1. Introduction

Hi! It's been a while since you heard from us. I am Ara Erickson with the University of Washington. We are taking this opportunity to do a quick introduction and explain a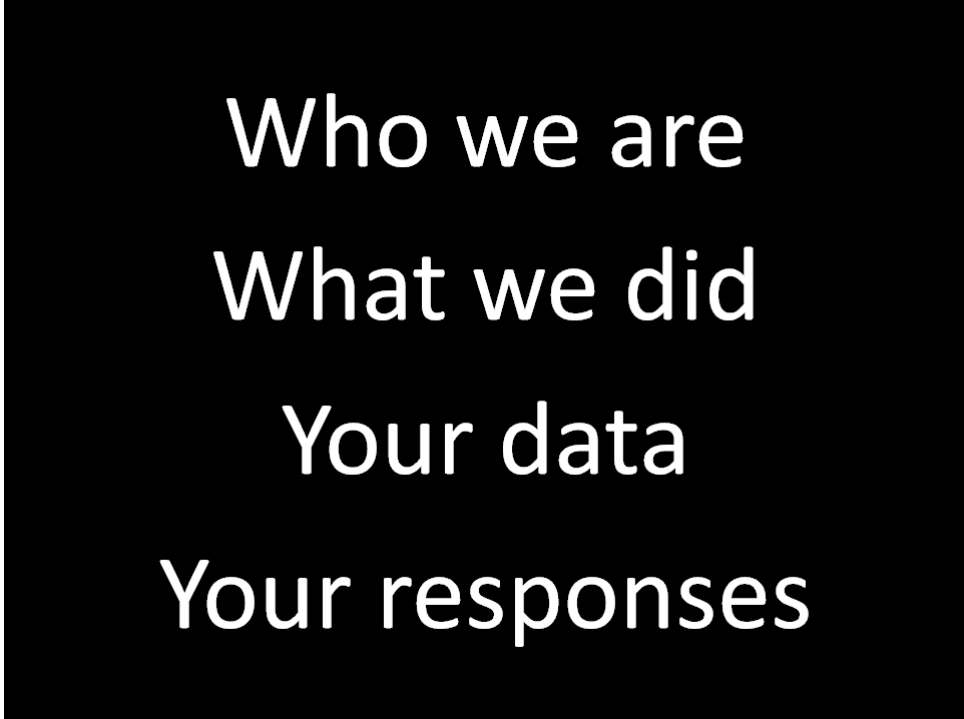 few things about the next steps of the Washington State Parcel Database project. We've opted for a video presentation today, rather than a long e-mail, that we hope you'll find informative and maybe even entertaining. Realizing that not everyone can watch a video while you are at work, however, the full text is also available.

To begin, we would like to thank all of you (and all of your colleagues) for being so helpful and patient since we started on this endeavor last year. We recognize that many of you spent hours, and even days, responding to our initial data requests and actively participating in this project. Thank you.

2. Overview

Today, we'd like to do 4 things:
- o Brief introductions;
- o An explanation of what we've done with the data you graciously provided us;
- o How we'd like you to look at your respective data sets; and
- o What we hope to have as responses from all of you.

We'll start with who we are and why we're collecting and using your data.

3. Who Are We?

We are researchers and staff technicians at the University of Washington. We work for a small research and outreach group called the Rural Technology Initiative, housed at the College of Forest Resources. Our current research projects focus on identifying family forest lands in Washington State and how to best provide incentives to maintain these, and other, forest lands in the state.

4. Parcel Database to Forest Land Database to Family Forest Database

To identify family forest lands in Washington, we first have to identify **all** of the forest lands, including industrial and other large landowners. We then extract the family forest information from this larger dataset. Thus, although we focus on family forest land, we are building a Forest Land Dataset for all of Washington State.

But, before we can build a forest land dataset, we have to build a parcel database since we rely on the spatial and tabular data that you produce and maintain to build the processes to identify the forest lands.

As you can imagine, it takes quite a bit of effort, both time and financial, to collect parcel data from all 39 counties, plus the various other state and federal agencies who manage parcel-related data.

5. Parcel Working Group

When we started this project we realized that in order to ensure that we have a reliable dataset to use for our research--now and in the future--we needed to have an updatable statewide parcel database built first. Therefore, we directed our efforts toward the **Washington State Parcel Database project** as a means to eventually arrive at the Forest Land Dataset.

In order to do this, we knew we needed help -- and thought that we might be able to share in some of the efforts that others were expending as well.

We gathered together a group of public agencies which regularly use and request parcel data from the many data producers in Washington State and now have an active and growing Parcels Working Group.

We meet monthly to discuss and brainstorm how we can work together to collect and share parcel data. My colleague Luke Rogers is currently co-chairing the group with a representative from the Washington State Department of Health.

6. Review of Who and Why

We just threw out a lot of projects, databases, and names. Here's a quick review of who we are, who the other people involved in this project are, and what we are all doing:

