows out of the acts or omissions of any of the officers, employees or agents of the state of Wisconsin and its private contractors including the University of Wisconsin System while acting within the scope of their employment and where such protection is afforded by ss. 893.82 and 895.46 (1), Wis. Stats."
Oneida	Struck fee
Vilas	Altered indemnification clause, same as Marathon County above
Sheboygan	No agreement signed; fee only

4.1.2 Response Times for Data Submission

Sixty-four of the 72 counties submitted data within one month of the original call for data that was sent on November 13, 2013. This rate of submittal exceeded expectations, which were based on the 2013 WLIP Survey responses and past experiences of other state agencies. However, it was not until June 2014 that all available county and municipal data had been collected. It took seven months largely due to the need for follow-up data requests once the technical team had identified gaps in data submitted and for processing license agreements.

The processing of license agreements and one invoice for a small fee added to the amount of time needed to acquire the data, because agreements required DOA and UW-Madison legal review and briefing of the DOA Division Administrator before signature, as well as processing time on the county's end. Because the address point datasets are stewarded by county staff other than LIOs in some counties, in some cases acquiring address points took much more time than the county's parcel dataset.



Map 09. Data Submission Response Times

4.2 Gaps In County Datasets

For this project, a density calculation aggregated to boundaries of Census-designated places was used to identify places where parcels or address points were not found but expected. Once gaps were identified, follow-up requests were made to 11 counties and requests for data were made to six municipalities that were identified as stewards for address point or parcel data not incorporated into county datasets.

Independent Municipal Data Stewards		
Municipality	**County**	**Not Incorporated into County Dataset**
City of Antigo	Langlade	Parcels
City of Ashland	Ashland/Bayfield	Address Points
City of Beloit	Rock	Parcels
City of Eau Claire	Eau Claire/Chippewa	Parcels
City of Fond du Lac	Fond du Lac	Parcels
City of Janesville	Rock	Parcels

4.2.1 Gaps Remaining in Address Point and Parcel Data

In some local cases, digital data simply does not exist. For example, the 2013 WLIP Report estimated that 3.4% of parcels in Wisconsin are not available in digital format. The team identified 43 cities and villages that were missing address point data and 22 cities and villages missing parcel data, as depicted in the table below and the images on the next two pages.

City and Village Gaps in Data Submitted	
Address Points	**43** do not exist or stand missing in final deliverable
Parcels	**22** do not exist or stand missing in final deliverable

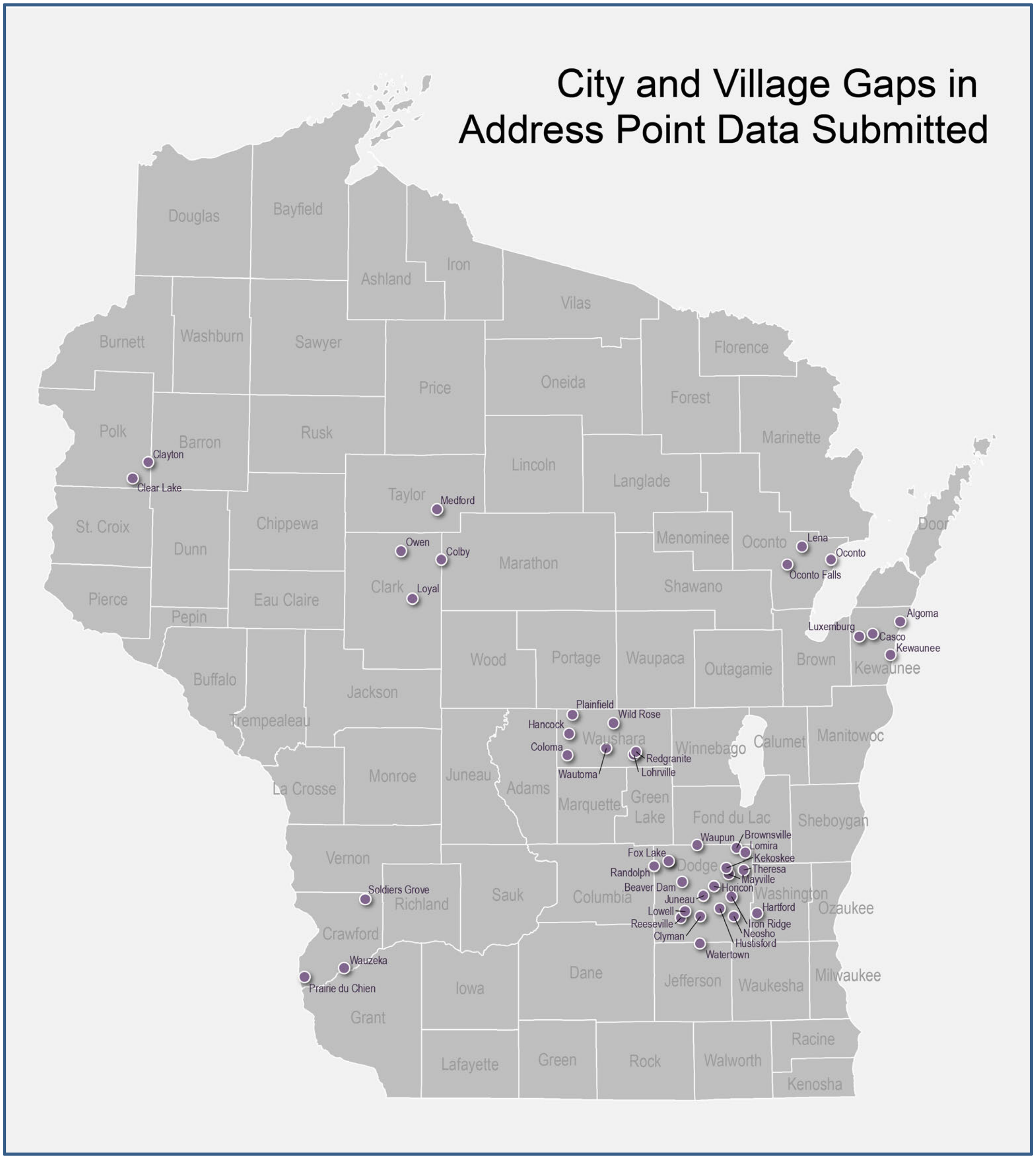
Address points were not available across the entire county for Dane, Grant, Racine, and Rusk counties. About 12 counties have incomplete parcel coverage.

Where address points were not available, but digital parcel data exists, the team substituted parcel centroids for address points. This was done by creating parcel centroid points for all parcels within a given jurisdiction absent of address points and then removing the centroids that did not have any associated property address. Although these points were not as spatially precise as actual address points, they were the best available information that could be obtained for this project.

