Training and Technical Assistance Webinar Series

Maintaining a Digital Presence - The Illinois SAC Experience

February 25, 2016
Maintaining a Digital Presence - The Illinois SAC Experience

Mark Powers
Research Analyst
What do we do at the SAC?

• Collect, house, and disseminate data over geography and time.
• Publish reports on data trends.
• Conduct evaluations and publish results.
• Provide technical assistance.

The actual services we provide haven’t changed much over time, but the technology and efficiency have.
Presentation

• SAC website redesign, 2010-2012
• SAC website redesign, 2015
• User base considerations
• Data considerations
• Examples of data tools and software
• Recommendations from our experience
ICJIA WEBSITE REDESIGN, 2010-2012
2010

• The good news – we made a top 12 websites list!

• The bad news – it was a top 12 worst websites list

• Illinois had a template that all state government sites had to use while Blagojevich was governor; Next governor gave agencies much more freedom
Technology and Funding in 2010

• Adobe Flash
  – Used primarily for videos and games previously
  – Developers created interactive data tools that would be difficult to produce consistently in standard HTML

• Jquery
  – Greatly simplified JavaScript code and enabled people with less coding experience
  – Allowed developers to write one set of code for all browsers instead of ten different versions

• SJS grants
  – Opportunity to primarily fund dedicated staff and software purchase
Steps taken

• Developed concept paper
• Applied for and received SJS grant
• inventoried data
• Converted data from spreadsheets to normalized database (Access)
• Found software to allow mapping and other analyses (Instant Atlas)
• Built accessible alternative (Instant Atlas provided data tables generator)
• Produced Instant Atlas templates and set up update procedures, put excel files online, and made our publications easier to find
Concept Paper

• This would be developed prior to seeking grant
  – Products & Goals
  – Audience
  – Feasibility
  – Benefits
  – Estimated cost (rough) and how to pay
Product & Goals Examples

2010 - 2015

• Data downloads
• Interactive mapping and time series
• Accessible alternative under state guidelines
• Automated data profiles
• Better way to deliver reports
Feasibility

• Do you have the right staff? Do current staff need training?

• What threats are there to impeding the progress?

• Even if the deliverable is feasible, can it be updated?
Benefits & Costs

• What do the SAC and the users gain from this?
  – Relative to current products
  – Relative to other opportunities
  – You will probably overestimate and underestimate some benefits

• Costs
  – Software purchase and maintenance
  – Staffing to develop and update
Website after redesign

• SAC had more presence on the web, including
  – Downloadable excel files
  – Instant Atlas tools (and later a Criminal History and Recidivism Tool)
  – Current projects descriptions
  – Publications
If the website still looked good, why change it?

People who accessed our website began accessing it from a mobile device….

…and that changed a lot for our organization’s website.
Technology Changes

• Smart phones became less of a luxury item and instead a common device
  – Smaller screens
  – Stylus’ were out, touch screens were in
• Applications needed to be “lightweight” and able to run on modern devices
  – War between Apple and Adobe over Flash (guess who won?)
  – HTML5 replaced Flash in many applications, including our tools
  – HTML5 does not have the accessibility problems that Flash did
• Code examples and libraries multiplied online, making developing easier
• We did not think we had fully engaged our audience with our content
Priorities

Hypothetical Scenario

User Base

Content Consumed
A variation of the Pareto Principle

80% of users

• May not know terminology
  – Prison vs. jail
  – Parole vs. probation
  – Arrests vs. offenses

• May not even know what they are really looking for

• May just need a single number for just their county/circuit, or many numbers

20% of users

• Are familiar with terminology

• Typically know what they are looking for and want a large amount of detailed, historical data

• May use web tools or may prefer to just have the data to download and analyze on their own
Priorities For 2nd Redesign

Hypothetical Scenario

We thought we had a lot here

We might be lacking on content here
Good News on the 2nd Redesign

• The Agency website underwent a considerable redesign, but the existing SAC content did not require much alteration
  – Data tools worked without interruption
  – Reports had a different entry page
  – We inventoried key words and tagged publications to improve search results

• We still needed to come up with a new way to condense our longer pdf reports into shorter articles

• We needed something to show our data in a simple format that a lay-person could understand, such as a county profile
Provide different products for different users

80% of users - What we provide

- Get the Facts and Walkthrough series
  - Short pdf’s that explain terminology and the processes of various part of the CJ system
- For our reports, a short HTML publication (new)
- Data profiles (new)

20% of users - What we provide

- Full reports that are fairly heavy on details and numbers
- Interactive mapping tools
- Criminal History and Recidivism Tool
- Downloadable data in excel
Product & Goals Examples