UW
- Compile statewide parcel database

Parcels Working Group
- Help facilitate needs assessment and sharing of database

UW and Family Forest Foundation
- Build a forest land dataset on top of the statewide parcel database
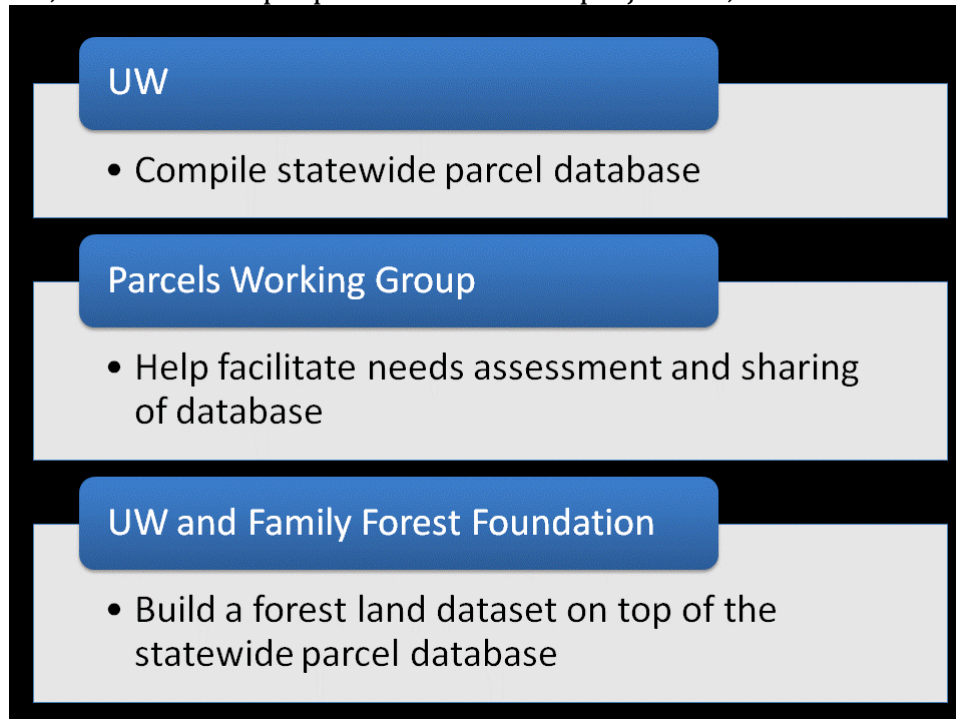
- We, the University of Washington, are compiling a statewide spatially-based parcel database from your data;
- We are working with the Parcels Working Group to facilitate a needs assessment and sharing of this database;
- We are also working with the Family Forest Foundation, a non-profit group, to build and use the forest land dataset and, as an extract of that, the family forest dataset.

So, three groups: the University of Washington, the Parcels Working Group, and the Family Forest Foundation. And, three datasets: a statewide parcel database, a forest land dataset, and a subset of the forest land dataset – the family forest dataset.

7. Collaboration

That's a lot of collaboration! But, nothing like the work we are trying to put into building relationships with the data producers -- you. We realize that none of these projects could happen without your continued production and management of such valuable and detailed information.

When we started this project, we made it our primary objective to find a way that everyone could benefit. One of our goals is to help find more resources for data producers to continue to manage and update GIS-based parcel data and help those counties without a GIS find support to start on the path.

As a first step toward this collaboration, we told all of you that we would be back in touch after we assembled all of the data and before we did anything else with it. Well, here we are!

8. Where Are We in the Project?

We spent last fall making connections with the county assessors and GIS departments across the state. Most of you received an introductory e-mail explaining our project and then we either followed up with phone calls, e-mails, faxes, or anything else that allowed us to start forming a relationship with the parcel data producers. We also visited a few counties to work directly with the assessors and staff.

We acquired data in a variety of formats and methods…and, we are pleased to report that we acquired data from 37 of the 39 counties (that's a 95% participation rate!).

9. Concerns

As we expected, there were concerns about privacy issues, timeliness of the data, sharing data, and unfunded mandates. We hope that we were able to address some of those issues through the many phone calls and other forms of communication we had with you. We recognize, however, that there are still some lingering concerns. We will continue to work with each county and all other data producers to address these concerns.

10. Data Producers Information
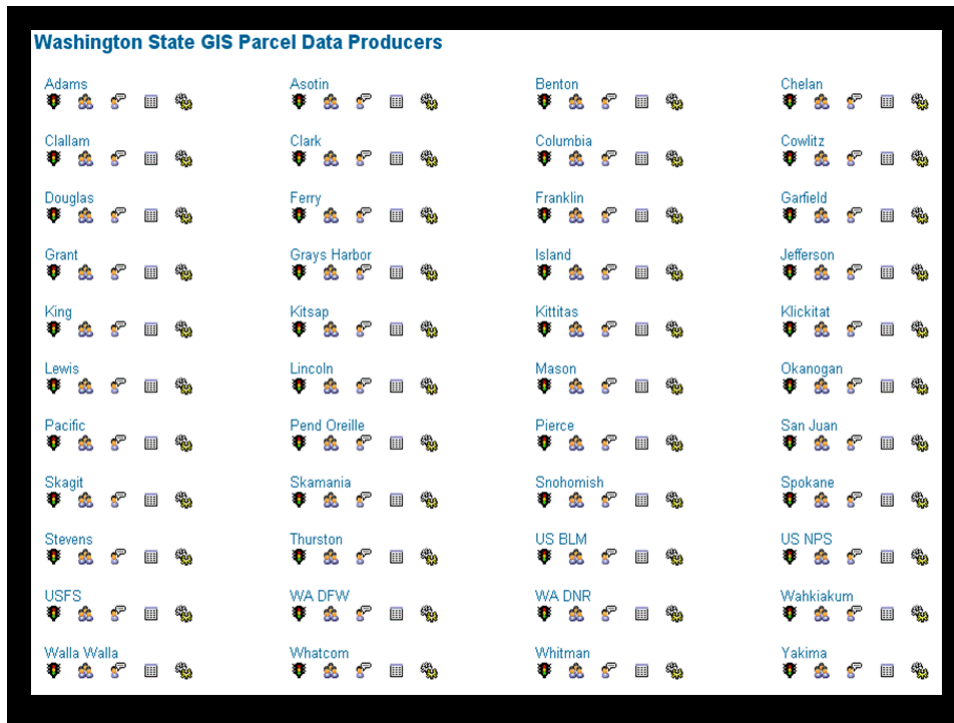
As we began contacting data producers we documented contacts for each county, all of our communications with the contacts, GIS status, and more. We built internal web pages for each data producer that includes all of this information, and a public-facing site that includes status, main contacts, and information about each data set collected.

11. Attributes Documentation

Once we began receiving data, we documented each attribute that we received. If metadata was provided with the data, we used that; if not, we made educated guesses based on on-line parcel information, other documentation we could track down, and phone calls or e-mails back to the counties.

12. Common Attributes

After all of the attributes were documented, we assigned each attribute to a "common" attribute category. These common categories were general themes that the attributes were related to, such as tax information, parcel address information, etc. The common attributes were summarized to give us an idea of how many data producers provided similar attributes.