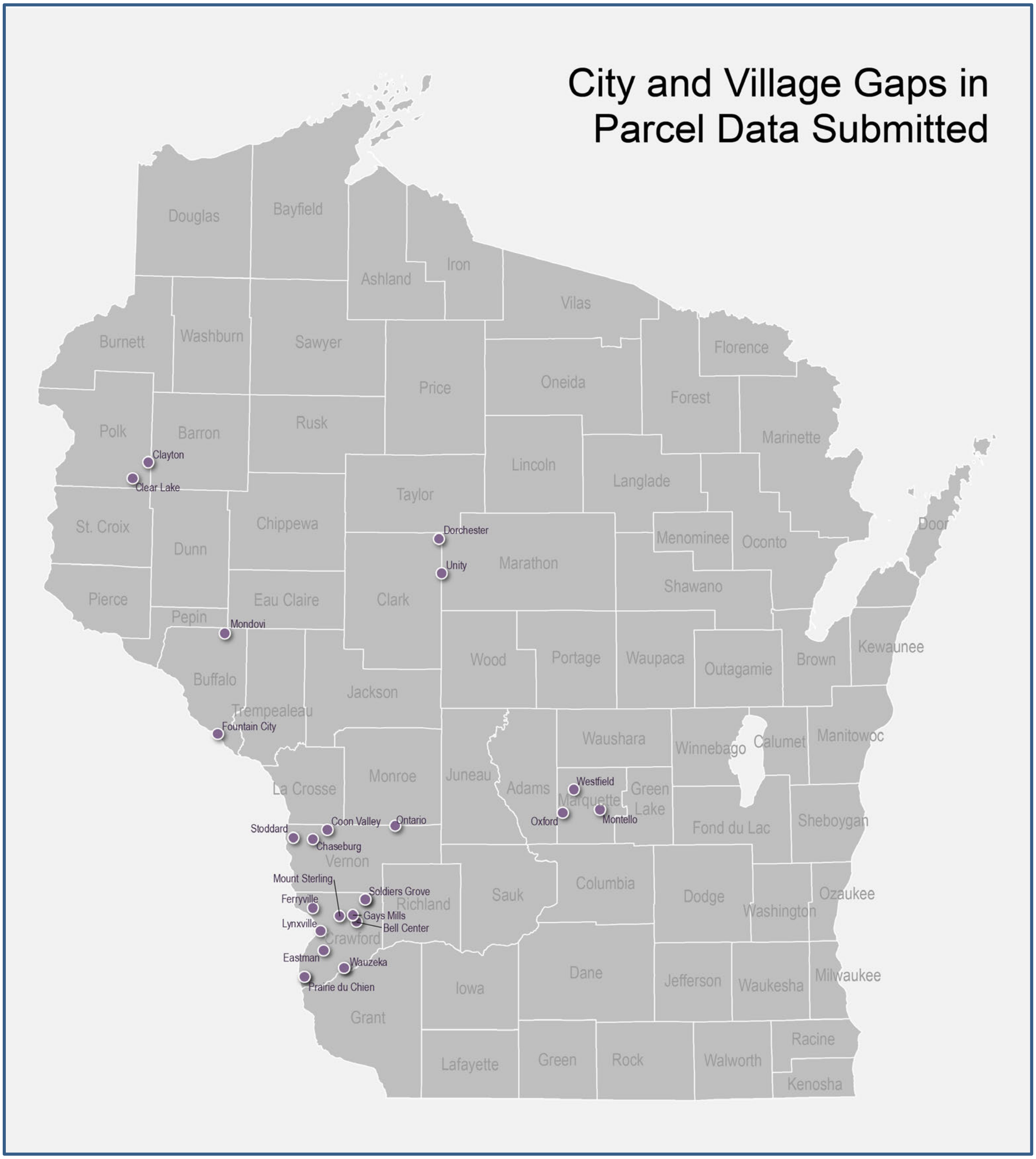
While 86.4% of features provided in the deliverable came directly from local-level address point layers, 13.6% of the address point deliverable was comprised of parcel centroid data from county or municipal sources.

Other noteworthy characteristics of the address point deliverable include:

- 0.2% of address point data was constructed from municipal-level parcel data (as parcel centroids)
- City of Ashland address points were calculated from building footprint centroids, making up 0.1% of the final deliverable



Map 10. City and Village Gaps in County Address Point Data

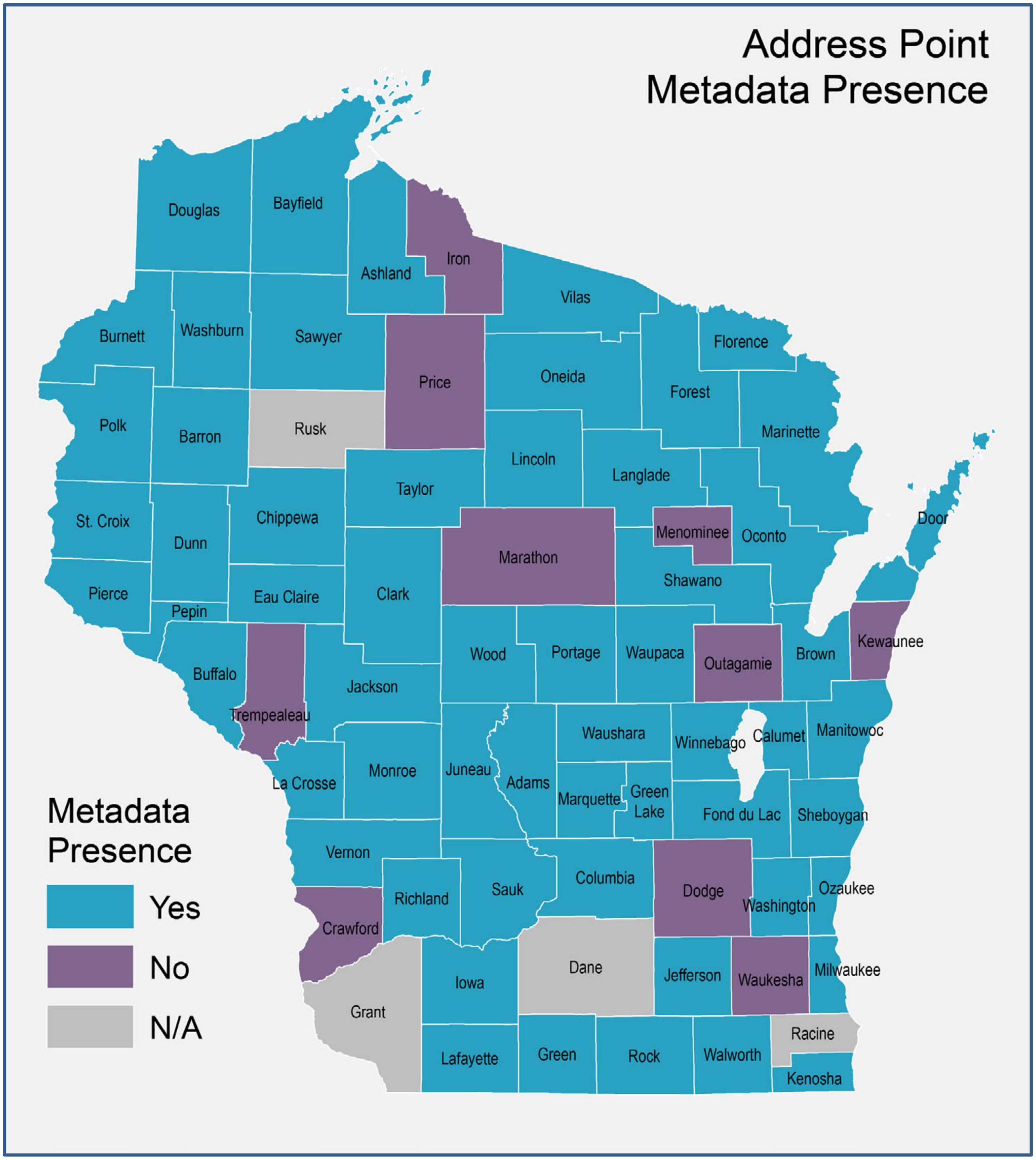


Map 11. City and Village Gaps in Parcel Data Submitted

4.3 Technical Hurdles

As with any project, there were several technical hurdles encountered throughout the course of the LinkWISCONSIN Address Point and Parcel Mapping Project. Parsing and concatenation challenges are addressed in Chapter 3. Here, technical hurdles are defined as issues with participating data, logic, or tools that called for unanticipated troubleshooting. They can range from large and complex to small and simple. Some of the most common technical challenges include lack of metadata and ETL tool weaknesses.

