Generating Legislative Reports from the Wisconsin Idea Database
Last updated 12/14/2021

The Wisconsin Idea Database website is a public-facing site and a tool used to generate reports for the Office of State Relations to distribute to legislators.

This document is for the person in University Communications charged with managing the data collection and production of Wisconsin Idea reports and serving as a liaison between State Relations and University Marketing. Read it through before starting. Ask Jason Gohlke or Bill Graf (among others) if you have questions.

Project Summary

- In this project, you will collect data on UW–Madison's activity around the state from a wide variety of campus units.

- These data points will ultimately be associated with Assembly and Senate districts and counties.

- These activities will be displayed as *Partnerships*, *Placements*, and *Key Facts* on https://wisconsinidea.wisc.edu/, when users browse county profiles.

- That data will also populate District Reports for the Office of State Relations.

- Your team will use the report generator to create 132 reports, one for each state legislator, at the beginning of each legislative session in January.

- During the report generation process, you will be able to select and edit the output to a limited extent, but the report generator does NOT edit the data in the database.

Task overview

1. Meet with State Relations in the summer to discuss potential changes to last year's reports (content, data, layout).

2. Communicate with University Marketing on any proposed changes to reports and/or the website.

3. In early September or sooner, reach out to sources of data around campus (and at UW System) for updated data (for the past fiscal year, generally). Follow up with them and attempt to have data in hand by at least October 31.

4. By the end of November, review all of the data for spelling, capitalization, company names and company addresses. Send it to UMark for geomatching (when applicable).

5. When the data has been uploaded to the website, review content and layout of county profile pages (now populated with the new data), and content and layout of reports (run through State Relations).

6. When ready, run 132 reports for Wisconsin legislators and send to State Relations (by January 1 if possible, but in general before the next year's legislative session starts). If time is a factor, generate them in a specific order (covered later on in this document).

About the project

Data collection

The first step is a massive data collection project. You will get data for the most recently completed academic year (sometimes including the summer just completed) from various locations on campus. Which data do you get from whom? Find out on the spreadsheet at https://docs.google.com/spreadsheets/d/1H-tERYtsaS0uiiAFZPRH_8dnCVk3dQTYNNA7coD4AMM/edit#gid=2082395055

Start reaching out in September. Give people a couple weeks at minimum. Some will take longer. Try to get data in by the end of October at the *latest*, in order to give yourself sufficient time to review the data.

The Box folder that includes data from 2020-21 data and reports generated in the 2021-22 academic year is currently at https://uwmadison.box.com/s/gvr4h1sojjz8qaa1yszo773qwnfr77y7. Replicate this structure for the new year and share with UMark after you review it.

Generating Reports

The reports feature lets you create and customize reports for every Assembly, Senate, and Congressional district in the state of Wisconsin.

To access the Reports generator:

1. If you have *not* previously done this, register with the Wisconsin Idea Database by going to https://wisconsinidea.wisc.edu/.

2. Click on the "Register" link at the top right, fill out and submit the registration form.

3. Upon registering, you will be logged in automatically and taken to your "dashboard" page. During subsequent visits, you will need to log in to the application (the "Login" link is at the top of the home page). This will take you to your dashboard page.

To create a report:

4. Log in to wisconsinidea.wisc.edu using your NetID.

5. Click on "Reports" in the application navigation. (If you don't see "Reports" in the navigation, contact Peter Weil, Joyce Johnston, or Nick Weaver in UMark to get the correct level of access.)

6. On the "Reports" page, first select district type, then select the district number and the year for which you want to generate the report.

7. Click the "submit" button, which will take you to the "Report Preview" page.

8. From the "Report Preview" page, you can customize most of the information that will appear on the report. Some of the data in the report cannot be changed at this stage.

   a. Select the items you would like to appear on the final report by checking the checkbox next to the item. You may adjust the spelling and capitalization of names and cities.
   b. Generally, you will include the top 5 non-UW spending recipients and as many under other categories that fit on the page.
   c. Since recipients of UW–Madison spending sometimes appear more than once in the list (usually under slightly different spellings), you may also add amounts together.

9. Once you've reviewed the report, click "Generate Report" at the bottom of the page. The application will generate a PDF version of the report that you can download. (Use the back button to re-generate a report if it's longer than two pages.)

Lesson Learned: Start data requests early

- **Ask early.** Asking people for data about two weeks after instruction is good timing, since important data is available at that point (enrollment data). You may wish to reach out in August just to ensure they're still the right people to ask.
  - ○ In any case, try to have all data in by the end of October. Some folks will require more follow-up than others.

- **Don't skimp on reviewing data.** Getting data in before November gives you time to review the data sufficiently. Turn over to UMark only after reviewing at *least* for duplicates and ideally for:
    - Inaccurate addresses of companies
        - misspelled city names
        - companies that are inaccurately matched to employee addresses
        - companies with PO Box addresses
    - Outdated company names (hospital mergers, etc.)
    - Duplicates or near-duplicates
        - Note the possibility that duplicate company names with a number of addresses may indicate that those are employee addresses (professional development training, for example)
        - Some companies have
    - Punctuation and spacing
    - [optional] "Inc," "LLC," and other extraneous suffixes (left them in in 2021)
    - [optional] If you're *really* good, make the names consistent across datasets. Maybe start by establishing standard styles first.