Attribute	Required	Desired
Landowner Name	110	24
Parcel Identification Number	108	25
Landowner Address	94	35
Site Address	96	31
Township, Range, Section	78	43
Critical Areas Designation	48	53
Current Zoning Information	47	50
Land Use Code	40	55
Taxpayer Name	43	39
Taxable Area	51	26
Current Use Participation	37	40
Taxpayer Address	33	44
Date	34	40
Grantor Name	33	33
Improvement Value	17	43
Land Value	19	37
Market Value	20	33

We were also looking at what attributes appeared to be most requested by participants in the Parcels Working Group. We distributed a survey and found that the most requested parcel-related attributes were landowner name, parcel identification number, landowner address, and site address.

This table lists these and other attributes along with the number of respondents who required versus desired the attribute.

13. Attribute Normalization

SITUS
123 Main St., Ritzville, WA, 98113

HOUSE NO	STREET	CITY	STATE	ZIP
123	Main St	Ritzville	WA	98113

We also needed to make the attributes from each county work with data from other counties. For example, some counties store parcel address information (also known as situs) in one field:

Others store each bit of information in many different fields:

Normalized attributes

Address_1
Address_2
Address_3
Address_City
Address_Country
Address_State
Address_Zip
Name
Name_Code
Name_ID
Acquisition_Date
Aggregated
Duplicate
Integration_Date
Organization_ID
Poly_ID
Flattened

Acres_Improved
Acres_Tabular
Acres_Unimproved
Land_Use_Code
Land_Use_Description
Legal_Description
Market_Value_Crop
Market_Value_Improved_Land
Market_Value_Improvements
Market_Value_Land
Market_Value_Timber_Land
Market_Value_Total
Market_Value_Unimproved_Land
Name_Role
Parcel_ID
Property_Type
Quarter_Section

Range
Section
Situs
Taxable_Value_Crop
Taxable_Value_Improved_Land
Taxable_Value_Improvements
Taxable_Value_Land
TaxRoll_ID
Taxable_Value_Timber_Land
Taxable_Value_Total
Taxable_Value_Unimproved_Land
Tax_Account_ID
Tax_Code_Area
Township
URL
Name_Role

By focusing on our common and desired attribute information, we were able to determine which data storage format was most consistent among the data producers and were eventually able to arrive at a list of normalized attributes that we would work with.

These attributes are shown in this slide.

We know that since we didn't ask for specific attributes last fall, wanting to make this as easy as possible for you and your offices, we are not fully capturing all of the attributes that your office may collect and store. For Round Two of this project, we hope to fill in some of the missing attributes.

After documenting and summarizing the attribute information, we had to figure out how the data was stored and related to other datasets.

14. Figuring out Related Data

Some counties provided us with flat files (one record for each parcel), some provided semi-flat files (multiple records for one parcel, one per owner for example if the parcel has two owners), some provided relational tables, and some provided multiple tables that didn't relate to each other quite as we expected. We spent many months sorting through all of the possibilities and tried to figure out the best way to deal with issues that come up when trying to go from so many different formats to one format.

15. Structure of the Statewide Parcel Database

After sorting through the many possibilities of how data is stored, we began to design how we wanted the normalized, statewide database to function. Although we want to maintain as much of the original data and format as possible, we realized that whatever we did, we were going to have to alter the format of some county's data.

We opted to build a relational database. Many counties already have relational databases, or at least related tables, so we knew we would be sticking with something already in existence.
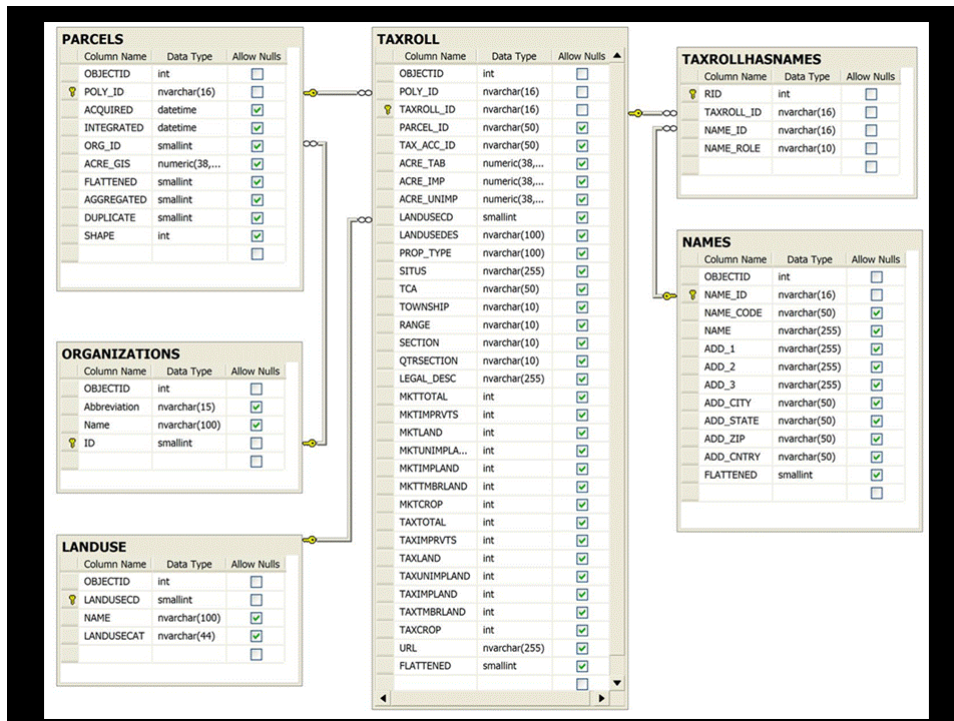
Similar to ESRI's parcel data model, which is quite complex and thorough, we designed a database that would allow for future flexibility. The current database includes three main tables: PARCELS, TAX ROLL, and NAMES. Each table contains fields that are related to the other tables. These relationships are stored as relationship classes.