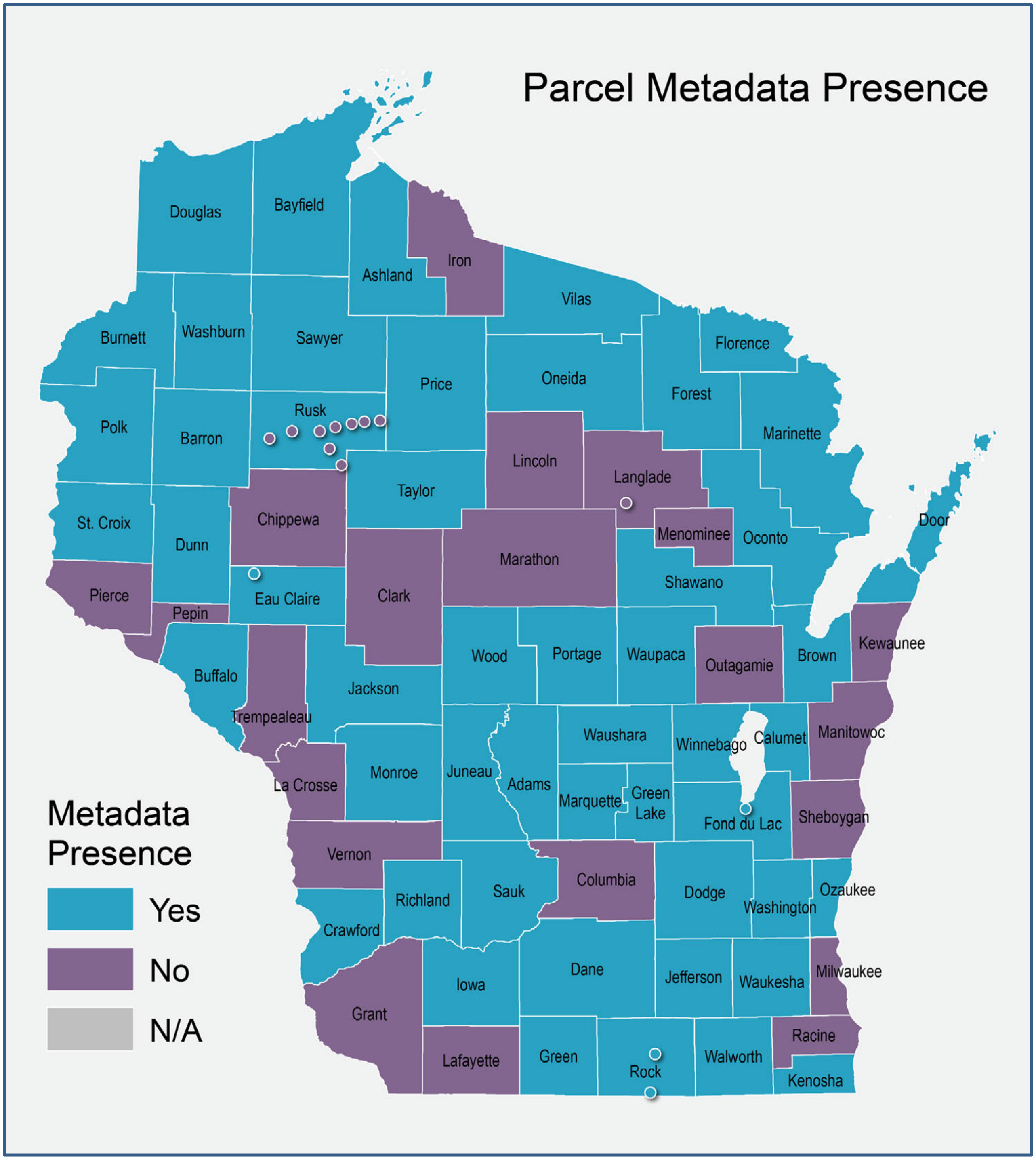
4.3.1 Lack of Metadata



Map 12. Address Point Metadata Presence

The value and importance of geospatial metadata cannot be over stated. In the simplest terms, metadata allows for documentation of the who, what, where, when, and how of the data. Taking into account the historic nature of cadastral data, this type of information is invaluable considering the number of individuals that have had a hand in updating, maintaining, and stewarding this data. Stewards of digital land information should be aware of practices for developing sound metadata standards, based on the FGDC resources at www.fgdc.gov/metadata.

During the project ingest phase, an assessment was made as to whether metadata was present or not. Maps 12 and 13 illustrate the presence or lack of metadata on a contributor level.



Map 13. Parcel Metadata Presence

To reiterate, metadata was only assessed as present or not present—no qualitative assessment of the nature of the metadata was made. The team's brief assessment suggests that metadata richness varies a great deal amongst contributors and would be an important place to focus future efforts when developing other statewide data layers.

4.3.2 ETL Tool Weaknesses

The Data Interoperability Extension was an effective tool for aggregating a large number of independent datasets into a single data model for this project. However, it did have a few drawbacks.

A major challenge was related to the loss of mappings and connections made within a given county tool after saving and reopening that tool at a later time. After consulting Esri customer service and doing some internet research of their own, the team uncovered that it was a bug in ArgGIS 10.1. Apparently this issue has been resolved in ArcGIS 10.2, although the team has not been able to evaluate and comment on this bug fix.

The other feature that was the cause of some frustration was an auto-connect feature within the workbench. Any attribute names that exist in the reader (input native dataset) as well as in the writer (output dataset) were automatically connected. Unfortunately it was not possible to turn this feature off. The team mitigated this issue by running a comparison of the final attribute schema against each individual contributor's attribute schema. When matching attribute names were discovered, the team created a new attribute field with an altered name, copied the matching field over, and removed the native field to prevent the auto-connection. This solution was not ideal and if this project were ever repeated, steps should be taken to add a unique character set to each output attribute field to prevent these types of auto-connection occurrences.

4.3.3 Other Technical Challenges

Processing of Custom Tools in Staging Database

Automated processing was crucial for this project due to the volume of data transformed. For context on the amount of processing the project required for the address records:

- There were 2.7 million records fitting the project's attribute schema of 39 features
- 2.7 million words would be the equivalent of reading *The Lord of the Rings* trilogy, 6 times
- 2.7 million records fitting the project's attribute schema of 39 features equals a total of 105.3 million individual elements

Size of Post-ETL Data

Another challenge that the technical team faced was the size of the post-ETL data. With the highest feature counts in Milwaukee County and Dane County (with 460,360 address points and 309,103 parcels, respectively) amongst all data contributors, county-level data manipulation was easily handled with machines running mid-grade specifications (Intel Core i7-3770 Processor, quad core, 8 thread, 3.40 GHz frequency with 8GB of RAM). The team experienced performance issues when working with the aggregated statewide layers which were roughly 72 times the disk size and feature count of the average county dataset.

To remedy the problem of dataset size, several actions were taken:

- Establishing an ArcSDE service was considered for containing the point and parcel layers but the project team established that the lack of flexibility, initial overhead, and lack of need for versioned editing as determinants against pursuing this option. The project team pursued the file geodatabase as the file format for implementing state-level logic and QA/QC, which was consistent with the deliverable and staging database formats.
- The project team was able to invest in an additional 4GB of RAM for all machines, bringing them from 8 to 12GB RAM each. This played a significant improvement in processing time and rendering performance. The team sought out a machine with 32GB of RAM for execution of the state-level logic, which was highly processor intensive.
- There are several methods that can greatly speed up geoprocessing of large datasets in ArcGIS. The project team utilized these strategies in order to perform the needed processing on 2,741,467 address point features and 3,723,392 parcel features.
- Python and Model builder were both leveraged to write custom geoprocessing and to string several processes into a single model. This approach was beneficial as it allowed the team to run processor-

intensive models over the course of a night or weekend while also allowing them to customize logic to directly and efficiently fit project workflow.

In addition to focusing and stringing geoprocesses, the technical team adhered to some general guidelines when processing on the statewide datasets, as elaborated in the next section.