2010 - 2015

• Data downloads
• Interactive mapping and time series
• Accessible alternative under state guidelines
• Automated data profiles
• Better way to deliver reports
Producing Web Articles with Markdown and HighCharts

- Research staff (not professional web developers) use three tools to generate articles
  - Markdown
    - Allows you to type as if using a word processor and it automatically creates HTML
    - Free online version
      [https://stackedit.io/](https://stackedit.io/)
  - High Charts ([www.highcharts.com](http://www.highcharts.com))
    - Create interactive charts in JavaScript
    - Generally one can find an example and just edit the example
    - Free for non-commercial, government site requires a license ($150)
  - JsFiddle (free, [https://jsfiddle.net](https://jsfiddle.net)) acts as a temporary container/sandbox to test it
    - With the html generated from markdown and the charts included from high charts, one can send a single text file or a link to JsFiddle to webmaster.
    - Webmaster links the styling sheet, appearance will match the rest of the site
ICJIA Publications

**Articles**
- Short analysis project or a condensed version of a longer report
- May present a small number of graphics and tables
- Requires no download
- The sausage, not the sausage factory

**Reports**
- Long (often 100+ pages)
- Usually has numerous graphics and data tables
- Statistical models
- Downloadable pdf
- The sausage factory
Data Profiles

• What can users do if they want a lot of data for just one county, circuit, or the state?
  – The data can be downloaded in our excel files…but the user would have to string together dozens of files!
  – We used to produce county profiles in static pdf’s...took too long
• Original plan was to show criminal history database arrest records and prison data in aggregate formats
• Data structure allows us to show other data as well
• Coded in Cold Fusion and HighCharts using SJS grant funds
Current ICJIA Data and Data Tools

Targeting users based on their knowledge of the criminal justice system…

Novice Users
- Data Profiles
- Excel Files

Advanced Users
- Instant Atlas
- Criminal History and Recidivism Tool
Future ICJIA Data and Data Tools

Targeting users based on their knowledge of the criminal justice system…

Novice Users
- Data Profiles
- Excel Files

Advanced Users
- Instant Atlas
- Criminal History and Recidivism Tool
# Two advanced data tool options

## Instant Atlas Desktop

[www.instantatlas.com/](http://www.instantatlas.com/)

- Requires a license to be purchased
  - Optional upgrade and support annual fees
- Very easy to set up
  - Data can be imported via Excel or Access, can be in denormalized (wide) format in excel or normalized (Access)
- Several different templates (single map, scatter plot, dual time series, etc.)

## Tableau Public

[http://public.tableau.com](http://public.tableau.com)

- Free version allows up to 10gb of data in the cloud
- More difficult to set up, there are many guides and videos, as well as examples
  - Preferred data format is normalized (long)
  - Denormalized data is easier to use for certain charts
- Templates are less defined, but this allows more flexibility
Instant Atlas Basics

- Data manager to import data into; exports into format for website (json)
- Designer to build and arrange your templates
- Styles editor to customize colors, fonts, etc.
- Publisher ties this all together to build an atlas
Instant Atlas Data Manager

• Excel version (start with this)
• Access version (maybe move to this)
• Data in Instant Atlas is independent of the configuration file and styling files
  – Allows you to create one set of data files that can be used across several templates
  – Likewise, if you change the configuration/layout of your report, you probably don’t need to change the data (and vice versa)
Access Data Manager
Instant Atlas Designer Layout

- Drag and drop interface that has underlying config file
- Simple to use, no “For Dummies” manual needed
- Some knowledge of html/css/javascript useful, but not required
- If you see something another organization has and want the same, you can download their config.xml file
Illinois Uniform Crime Reports (I-UCR) Murder (and Nonnegligent Manslaughter) offenses reported to police by calendar year. Murder and Nonnegligent Manslaughter is the willful (nonnegligent) killing of one human being by another. Attempts are not included but are included as Aggravated Assaults. Rates are per 100,000 people.
Instant Atlas Publisher

• Requires
  – Map file
  – Template license
  – Data files from the data manager

• Generates a folder of files, which you deploy to web server. That’s it!