16. The PARCELS Table

The PARCELS table only includes the information about the individual spatial feature: acres as calculated by the GIS, POLY_ID, and information about when the features was integrated into the database and acquired from the data provider. We also assigned each data provider a unique Organization ID (Org_ID). The POLY_ID is a concatenation of the ORG_ID and a random unique number. Thus, each of the 2.9 million parcel features in Washington has a unique identifier in the statewide parcel database.

17. The TAXROLL Table

Each feature in the PARCELS table is related to zero, one, or more records in the TAXROLL table. The TAXROLL table has a unique identifier as well, TAXROLL_ID, as well as all of the information about the tax record, minus the owner and taxpayer name(s) and addresses.

18. The NAMES Table

The NAMES table includes the name and addresses of the owners and/or taxpayers associated with each record in the TAXROLL table. The NAMES table and the TAXROLL table are linked by a TAXROLLHASNAMES table, which allows multiple names to be associated with a tax record as well as one name to be associated with multiple tax records.

19. The Trail of "Crumbs" – Tabular Information

Additional information included in all of these tables is what we call the "crumbs" of our data processing. We wanted to ensure that anyone using the database understood what processing steps were taken in order to normalize the data to the statewide format. Thus, there are multiple fields that include counts of the number of times we did some sort of spatial or tabular processing. For example, we wanted to have one instance of each name per county in the NAMES table, so we matched for name and address information to make one unique record per name. If one person owns four parcels in one county, and their name and addresses match, their name would only be in the NAMES table once and the NAMES.FLATTENED field would say "4" – meaning that we found 4 instances where this name was included in the original data, but we are only including it once.

## 20. The Trail of "Crumbs" – Spatial Information
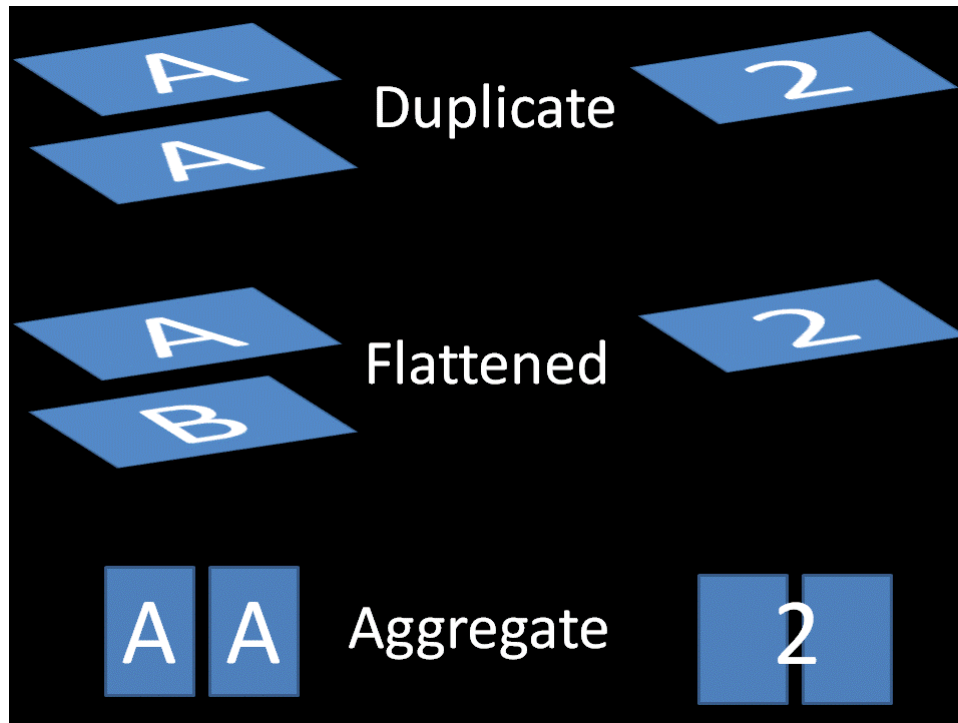
Other times that we include this type of information is when we run the spatial features through processing steps to identify and remove duplicate records (PARCELS.DUPLICATE), identify and flatten multiple instances of the same spatial feature with different tax roll information (PARCELS.FLATTENED), and identifying and merging two features that share the same tax roll information, including the parcel identification number, into a multipart shape (PARCELS.AGGREGATE).

21. Database Documentation

This level of documentation, within the database and included on the website, allows the users full access to understand how we built the database and how to use it properly. It also allows us the opportunity to share with you the detailed methodology and work together to ensure that we are interpreting your data correctly.

22. Tabular Data
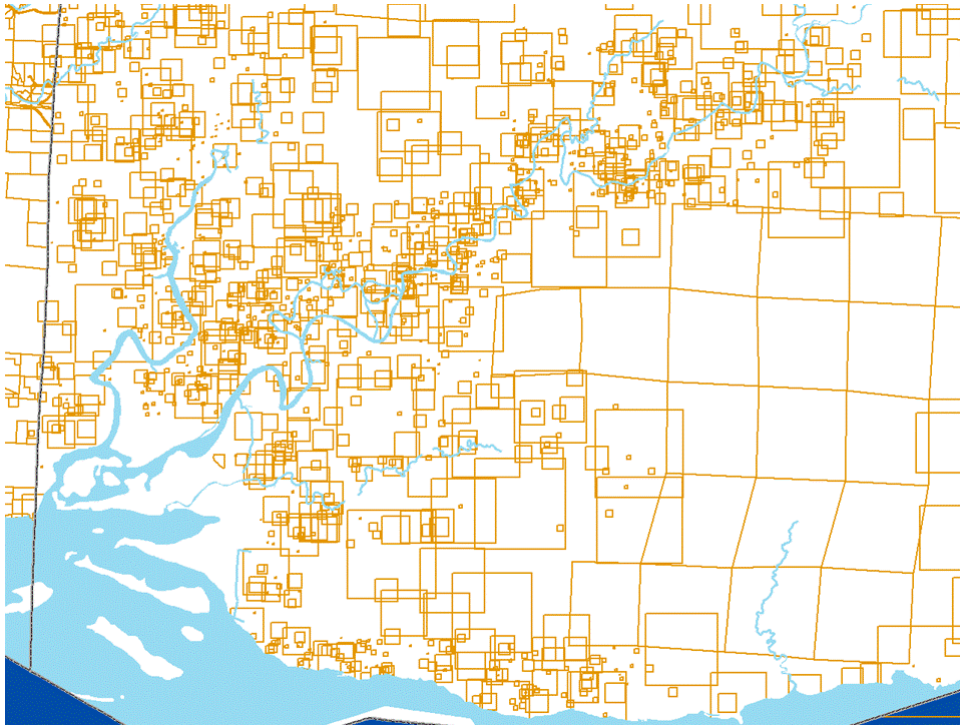
Some counties do not yet have parcel data in a GIS. A major goal of the Parcels Working Group is to help those counties develop GIS parcels. In the interim to include tabular tax-roll data into the statewide parcel database we built quasi-spatial parcel polygons from the tax-roll acreages, legal descriptions and/or parcel numbers. By using the Township, Range and Section information we wrote a program that constructs polygons randomly near the true location of the parcel. While not suitable for analyses at the local level, for statewide statistics, mapping and trends it is a reasonable substitute for spatially explicit GIS data.