Writing to the "In-Memory" Workspace

In lieu of writing geoprocessing outputs to a location on a disk, ArcGIS allows a user to write outputs to an in-memory workspace. Writing to memory versus to a disk location can significantly reduce geoprocessing time.

Writing to memory is not difficult. The technician simply replaces what would otherwise be a path to an output directory with the "in_memory\" path syntax, appending an output name to this path and omitting any file extension. There are some caveats to be aware of when working with the in-memory workspace. A few key tips include:

- Data written to memory is only available for subsequent processing until the application is closed.
- Writing to memory will effectively consume a machine's RAM, or temporary storage areas. RAM is classified under two types, physical RAM and virtual memory. Your machine will consume physical RAM until it is used up, at which point, a machine will then create virtual memory by consuming hard drive space and trying to use it as if it were real (physical RAM) memory. When this point is reached, you will likely notice slower geoprocessing because virtual memory is much slower than physical RAM. Upgrading your machine's physical RAM may be something to consider if you plan to routinely process large datasets.
- Note that there is a point where the advantage of using the in-memory workspace will grant a diminishing return. This point is met when the computational enhancements are outweighed by the application slowdown inherent of using virtual memory.
- If a geoprocessing model consumes the entirety of your machine's RAM (physical RAM and virtual memory), the process will fail. But there are steps that you can take to increase chances that the process will complete successfully. Increasing your machine's physical RAM can be a costly and time consuming option, but you can increase virtual memory easily by increasing your machine's paging file size to any amount available (open disk space). The paging file allocates a certain area on your hard drive to be used as if it were RAM. Virtual memory is slower than physical RAM but increasing your paging file can help ensure that your in-memory processes complete successfully.
- It is good practice to clear outputs from memory after the completion of intermediary processes within a model. To do this, utilize the "Delete Features" tool within model builder to delete old outputs and free-up RAM.
- Avoid using other memory-intensive programs while performing memory-intensive geoprocessing.

Scrutinize Data Format and Geometries

The nature of data that is plugged into geoprocessing can also play an important role in the outcome or progress of the process. Some things to think about when scrutinizing data inputs include:

- Removal of non-needed attribute fields can reduce file size, speed up processing, and free up RAM.
- Be aware of features with an excessive amount of vertices when processing a dataset, they can cause the process to hang up or fail.
- Check the feature class for bad geometry. Bad geometries can hang-up a process or cause it to fail. The "Check Geometry" and "Repair Geometry" tools can help identify and solve problems of this nature. One common example of a bad geometry includes tabular features without corresponding geometries.
- Consider the spatial reference of inputs. In order for spatial relationships to be applied across two or more datasets, a common coordinate system must be established. In order to achieve this, a re-project must be performed so that all participating inputs meet the projection of the primary input. This re-project can consume time and resources and can also affect the output result.
- Avoid working across network drives and work on local disks when possible.
- File geodatabase is recommended for best performance.

5 BEST PRACTICES & RECOMMENDATIONS

5.1 Repeatability

The LinkWISCONSIN Address Point and Parcel Mapping Project was designed to be repeatable. To that end, each step of the process was documented thoroughly so as to preserve the broad, statewide process the project team administered, and the data sharing lessons learned. Through documentation, custom tools, stored tools and models, the systematic approach taken is available to advise similar projects in the future.

It is important to note that repeatability does not necessarily mean a future project would entail full automation. It would be feasible for the project team to create a tool that would process all data from start to finish with the push of a button, but such a workflow would greatly compromise the integrity and utility of the deliverable. Due to the variation in contributing datasets, an important degree of human knowledge must be applied to each dataset, even for a second iteration of the project process. For this reason, the project's repeatable processes can be broken down into two types—documented procedures and stored procedures.

5.1.1 Documented Procedures

Each contributing dataset contains a unique configuration of address point and parcel data. While the these county-level inputs will likely stay reasonably constant through time, there are essential decisions that need to be made in a non-automated fashion. Such examples include:

- Resolving domains
- Removal of extraneous content from attribute data (such as "Private" from "Matthew (Private) St.")
- Removal of extraneous content from spatial data (such as data that reside outside of jurisdiction)
- Custom (various kinds) of parsing;
    - Street Name from Street Type
    - AddNumPrefix from AddNum within a grid address
    - Full Address parsing into respective elements fitting
    - Parsing prefix (USH) from Street Name (151) and properly identifying/categorizing these Highway address elements
    - Several others

5.1.2 Stored Procedures

In contrast to documented procedures, stored procedures are those that can be modeled either through a script or a model with their parameters stored. There are several tools created and used throughout the project that contain stored parameters and, in theory, can be run again as long as the input data does not change. Stored procedures include:

- All ETL tools
- State-level logic model
- Standardization/normalization model

By following the documented procedures and the stored procedures set forth by this project, a more efficient iteration of this project's workflow could be achieved, should a similar project be undertaken in the future.

5.2 Recommendations

This project provided many valuable insights pertaining to aggregating local level datasets into a statewide data layer. The recommendations contained in this section are based on observations and interpretations made by the project team, and are divided into recommendations for counties, state agencies, and general lessons learned.

5.2.1 Recommendations for Counties and Municipalities

Maintain and Distribute Current Metadata

As discussed in Chapter 4, the value of metadata cannot be overstated, especially when it comes to parcel datasets. Robust metadata can greatly assist individuals using the data and ensure that any qualifications or restrictions are well known to the user. Issues of cryptic attribute names and coded domain descriptions or values can be defined within the metadata and can help to ensure that a user is not making incorrect assumptions when trying to use the data. At a minimum, it is strongly recommended that metadata include the minimum essential components below.

Minimum Metadata Essentials	
Citation	Information to be used to reference the dataset
Description/Abstract	Characterization of the data set, including intended use and limitations
Time Period of Content	Time period to which the data set corresponds
Status	State of and maintenance information for the data set
Spatial Domain	The geographic areal domain of dataset including bounding coordinates, coordinate system, datum, units of measure
Keywords	Words or phrases summarizing an aspect of the data set
Access Constraints	Restrictions and legal prerequisites for accessing the data set
Use Constraints	Restrictions and legal prerequisites for using the data set after access is granted

A data dictionary is also essential. Data dictionary contains information on a dataset, such as full attribute names, meanings of codes, scale of source data, and accuracy of locations.

Maintaining and updating this information helps not only the internal users of the data, but also can help prevent the misuse of the data by external users. A well-defined explanation of a dataset and its limitations and restrictions can help to prevent misuse and misrepresentation of what an individual dataset can provide.

Look Towards Standards and Best Practices In Addressing

The workshop offered at the 2014 WLIA Annual Conference, *Building an Address Repository Using the FGDC Standard: Implementing Data Quality and Data Sharing* provided many valuable insights about addressing and the best ways to build and maintain an address repository. The team strongly recommends that managing entities look to and consider similar guidelines when building and maintaining their own address repository and assigning new addresses.

Considerations for addressing should focus on:

- Consider all users who could benefit from central address repository, focusing on flexible use between different user groups or governmental departments
- Segment and isolate individual address elements
- Focus efforts on normalizing data across address point and parcel datasets
    - Construct relationship tables for fields with common domains (e.g., Street Name, Street Type, Prefix, Suffix, PlaceName)
- Maintain retired addresses in a repository as "retired"

Such standardization efforts could enhance the efficiency of any future efforts to build or update statewide aggregations of local address data. Processing local address data into a consistent schema was one of the most time-consuming aspects of this project. If local governments moved toward a standard encompassing the recommended elements listed above, statewide integration efforts could be done more quickly and at lower cost.

Understand, Formalize, and Maintain Governing Rules for Addressing

These rules provide a standard process for assigning new addresses and also define the method for maintaining all address elements. Governing rules for addressing vary on a local basis. Local data stewards should take to formally document the rules used for address assignment, to ensure consistency over time.