- **Why is this important?**
    - UMark developers convert address information to legislative districts and counties, by way of first calculating latitude and longitude and then matching those points to geographic boundaries. (Some of the data you receive will already be aggregated into district & county totals.)
    - When company names and addresses are wrong, activities are assigned to incorrect districts (and even counties), which isn't great when you give reports to legislators or post the data publicly.

- Sufficient review of the data would allow us to be completely consistent about punctuation; do some intensive Googling to ensure addresses correspond to the associated company; and to do some serious de-duping (now that I've seen the output, *all* these things are worth spending significant time on before sending to UMark).

Notes on programs that train individuals on behalf of companies (such as CPED and InterPro):
Because of how registration and payment work for various programs in which UW–Madison trains the public, these programs store the residential addresses of people who participate in training rather than their companies' business addresses.

This results in a need to review almost every line of these files, because employee addresses are seldom in all the same districts as their employers.

While State Relations staff could explain to legislators, "Though Company X, which is listed in your report, is outside your district, people who live in your district who work for them received CPED training"), this is not ideal.

When this was discovered in Fall 2020 while writing reports, I verified addresses of companies and matched them to districts on the Wisconsin Legislature's page before including them in any given report. In Fall 2021, I verified addresses of companies before sending the data to UMark, which means the report generation process will take less time. I am confident that major companies in the report generator will be listed in the correct districts; however, companies that didn't participate in multiple trainings were not *always* verified.

Lessons learned for 2022-23 and beyond:

Report generation: When time is short

- **Generating the reports:** In 2020-21, this order was prioritized: 1. New legislators; 2. Leadership; 3. Members of the Joint Finance Committee. In January of odd-numbered years, these should be the first to be generated because they are needed most for State Relations' purposes.

  - Doing them one at a time in a random order, or in numerical order, is perfectly OK.
  - However, the report generator form generally "remembers" the changes you made to each field in any given session. That means the *really* efficient way to generate the reports, and save time overall (especially for reports that will be distributed at the beginning of odd-numbered years), is to do this:
    - a. After election results are certified, make a list of districts of new legislators (see what you can get from State Relations staff first – don't reinvent the wheel)
    - b. Make a list of districts of JFC members
    - c. Make a list of districts of leadership in each house
    - d. Generate one Senate district report at a time
    - e. After generating each Senate district report, generate the reports for the Assembly districts that are fully contained within that Senate district
    - f. Generate reports for Senate districts that contain Assembly districts from the above lists that haven't yet been generated
    - g. Generate reports for both Assembly districts contained in those Senate districts
    - h. Repeat for all other districts: Senate, Assembly, Assembly; etc.
  - In other words, make a list of Senate districts that are in Jt. Finance & leadership; generate those and their component Assembly districts; make a list of Assembly leadership and Jt. Finance districts that haven't already been generated; generate their corresponding Senate districts first as well as the neighboring Assembly districts within the same Senate district; do the rest, Senate districts first.

- Update notes on the data as needed and send to UMark.

  o Part of it is this:
    - UMark uses Bing Maps API to geocode addresses to latitude and longitude
    - QGIS3 – opensource Geographic Information System – used to map latitude and longitude to Assembly, Senate and Congressional districts and to county

Things to think about in the future

- Consider setting up a timeline document in addition to the "who to contact" document – do as a Microsoft Planner timeline. That could be used to assign tasks and generate email reminders (incomplete example): https://tasks.office.com/uwprod.onmicrosoft.com/Home/PlanViews/LpqAFio3OU2qmpgsNMRSN2QAFBST?Type=PlanLink&Channel=Link&CreatedTime=637629295154110000)

- Does sponsored research show up on reports? Where? Should it? It does at least show up on the website.

- **CONCATENATION OF NAME AND CITY:** Question for UMark: Is there a benefit to concatenating company name and city within the database? Why not treat them like any other data points, keeping them in separate fields (as with the rest of the data) and formatting them upon display on the site and report generation? Not doing this could make scanning and manual editing of data in the report generation field easier.

Current interface:

Major Vendors

INCLUDE IN REPORT?	VENDOR NAME AND CITY	AMOUNT SPENT
☐	Ge Healthcare, Wauwatosa	$ 1,452,267.31
☐	Capital Data Inc, Brookfield	$ 1,134,636.92
☐	Ricoh Usa Inc, Brookfield	$ 802,242.34
☐	United Parcel Service Inc, Elm Grove	$ 745,164.71

Proposed:

Major Vendors

INCLUDE IN REPORT?	VENDOR NAME	CITY	AMOUNT SPENT
☐	GE Healthcare	Wauwatosa	$ 1,452,267.31
☐	Capital Data Inc	Brookfield	$ 1,134,636.92
☐	Ricoh USA Inc	Brookfield	$ 802,242.34
☐	United Parcel Service Inc	Elm Grove	$ 745,164.71