23. The Process

After all of this documentation and database design was worked out, we were ready to build the first version of the database. We used a very cool software program called FME (Safe Software's Feature Manipulation Engine). This program takes all types of data and runs it through a series of steps to eventually end up with all of the different data imported into the same format and stored in the Parcel Database.

Our sample workbench looks like this:
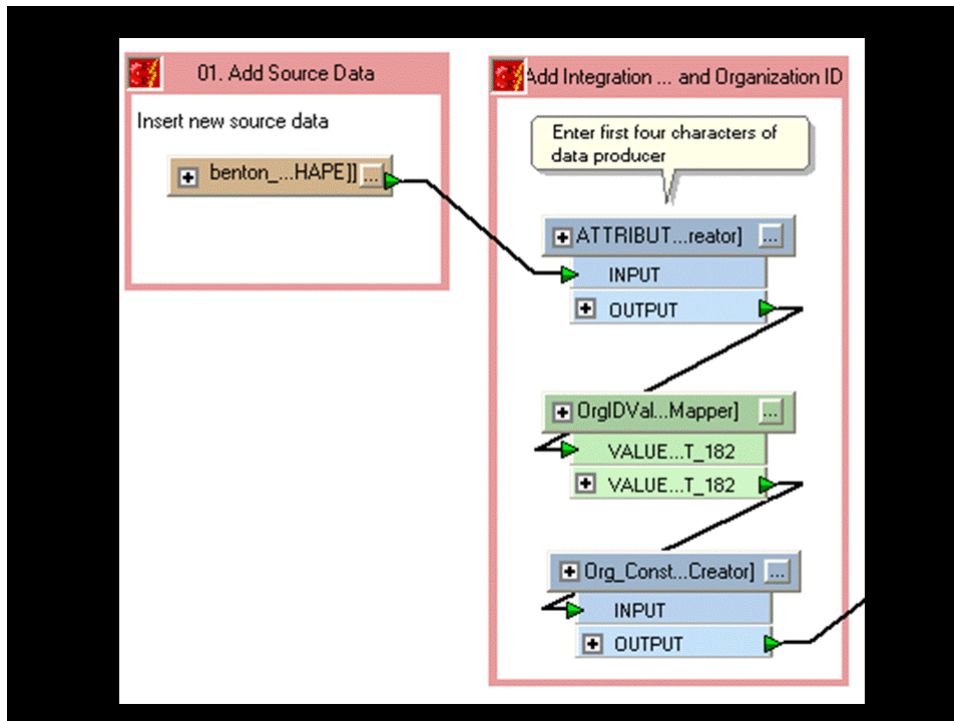
- We import the original data
- We add constants to identify each data producer: Organization ID, date the data was acquired, and date it was integrated.
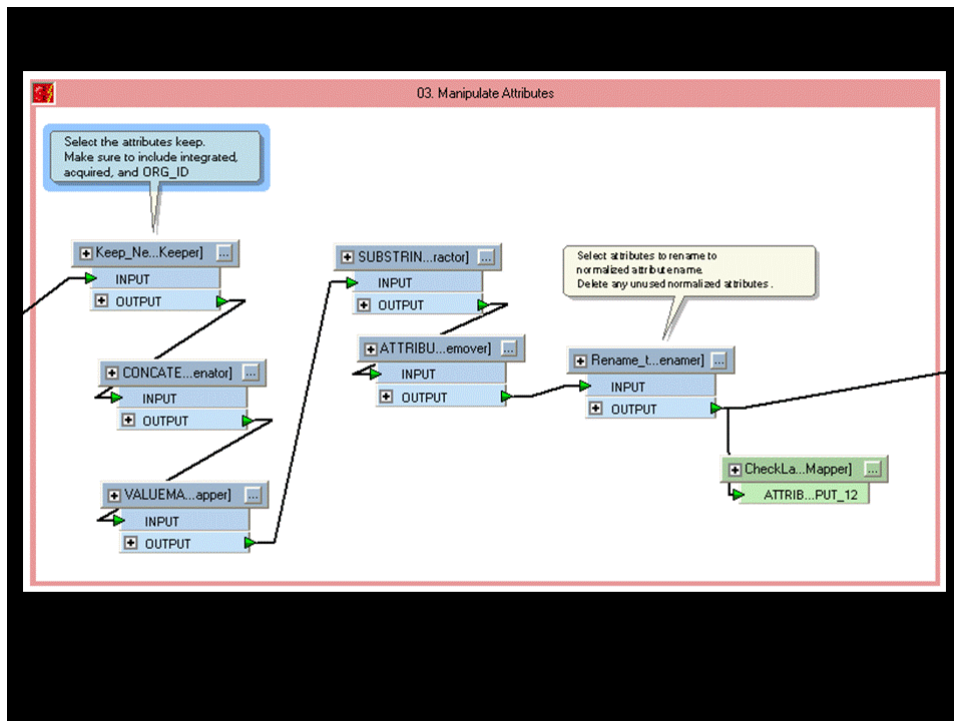