Maintain a Link Between Address Points and Parcels

In many cases this would be the parcel ID associated with the address, but this could be just a unique property ID that is maintained that allows for the linkage of the two layers. A link like this can eliminate some of the vagueness that can occur when address points fall outside of their associated parcel geometry (e.g., in the right-of-way, in a stream or lake, as spatially misplaced points, etc.).

5.2.2 Recommendations for Local GIS Data Acquisition

As a statewide aggregation effort, this project yielded many observations that can potentially inform GIS data efforts for state agencies and those working on statewide mapping initiatives in the future.

Focus on Outreach and Communications

Efforts made to explain a given project and its benefits can go a long way in encouraging contributor participation. Explanations of how a project has benefits for the entire state, as well as regional and local areas, can help build public support and could encourage unconvinced stakeholders to participate. A degree of transparency and a willingness to provide project updates and progress maps can also help boost participation.

Local governments, as possible project participants, are likely to first judge a project by the way it is presented and how the goals of the project are communicated. Professional communication is important to instill confidence in the project. Outreach helps to ensure that the message reaches all likely participants.

For this project, communication and outreach occurred *prior to* the call for data, including a one-page project announcement, conference calls with WLIA and LION representatives, a Webinar, the launch of a project Website, and a conference presentation. These outreach measures conveyed the value of the project to potential participants and likely moved some that may have been non-responsive or reluctant to participate.

Build Relationships with County and Municipal Participants

Communication should go both ways. In order to convey potential benefits to local governments and speak to their concerns, one must first understand the participants and their contexts. Even though the technology exists to share data without human interaction, personal relationships play an important role in the willingness for local data stewards to participate. Personally interacting with municipal data stewards and land information officers, in such practices as attending Land Information Officer Network meetings, making phone calls, and visiting counties is important to gaining confidence.

Make Data Requests Precise and Submission Easy

Data stewards are more likely to contribute when requests for data are concise, specific, and provide clear directions. The focus here should be on minimizing the effort needed to understand and respond to the data request. Make it straightforward and as easy as possible.

The LinkWISCONSIN Address Point and Parcel Mapping Project team used an upload widget from a company called Box that was placed on the project website. Contributors simply had to zip their data in a file folder, drag and drop it into the widget, and provide an email address in order to upload data.

Upon receiving data, a brief inspection and assessment should be undertaken, so that if subsequent requests are required, those follow-up communications can happen shortly after the initial response to the data request.

Emphasize the Benefits of Participation

In communications about the project, emphasize the benefits of participating to the state as a whole and to the local government in particular. This may require designing the project so that there are specific deliverables or ways in which a county will benefit from participating.

In this case, the project team made clear that there were statewide benefits to the project, including the enrichment of broadband-related mapping capacities at the PSCW and serving as a source of information for the upcoming Statewide Parcel Map Initiative. Benefits to counties included the potential for PSCW to better identify underserved areas for broadband in some counties and a free addressing workshop as part of the project.

County Assessment and Observation Reports were a major way the team provided counties with feedback for the LinkWISCONSIN Address Point and Parcel Mapping Project. These reports were individualized assessments of county address point and parcel datasets, unique to each county. A sample appears in Appendix C.

Encourage Participation

Those who are hesitant about getting involved in such a project may be assuaged when learning about the participation of others who are similar to them. Therefore facilitating communication between potential contributors is an important aspect of project success. For example, creating a status map of participation helps inform potential contributors about how others are responding. In addition, it is important to listen to the concerns of potential contributors and address these when possible.

Persistence Pays

Even if the benefits of participation seem straightforward, the risks minimal or non-existent, and the method easy, it may still be difficult to garner support or participation from some data stewards. This is why it is important to send follow-up communications after the original request, to follow emails with phone calls, and even make site visits. Taking the time to travel across Wisconsin to meet someone in-person conveys the value of the project.

It is imperative to first exhaust efforts at collaboration with staff who are supposed to be able to handle your request, before moving up the decision-making ladder to another individual in the hierarchy who might influence the outcome. It is also important to document any repeated requests for data and the individuals you have dealt with, so that when dealing with a new staff person or administrator, you can reference your previous efforts.

5.3 A Collaborative Statewide GIS Success

This project can be interpreted as a success not just based on the final deliverable of statewide GIS layers, but also on the process itself. By bringing together GIS experts from the PSCW's broadband mapping team, the State Cartographer's Office at the University of Wisconsin-Madison, county land information offices, and WLIP staff, the project team was able to achieve true intergovernmental coordination, and a final result that would not have been possible otherwise.

This project was completely dependent on collaboration from Wisconsin's counties and municipalities. All known existing address point and parcel datasets in the state were contributed for this project, a truly unprecedented measure of GIS data sharing between state agencies and local governments which defied expectations, and bodes well for creating statewide GIS layers in the future.

Lessons Learned

This project shows that the “just do it” approach to initiatives such as creating statewide maps layers can have long-term beneficial effects. In 2012, the WLIA Parcel Team aggregated parcel datasets to make a contiguous parcel layer viewable on any desktop or smartphone, a proof-of-concept that provided motivation for DOA to work with PSCW to get the LinkWISCONSIN Address Point and Parcel Mapping Project off the ground.

Persistence is necessary at all phases of a project of this size, and the initial stages of getting a project off the ground can sometimes be the most difficult. While the original DOA portion of the project had been approved by the PSCW and National Telecommunications and Information Administration in 2010, DOA lacked a project plan and pathway to completion. It took a year of collaboration between PSCW, DOA, SCO, and the WLIA Parcel Team to create a project plan and arrive at the MOU between DOA and PSCW to begin the project in July of 2013.

There are important implications that stem from the achievements of this project. Not only was the PSCW provided with the address point and parcel layer deliverables, but other state agencies also can, and will, benefit from the lessons of this project. Perhaps most significantly, the DOA effort of the Statewide Parcel Map Initiative will be able to build from the experience and the success of the LinkWISCONSIN Address Point and Parcel Mapping Project.

APPENDICES

Appendix A. Address Point Schema Definitions

Legend	
ElementName	Denotes database field name
(Element Name)	Full database field name
[FGDC]	Denotes database field name that is equivalent to FGDC element name
[FGDC: <element name from FGDC standard>]	Denotes field name translated to FGDC element name
[LinkWISC]	Denotes database field name distinct from FGDC that serves LinkWISCONSIN project needs

RelateID (Relate ID) *[LinkWISC]*
This string field contains a 10 digit key which correlates with the "RelateID" that exists in the project parcel layer. This field will facilitate joins between the project Parcel Layer and the project Address Point Layer. Multiple points residing inside one parcel are given the same identification value.