• When you update data but keep the template the same, you can just replace the data files on the web server
Tableau Public Basics

• Allows numerous types of charts, dependent on data
• Worksheets can be combined into dashboards and stories
• Drag and drop interface
• Be aware of what “Public” entails
Recommendations - Try both

• Instant Atlas has a free evaluation version that will allow you to develop and test
• Tableau Public is free if you can figure out how to use it and if the “Public” issues are ok with your organization
• Other options exist
Recommendation - Promote the Work of the SAC

• Develop social media accounts
  – Facebook
  – Twitter
• Maybe create videos of data analyses and/or presentations
  – I’m not sure people will actually watch these
  – If you need a video explaining how to use a tool, it’s probably too complicated
  – A video that shows interesting results by using a tool makes more sense
Recommendation - Use a database for storage instead of spreadsheets

• Flexible data structure (normalized)
• Enforces data integrity and consistency
• Updates are easier and safer
• Avoids storing calculated values and/or references that can go bad
• Very, very (!) large productivity increase for staff as they get used to database
## Data Preparation - Normalization

### Normalized Example Data

<table>
<thead>
<tr>
<th>Indicator_Number</th>
<th>County_Number</th>
<th>Data_Year</th>
<th>Data_Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>4</td>
<td>2011</td>
<td>244</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>2012</td>
<td>213</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>2013</td>
<td>203</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>2014</td>
<td>227</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>2001</td>
<td>25</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>2002</td>
<td>13</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>2003</td>
<td>10</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>2004</td>
<td>21</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>2005</td>
<td>29</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>2006</td>
<td>28</td>
</tr>
</tbody>
</table>

### Avoid storing data like....

<table>
<thead>
<tr>
<th>County</th>
<th>Year</th>
<th>Murder Arrests</th>
<th>Robbery Arrests</th>
<th>AggAssault Arrests</th>
<th>CrimSexAsslt Arrests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Will</td>
<td>2000</td>
<td>60</td>
<td>102</td>
<td>428</td>
<td>82</td>
</tr>
<tr>
<td>Will</td>
<td>2001</td>
<td>57</td>
<td>90</td>
<td>397</td>
<td>68</td>
</tr>
<tr>
<td>Will</td>
<td>2002</td>
<td>52</td>
<td>83</td>
<td>406</td>
<td>73</td>
</tr>
</tbody>
</table>

### ...Or like....

<table>
<thead>
<tr>
<th>County</th>
<th>Indicator</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>McHenry</td>
<td>Murder_Arrests</td>
<td>52</td>
<td>47</td>
<td>56</td>
</tr>
<tr>
<td>McHenry</td>
<td>Robbery_Arrests</td>
<td>89</td>
<td>99</td>
<td>85</td>
</tr>
<tr>
<td>McHenry</td>
<td>AggAssault_Arrests</td>
<td>421</td>
<td>444</td>
<td>443</td>
</tr>
<tr>
<td>McHenry</td>
<td>CrimSexAsslt_Arrests</td>
<td>83</td>
<td>73</td>
<td>80</td>
</tr>
</tbody>
</table>

You can **store** data like above but **present** it like the data to the right easily.
Example of problems from spreadsheets

- IDOC admission rates per 100,000 adults, age 17+
  - The rates were calculated using a reference to a population worksheet.
  - When someone updated the population worksheet, they pasted in the total population, including juveniles.
  - As a result, we have been using the wrong population for years!
- McHenry County left one judicial circuit and became it’s own circuit
  - Requires changing calculations in dozens of spreadsheets
- Both problems above required just changing a single entry in the database
- Database approach
  - For each indicator, define what the population is for a rate calculation. Store that in a table and use just that population value.
  - More advanced, for each indicator-year pair, store the population type to apply for a rate calculation.
  - Same thing for storing the judicial circuit a county belongs to
<table>
<thead>
<tr>
<th>Indicator_ID</th>
<th>Indicator_Desc</th>
<th>Population_Indicator_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>700</td>
<td>Prison Admissions - New Court</td>
<td>1000</td>
</tr>
<tr>
<td>1000</td>
<td>Total population</td>
<td></td>
</tr>
<tr>
<td>1008</td>
<td>Population age 17+</td>
<td></td>
</tr>
<tr>
<td>1410</td>
<td>Murder Offenses</td>
<td>1000</td>
</tr>
<tr>
<td>1411</td>
<td>Criminal Sexual Assault Offenses</td>
<td>1000</td>
</tr>
<tr>
<td>1412</td>
<td>Robbery Offenses</td>
<td>1000</td>
</tr>
<tr>
<td>1413</td>
<td>Aggravated Assault Offenses</td>
<td>1000</td>
</tr>
</tbody>
</table>
Keeping consistent data

- Storing data in one database keeps it consistent.
- Other things can connect to database and will have consistent values across different data products (with some delays just due to human work).
- Example: Our downloadable datasets are actually just a copy of an excel file linked to the database.
- New data steps:
  - Download data or extract from pdf
  - Normalize
  - Add to database
  - Refresh in Excel file
  - Run macros
  - Post online
- Example – total time taken to extract, transform and load 42 court indicators and produce downloadable excel files and updated data tools was about one day.
Questions?