- We do a series of manipulations to normalize the attributes, such as concatenating five different legal description fields into one. We also make sure that the land use

codes provided match the statewide standardized land use codes or use a crosswalk to make them consistent.
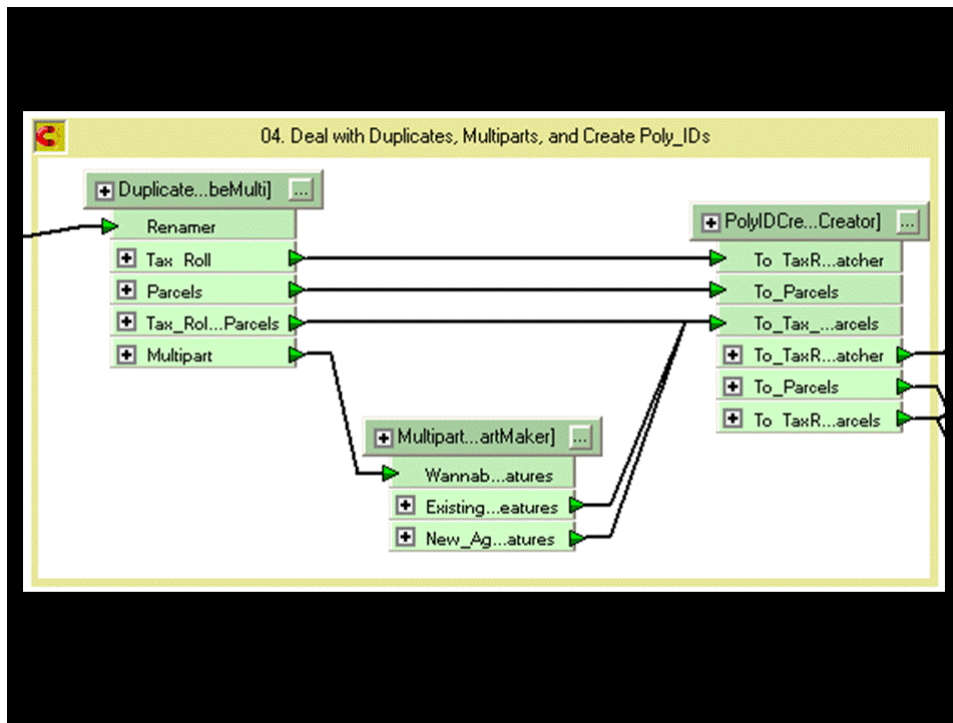
- We deal with duplicate records (the exact same shape and information) and stacks (same shape different information).
- We make single-part records that should be stored as multi-part records (which we call wanna-be multis) by finding records that have different shapes but the same tax roll information.
- We create a unique Poly_ID, of which the first 2-digits are the Org_ID, thus ensuring a unique identifier throughout the entire database.
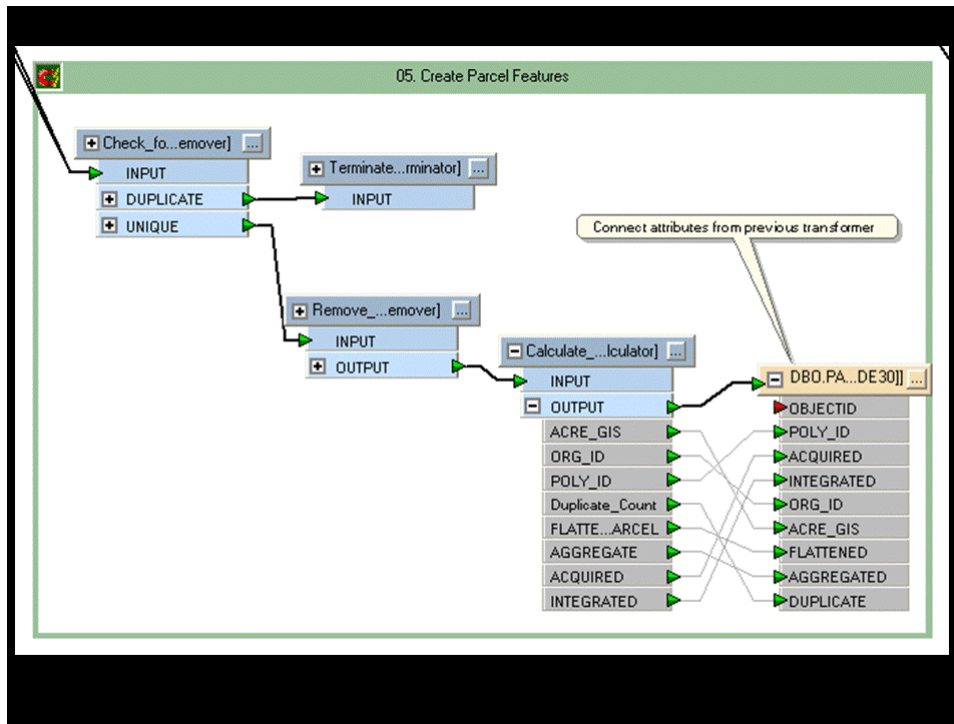
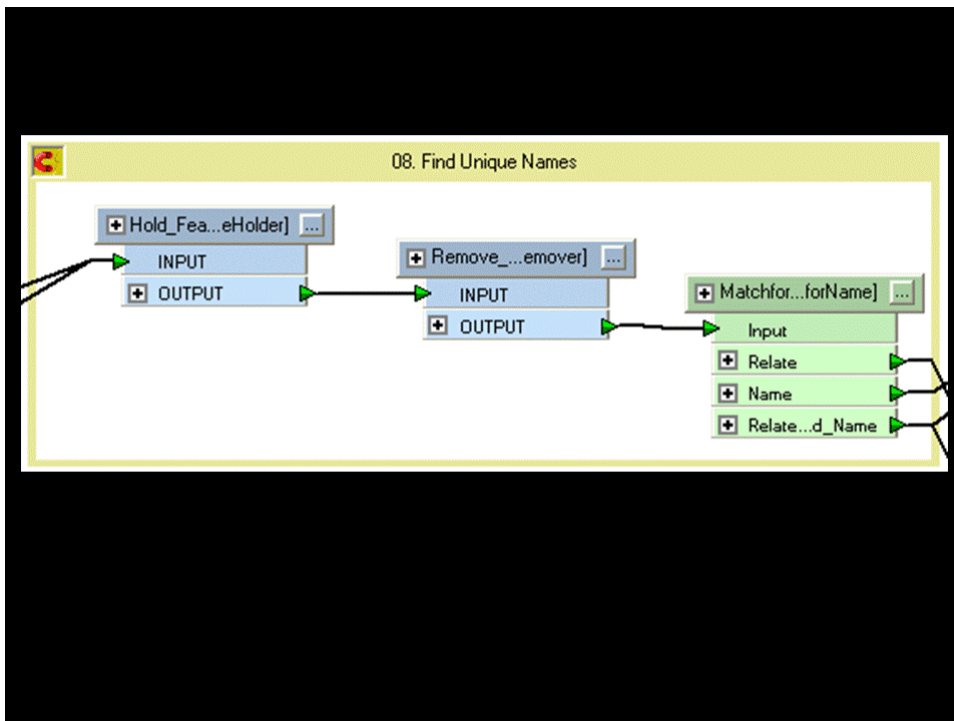- We send the unique parcel features to the PARCELS table.
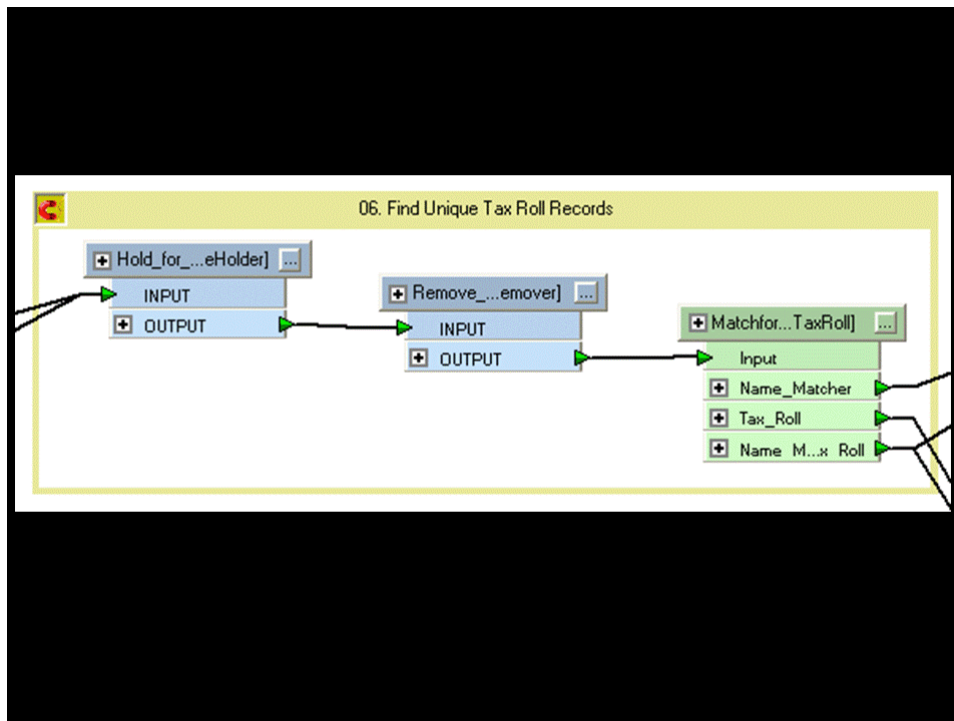
- We take the tax roll information associated with the parcels and find records which have the same tax roll information, such as taxable value and land use code, but possibly different owner names.
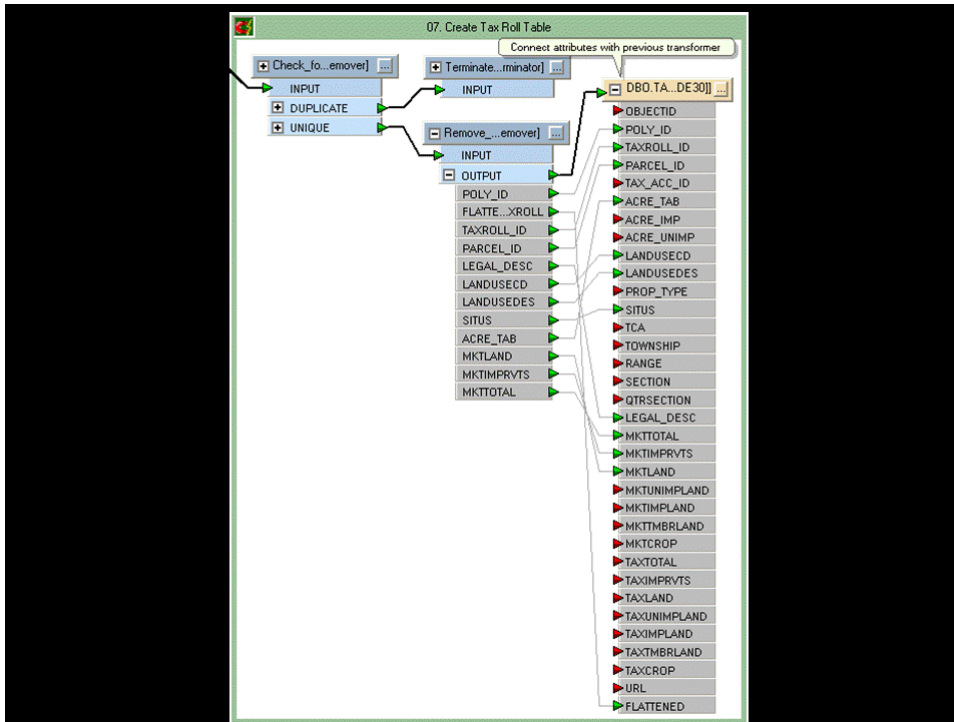
- We send the unique tax records to the TAXROLL table.