ParcelID (Parcel ID) *[FGDC: Address Parcel Identifier]*
Unique number assigned to a parcel by the local authority (can be tax roll ID number)

CountyID (County FIPS Code) *[FGDC: County Name]*
The FIPS code for each of Wisconsin's 72 counties. Codes are as follows:

ADAMS	001	JACKSON	053
ASHLAND	003	JEFFERSON	055
BARRON	005	JUNEAU	057
BAYFIELD	007	KENOSHA	059
BROWN	009	KEWAUNEE	061
BUFFALO	011	LA CROSSE	063
BURNETT	013	LAFAYETTE	065
CALUMET	015	*LANGLADE*	*067*
CHIPPEWA	017	*LINCOLN*	*069*
CLARK	019	*MANITOWOC*	*071*
COLUMBIA	021	*MARATHON*	*073*
CRAWFORD	023	*MARINETTE*	*075*
DANE	025	*MARQUETTE*	*077*
DODGE	027	*MENOMINEE*	*078*
DOOR	029	*MILWAUKEE*	*079*
DOUGLAS	031	*MONROE*	*081*
DUNN	033	*OCONTO*	*083*
EAU CLAIRE	035	*ONEIDA*	*085*
FLORENCE	037	*OUTAGAMIE*	*087*
FOND DU LAC	039	*OZAUKEE*	*089*
FOREST	041	*PEPIN*	*091*
GRANT	043	*PIERCE*	*093*
GREEN	045	*POLK*	*095*
GREEN LAKE	047	*PORTAGE*	*097*
IOWA	049	*PRICE*	*099*
IRON	051	*RACINE*	*101*

RICHLAND	*103*	*VERNON*	*123*
ROCK	*105*	*VILAS*	*125*
RUSK	*107*	*WALWORTH*	*127*
ST. CROIX	*109*	*WASHBURN*	*129*
SAUK	*111*	*WASHINGTON*	*131*
SAWYER	*113*	*WAUKESHA*	*133*
SHAWANO	*115*	*WAUPACA*	*135*
SHEBOYGAN	*117*	*WAUSHARA*	*137*
TAYLOR	*119*	*WINNEBAGO*	*139*
TREMPEALEAU	*121*	*WOOD*	*141*

FullMailAdd (Full Mailing Address) *[FGDC: USPS Address]*
The full mailing address of a point feature comprised of AddNumPrefix*, AddNum, AddNumSuffix*, PrefixDir*, StreetName, StreetType*, SuffixDir*, Building*, UnitType*, UnitID*, PlaceName*, State* and ZipCode* concatenated together. Natively provided full mailing addresses are included in this field whenever available.
*where applicable

FullAdd (Full Address) *[FGDC: Delivery Address]*
The full address of a point feature comprised of AddNumPrefix*, AddNum, AddNumSuffix*, PrefixDir*, StreetName, StreetType*, SuffixDir*, Building*, UnitType* and UnitID* concatenated together. Natively provided full addresses are included in this field whenever available. If a natively provided full address is not available, then a full address is constructed from the above address components.
*where applicable

AddNumPrefix (Address Number Prefix) *[FGDC]*
Is a rarely used prefix of the address number. (See FGDC Address Standard 2.2.1.1 for more compliance notes) In Wisconsin, this field is of particular interest due to grid address examples such as “W180N8085 TOWN HALL ROAD”
Examples:

- *N*
- *S*
- *W180N*
- *S379W*

AddNum (Address Number) *[FGDC]*
The whole number component of a posted building identifier

AddNumSuffix (Address Number Suffix) *[FGDC]*
Is a rarely used extension of the address number for a posted building identifier, not to be confused with unit divisions within a building (UnitID). *For example “798 A 26TH STREET”
Examples:

- -856
- *-2445A*
- *B*
- *C*
- *½*
- *.5*

Prefix (Prefix) *[FGDC: Street Name Predirectional]*
One letter street direction that precedes the street name
Examples:

- *N – North*
- *S – South*
- *E – East*
- *W – West*
- *NW – North West*
- *SW – South West*
- *NE – North East*

- *SE – South East*
- *SB – South Bound*
- *NB – North Bound*
- *EB – East Bound*
- *WB – West Bound*
- *CTH – County Highway*
- *STH – State Highway*
- *USH – United States Highway*
- *INTERSTATE – Interstate Highway*
- *W CTH – West County Highway*
- *N STH – North State Highway*

StreetName (Street Name) *[FGDC]*
The legal street name as assigned by local address authority. StreetName does not include the StreetType of a named street. Additionally, StreetName does not include the suffix direction of a coordinate street. The suffix direction of a coordinate street should be stored in the Suffix

StreetAlias1 (Street Name Alias 1) *[LinkWISC]*
The full street name as assigned by local address authority, as commonly used name in community, as historically used name or as name of possible intersecting or related corner-lot street. StreetAlias1 does include the StreetType of a named street. Additionally, StreetAlias1 does include the suffix direction of a coordinate street. This is alias field supports the full street name.
Examples:

- E. North St.
- Johnson St.
- West Washington PKWY
- HWY 151
- County Highway 31

StreetAlias2 (Street Name Alias 2) *[LinkWISC]*
The full street name as assigned by local address authority, as commonly used name in community, as historically used name or as name of possible intersecting or related corner-lot street. StreetAlias2 does include the StreetType of a named street. Additionally, StreetAlias2 does include the suffix direction of a coordinate street. This is alias field supports the full street name.
Examples:

- E. North St.
- Johnson St.
- West Washington PKWY
- HWY 151
- County Highway 31

StreetAlias3 (Street Name Alias 3) *[LinkWISC]*
The full street name as assigned by local address authority, as commonly used name in community, as historically used name or as name of possible intersecting or related corner-lot street. StreetAlias3 does include the StreetType of a named street. Additionally, StreetAlias3 does include the suffix direction of a coordinate street. This is alias field supports the full street name.
Examples:

- E. North St.
- Johnson St.
- West Washington PKWY
- HWY 151
- County Highway 31

StreetType (Street Type) *[FGDC: Street Name Posttype]*
Street type of a named street written to full name of type:

Example Domains:

ACCESS	CREST	GREEN	PARKWAY	SPRING
ACRES	CROSS	GROVE	PASS	SPRINGS
ALLEY	CROSSING	HARBOR	PASSAGE	SPUR
AVENUE	CURVE	HEIGHTS	PATH	SQUARE
BAY	DALE	HIGHWAY	RIDGE	STREET
BEACH	DRIVE	HILL	ROAD	STRIP
BEND	DRIVE N	HILLS	PATHWAY	SUMMIT
BLUFF	DRIVE W	HOLLOW	PIKE	TERRACE
BOULEVARD	DUGWAY	ISLAND	PLACE	TOWER
BOULVARD	EASEMENT	ISLE	PLAZA	TRACE
BRANCH	END	JUNCTION	POINT	TRAIL
BYPASS	ESTATE	KNOLL	PRAIRIE	TRAILS
CAUSEWAY	ESTATES	KNOLLS	PRIVATE DRIVE	TRAILWAY
CENTER	EXPRESSWAY	LAKE	R3	TURN
CHASE	HAVEN	LANDING	R4	TURNPIKE
CIRCLE	HEIGHT	LANE	RAPIDS	VALE
CLIFF	GATEWAY	LOOP	RESERVE	VALLEY
CLOSE	GLEN	MALL	RETREAT	VIEW
COMMON	GLENN	MANOR	ROUND	VISTA
COMMONS	EXTENSION	MEADOW	ROW	WALK
COURSE	FIELDS	MEADOWS	RUN	WAY
COURT	FOREST	MEWS	SCHOOL	WELLS
COVE	FORK	NEST	SETTLEMENT	
CREEK	GARDENS	OVERLOOK	SHORE	
CRESCENT	GATE	PARK	SHORES	

Suffix (Suffix) *[FGDC: Street Name Postdirectional]*
One letter street direction that follows the street name
Coded Value Domains:

- *N – North*
- *S – South*
- *E – East*
- *W – West*
- *NW – North West*
- *SW – South West*
- *NE – North East*
- *SE – South East*
- *40W*
- *2N*

LandmarkName (Landmark Name) *[FGDC]*
The common place name of a point feature. (Provided as available).

UnitType (Unit Type) *[FGDC: Subaddress Type]*
Indicates the unit type associated with a point feature (i.e. apartment, room, suite, unit, etc.)

UnitID (Unit ID) *[FGDC: Subaddress Identifier]*
UnitID includes the number or letter identification string for a building, apartment, room, suite, unit, room or desk (as well as other examples). Not to be confused with AddNumSuffix, as this is a component to the address number. UnitID delineates a unit within an address (i.e. “123 ½ Apt A” → “½” is the AddNumSuffix, “Apt” is the UnitType and “A” is the UnitID).