- All of the tax roll records are further searched to find matching owner or taxpayer information.
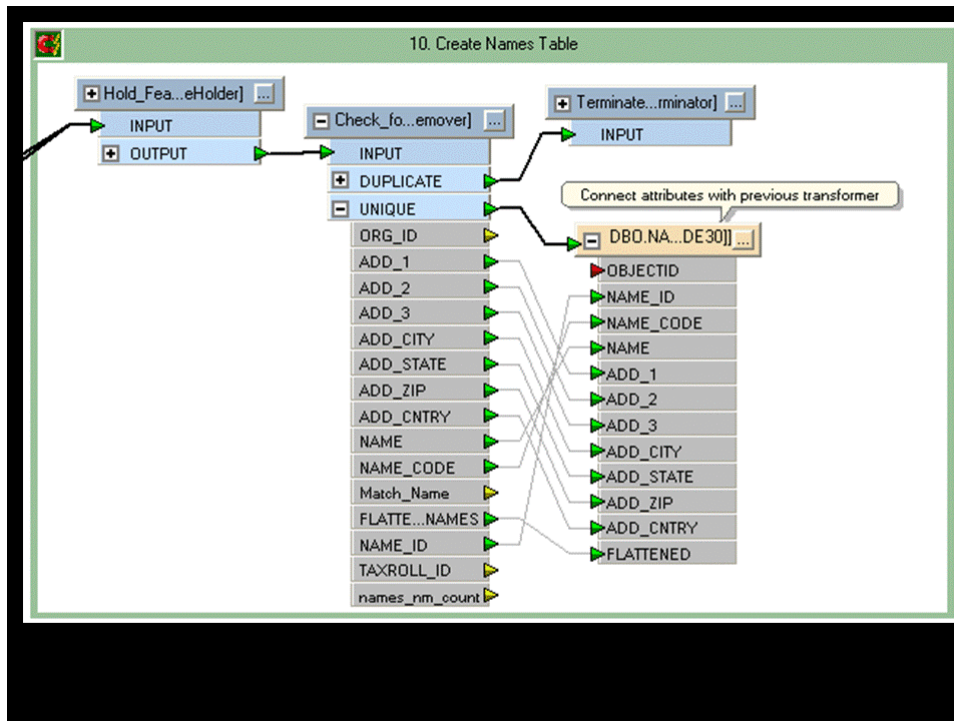
- We then send the unique name records to the NAMES table.