PlaceName (Place Name) *[FGDC: Complete Place Name]*

The name of an officially designated jurisdiction that the address point belongs to. The name is explicitly defined in the native dataset by the county or jurisdiction itself. PlaceName is provided where it is available in native datasets and standardized to include LSAD descriptors (CITY, TOWN, VILLAGE) when possible.

PlaceNameAlternate (Place Name Alternate) *[LinkWISC]*
The name of the USPS preferred place name* or alternate jurisdiction to which a point feature is located. PlaceNameAlternate is provided where it is available in native datasets and standardized to include LSAD descriptors (CITY, TOWN, VILLAGE) when possible.

*The USPS preferred place name refers to name of an area, sector, or development (such as a neighborhood or subdivision in a city, or a rural settlement in unincorporated area); incorporated municipality or other general purpose local governmental unit; county or county-equivalent; or region within which the address is physically located; or the name given by the U.S. Postal Service to the post office from which mail is delivered to the address.

CensusPlaceName (Census Place Name) *[LinkWISC]*
The name of the geographic area defined by legal boundaries gathered through the 2013 U.S. Census Bureau's Boundary and Annexation Survey (BAS). Point features are attributed with correlating place names through a spatial join to the BAS geometries.

ZipCode (Zip Code) *[FGDC: ZIP Code]*
The 5 digit zip code associated with a point feature

Zip4 (Zip Code) *[FGDC: ZIP Plus 4]*
The 4 additional digits appended to the 5 digit zip code of some point features

State (State) *[FGDC: State Name]*
Two letter state abbreviation of a point feature's address

PtType (Point Type) *[LinkWISC]*
Indicates the zoning type of a point feature
Examples:

- *RESIDENTIAL*
- *COMMERCIAL*
- *INDUSTRIAL*
- *ABANDONED*
- *RESIDENCE*
- *OTHER*
- *UNKNOWN*

AddSource (Address Source) *[FGDC: Address Parcel Identifier Source]*
Indicates the source or entity where a point feature originates (i.e. DANE COUNTY, EAU CLAIRE COUNTY, CITY OF EAU CLAIRE)

LoadDate (Load Date) *[LinkWISC]*
The mm/dd/yyyy when a point feature is loaded and conflated with the statewide dataset.

EntryDate (Entry Date) *[LinkWISC]*
The date of when a point feature was originally created in the native dataset.

CAI (Community Anchor Institutions) *[LinkWISC]*
Cites whether parcel/point is a CAI and summarizes the role of the CAI.

- Library
- Schools K-12
- Libraries
- Health Care
- Public Safety
- Safety and Professional Services
- Higher Education
- Government
- Non-Government
- Department of Corrections
- Ports
- Private Schools

Township (Township) *[LinkWISC]*

Distinguishes the township of the address point.

Section (Section) *[LinkWISC]*
Distinguishes the section of the address point.

Range (Range) *[LinkWISC]*
Distinguishes the range of the address point.

GEOID (GeoID) *[LinkWISC]*
Distinguishes the GeoID of the Census block that the address point resides within. *GeoID is a 15 digit code comprised of two FIPS id components, a tract ID and a block ID (STATE + COUNTY + TRACT + BLOCK).*

- Digits 1-2: State FIPS code
- Digits 3-5: County FIPS code
- Digits 6-11: Tract ID
- Digits 12-15: BLOCK ID – "Census Block Numbers" – Census blocks are numbered uniquely with a four-digit census block number from 0000 to 9999 within census tract, which nest within state and county. The first digit of the census block number identifies the block group. Block numbers beginning with a zero (in Block Group 0) are only associated with water-only areas.

-For More info: http://www.census.gov/prod/cen2010/doc/dpsf.pdf
-Blocks can be found here: http://www.census.gov/cgi-bin/geo/shapefiles2013/main
-About census blocks: http://blogs.census.gov/2011/07/20/what-are-census-blocks/

SourceType (Source Type) *[FGDC: Address Feature Type]*
Identifies the means in which the point was created or the status of the address

- ADDRESS POINT – Was entered as an address point provided by the local level data contributor.
- PARCEL CENTROID – Was created as a centroid from a parcel layer provided by the local level data contributor.
- FOOTPRINT CENTROID – Was created as a centroid from a building footprint layer provided by the local level data contributor.

AddLifeStatus (Address Lifecycle Status) *[FGDC]*
Identifies the status of an address as Active or Retired

- ACTIVE
- RETIRED
- UNKNOWN

AddressCompleteness *(Address Completeness) [LinkWISC]*
Is an assessment measurement that determines the completeness of the point feature's address attributes. This is a general measure that weight four address components AddNum, StreetName and StreetType (or Prefix in the case of highways).

- 0 = no address (poor quality)
- 1 = partial address (at least 1 address element)
- 2 = partial address (at least 2 address element)
- 3 = full address (high quality)

AddressLongitude (Address Longitude) *[FGDC]*
The longitude of the address location, in decimal degrees.
Example: -84.29049105

AddressLatitude (Address Latitude) *[FGDC]*
The latitude of the address location, in decimal degrees.
Example: 33.77603207

Appendix

B. Parcel Schema Definitions

Legend	
ElementName	Denotes database field name
(Element Name)	Full database field name
[FGDC]	Denotes database field name that is equivalent to FGDC element name
[FGDC: <element name from FGDC standard>]	Denotes field name translated to FGDC element name
[LinkWISC]	Denotes database field name distinct from FGDC that serves LinkWISCONSIN project needs

RelateID (Relate ID) *[LinkWISC]*
This string field contains a 10 digit key which correlates with the "RelateID" that exists in the project address point layer. This field will facilitate joins between the project Parcel Layer and the project Address Point Layer. Multiple points residing inside one parcel are given the same identification value.

ParcelID (Parcel ID) *[FGDC: Address Parcel Identifier]*
Unique number assigned to a parcel by the local authority (can be tax roll ID number)

CountyID (County FIPS Code) *[FGDC: County Name]*
The FIPS code for each of Wisconsin's 72 counties. Codes are as follows:

ADAMS	*001*	*KEWAUNEE*	*061*
ASHLAND	*003*	*LA CROSSE*	*063*
BARRON	*005*	*LAFAYETTE*	*065*
BAYFIELD	*007*	*LANGLADE*	*067*
BROWN	*009*	*LINCOLN*	*069*
BUFFALO	*011*	*MANITOWOC*	*071*
BURNETT	*013*	*MARATHON*	*073*
CALUMET	*015*	*MARINETTE*	*075*
CHIPPEWA	*017*	*MARQUETTE*	*077*
CLARK	*019*	*MENOMINEE*	*078*
COLUMBIA	*021*	*MILWAUKEE*	*079*
CRAWFORD	*023*	*MONROE*	*081*
DANE	*025*	*OCONTO*	*083*
DODGE	*027*	*ONEIDA*	*085*
DOOR	*029*	*OUTAGAMIE*	*087*
DOUGLAS	*031*	*OZAUKEE*	*089*
DUNN	*033*	*PEPIN*	*091*
EAU CLAIRE	*035*	*PIERCE*	*093*
FLORENCE	*037*	*POLK*	*095*
FOND DU LAC	*039*	*PORTAGE*	*097*
FOREST	*041*	*PRICE*	*099*
GRANT	*043*	*RACINE*	*101*
GREEN	*045*	*RICHLAND*	*103*
GREEN LAKE	*047*	*ROCK*	*105*
IOWA	*049*	*RUSK*	*107*
IRON	*051*	*ST. CROIX*	*109*
JACKSON	*053*	*SAUK*	*111*
JEFFERSON	*055*	*SAWYER*	*113*
JUNEAU	*057*	*SHAWANO*	*115*
KENOSHA	*059*	*SHEBOYGAN*	*117*

TAYLOR	*119*	*WASHINGTON*	*131*
TREMPEALEAU	*121*	*WAUKESHA*	*133*
VERNON	*123*	*WAUPACA*	*135*
VILAS	*125*	*WAUSHARA*	*137*
WALWORTH	*127*	*WINNEBAGO*	*139*
WASHBURN	*129*	*WOOD*	*141*

OwnerName (Owner Name) *[LinkWISC]*
The owner name of a parcel

FullMailAdd (Full Mailing Address) *[FGDC: USPS Address]*
The full address of a point feature comprised of AddNumPrefix*, AddNum, AddNumSuffix*, PrefixDir*, StreetName, StreetType*, SuffixDir*, Building*, UnitType*, UnitID*, PlaceName*, State* and ZipCode* concatenated together. This attribute is complete as provided from native datasets.
*where applicable

FullAdd (Full Address) *[FGDC: Delivery Address]*
The full address of a point feature comprised of AddNumPrefix*, AddNum, AddNumSuffix*, PrefixDir*, StreetName, StreetType*, SuffixDir*, Building*, UnitType* and UnitID* concatenated together. Natively provided full addresses are included in this field whenever available. If a natively provided full address is not available, then a full address is constructed from the above address components.
*where applicable

AddNumPrefix (Address Number Prefix) *[FGDC]*
Is a rarely used prefix of the address number. (See FGDC Address Standard 2.2.1.1 for more compliance notes) In Wisconsin, this field is of particular interest due to grid address examples such as “W180N8085 TOWN HALL ROAD”
Examples:

- N
- S
- W180N
- S379W

AddNum (Address Number) *[FGDC]*
The whole number component of a posted building identifier

AddNumSuffix (Address Number Suffix) *[FGDC]*
Is a rarely used extension of the address number for a posted building identifier, not to be confused with unit divisions within a building (UnitID). *For example “798 A 26TH STREET”
Examples:

- -856
- -2445A
- B
- C
- ½
- .5

PrefixDir (Prefix) *[FGDC: Street Name Predirectional]*
One letter street direction that precedes the street name
Examples:

- N – North
- S – South
- E – East
- W – West
- NW – North West
- SW – South West
- NE – North East

- SE – South East
- SB – South Bound
- NB – North Bound
- EB – East Bound
- WB – West Bound
- CTH – County Highway
- STH – State Highway
- USH – United States Highway
- INTERSTATE – Interstate Highway
- W CTH – West County Highway
- N STH – North State Highway

StreetName (Street Name) *[FGDC]*
The legal street name as assigned by local address authority. StreetName does not include the StreetType of a named street. Additionally, StreetName does not include the suffix direction of a coordinate street. The suffix direction of a coordinate street should be stored in the Suffix. Additionally, County, State, U.S. and Interstate highway identifiers are not included in the StreetName but identified in Prefix.

StreetType (Street Type) *[FGDC: Street Name Posttype]*
Street type of a named street written to full name of type:

Example Domains:

ACCESS	CREST	GREEN	PARKWAY	SPRING
ACRES	CROSS	GROVE	PASS	SPRINGS
ALLEY	CROSSING	HARBOR	PASSAGE	SPUR
AVENUE	CURVE	HEIGHTS	PATH	SQUARE
BAY	DALE	HIGHWAY	RIDGE	STREET
BEACH	DRIVE	HILL	ROAD	STRIP
BEND	DRIVE N	HILLS	PATHWAY	SUMMIT
BLUFF	DRIVE W	HOLLOW	PIKE	TERRACE
BOULEVARD	DUGWAY	ISLAND	PLACE	TOWER
BOULVARD	EASEMENT	ISLE	PLAZA	TRACE
BRANCH	END	JUNCTION	POINT	TRAIL
BYPASS	ESTATE	KNOLL	PRAIRIE	TRAILS
CAUSEWAY	ESTATES	KNOLLS	PRIVATE DRIVE	TRAILWAY
CENTER	EXPRESSWAY	LAKE	R3	TURN
CHASE	HAVEN	LANDING	R4	TURNPIKE
CIRCLE	HEIGHT	LANE	RAPIDS	VALE
CLIFF	GATEWAY	LOOP	RESERVE	VALLEY
CLOSE	GLEN	MALL	RETREAT	VIEW
COMMON	GLENN	MANOR	ROUND	VISTA
COMMONS	EXTENSION	MEADOW	ROW	WALK
COURSE	FIELDS	MEADOWS	RUN	WAY
COURT	FOREST	MEWS	SCHOOL	WELLS
COVE	FORK	NEST	SETTLEMENT	
CREEK	GARDENS	OVERLOOK	SHORE	
CRESCENT	GATE	PARK	SHORES	

SuffixDir (Suffix) *[FGDC: Street Name Postdirectional]*
One letter street direction that follows the street name
Coded Value Domains:
- N – North
- S – South
- E – East

- W – West
- NW – North West
- SW – South West
- NE – North East
- SE – South East
- 40W
- 2N

LandmarkName (Landmark Name) *[FGDC]*
The common place name of a parcel feature. (Provided as available).

UnitType (Unit Type) *[FGDC: Subaddress Type]*
Indicates the unit type associated with a point feature (i.e. apartment, room, suite, unit, etc.)

UnitID (Unit ID) *[FGDC: Subaddress Identifier]*
UnitID includes the number or letter identification string for a building, apartment, room, suite, unit, room or desk (as well as other examples). Not to be confused with AddNumSuffix, as this is a component to the address number. UnitID delineates a unit within an address (i.e. “123 ½ Apt A” → “½” is the AddNumSuffix, “Apt” is the UnitType and “A” is the UnitID).

PlaceName (Place Name) *[FGDC]*
The name of an officially designated jurisdiction that the parcel belongs to. The name is explicitly defined in the native dataset by the county or jurisdiction itself. PlaceName is provided where it is available in native datasets and standardized to include LSAD descriptors (CITY, TOWN, VILLAGE) when possible.

CensusPlaceName (Census Place Name) *[LinkWISC]*
The name of the geographic area defined by legal boundaries gathered through the 2013 U.S. Census Bureau’s Boundary and Annexation Survey (BAS). Point features are attributed with correlating place names through a spatial join to the BAS geometries.

ZipCode (Zip Code) *[FGDC: ZIP Code]*
The 5 or 9 digit zip code associated with a parcel feature

State (State) *[FGDC]*
Two letter state abbreviation of a parcel feature’s address

LoadDate (Load Date) *[LinkWISC]*
The mm/dd/yyyy when a parcel feature is loaded and conflated with the statewide dataset.

ParcelSource (Parcel Source) *[FGDC: Address Parcel Identifier Source]*
Indicates the source or entity where a parcel feature originates (i.e. DANE COUNTY, EAU CLAIRE COUNTY, CITY OF EAU CLAIRE)

Appendix

C. County Assessment/Observation Report – Sample

Fields Included in County Data Reports	
Field	**Criteria**
METADATA	Present or not for address points Present or not for parcels Data dictionary present for address points Unresolved coded domains present for address points General description of address points provided
CURRENTNESS	Date of data Mode used to determine date of data
GEOMETRIC COMPLETENESS	Address point completeness Parcel completeness
ATTRIBUTE COMPLETENESS	Data with question marks, N/As, etc. Half-filled columns, missing values Address completeness score PlaceName completeness
POSITIONAL ASSESSMENT	Parcel to address point relatability Boundary over/underlap
ATTRIBUTE CONSISTENCY	Consistency of attribute elements/controlled domains
OTHER OBSERVATIONS	Attribute assessment Extraneous data in attribute elements Consistency Extraneous geographic data Retiring versus deleting addresses Presence of zoning info

This project was funded by the Public Service Commission of Wisconsin with an American Recovery and Reinvestment Act grant awarded by the National Telecommunications and Information Administration.