24. Completed Normalized, Standardized Database

In the end, the database that is produced is a fully-relational database, with three tables and two relates. And, we can view, analyze, and use all of the data acquired now in the same format, with normalized attributes! It's pretty amazing, really.

25. What Now?

Although we tried our best to do the correct normalization and processing steps, we know that we have a few mistakes or oversights. We would like to work with you and your staff to identify and fix these errors so that we can be confident that we have produced the best available statewide parcel database.

26. Visit Your County's Documentation Page
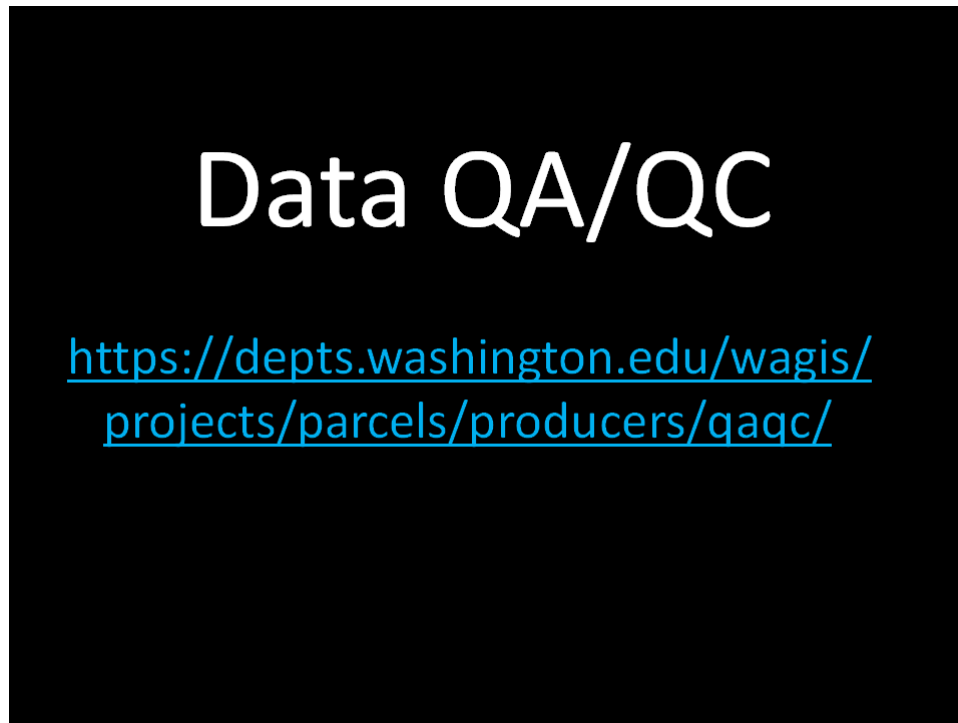
We'll start by asking you to visit your county's documentation and normalization page and look through the various information we have included, including detailed tables of attributes and normalization procedures.

27. The Survey

Then, we'll ask you to take a brief (no more than 10 minute) survey about what you saw on that site and answer a series of questions and help us identify and fix errors or put us in contact with someone who could help us sort through any issues.

This survey also includes a series of questions about how we can share the completed parcel database with members of the Parcel Working Group, other academic researchers, and non-profits interested in forest land information and other land information.

28. In Closing

Thank you for taking the time to learn more about the statewide parcel database. We hope it was informative and maybe even entertaining! A few key points before we sign off:

We would like to start working on fixing any problems and identifying any sharing concerns as soon as possible. Therefore, we are asking that you (or someone you identify as a surrogate) respond to the survey before the end of July. We will follow up with a phone call if we don't have a response from you before July 15th and a second call first week of August.

Thank you for your continued support of this project.

We look forward to future collaboration and successes.

Thank you again for your continued support of this project and we look forward to future collaboration and successes. As a member of the Parcels Working Group, remember that this is your project too!