

## **Inventing a Patent Database**

### **Lessons Learned while Creating the Wyoming Inventors Database**

**Dawn Rohan**  
**Intellectual Property Librarian & PTDL Representative**  
**Wyoming State Library**  
[drohan@state.wy.us](mailto:drohan@state.wy.us)

and

**Desiree Saunders**  
**Systems Librarian**  
**Wyoming State Library**  
[dsaund@state.wy.us](mailto:dsaund@state.wy.us)

#### **Abstract**

This paper discusses the creation of the *Wyoming Inventors Database* (WID) by the Patent and Trademark Depository Library (PTDL) at the Wyoming State Library. We discuss the process used to create the structure of the database and how we identified the data to include. Problems encountered during the creation are addressed to help others considering their own state patent database.

**Keywords:** Patents; Inventors; Database Creation, Wyoming State Library, Patent and Trademark Depository Library, PTDL

#### **Introduction**

Lack of a pre – 1976 bibliographic database from the U.S. Patent and Trademark Office (USPTO) increases the workload of the Patent and Trademark Depository Library (PTDL) and frustrates patrons who expect an easy electronic search of this important historic U.S. Patent Collection. A PTDL-created database is a wonderful outreach tool. Creating this database gave us a better understanding of the history of patents in Wyoming and the needs of our patrons.

Taking a cue from the State Library of Iowa's *Iowa Inventions Database* (<http://www.silo.lib.ia.us/app/cgi-bin/patents/>), librarians at the Wyoming State Library decided to create a state patent database for Wyoming. The Wyoming State Library became a PTDL in 2001 and the database was started shortly after Dawn Rohan was hired in the fall of 2001 to serve as the PTDL Representative and Intellectual Property Librarian. The database was completed in July of 2003.

The *Wyoming Inventors Database* (<http://cowgirl.state.wy.us/inventors/>) indexes U.S. patents in which one or more of the inventors were from Wyoming. The first patent in the database is patent no. 67,898 from 1867 which was issued to Anson Mills, a Captain of the Eighteenth Infantry stationed at Fort Bridger, Utah Territory. Fort Bridger is currently located in Uinta County, Wyoming.

Creating this database allowed us to get a handle on the patents that had come before and understand the needs of our patrons. For example, knowing that the majority of Wyoming patents are issued to independent inventors helped us to plan outreach and collection development efforts.

## **Creating the Database**

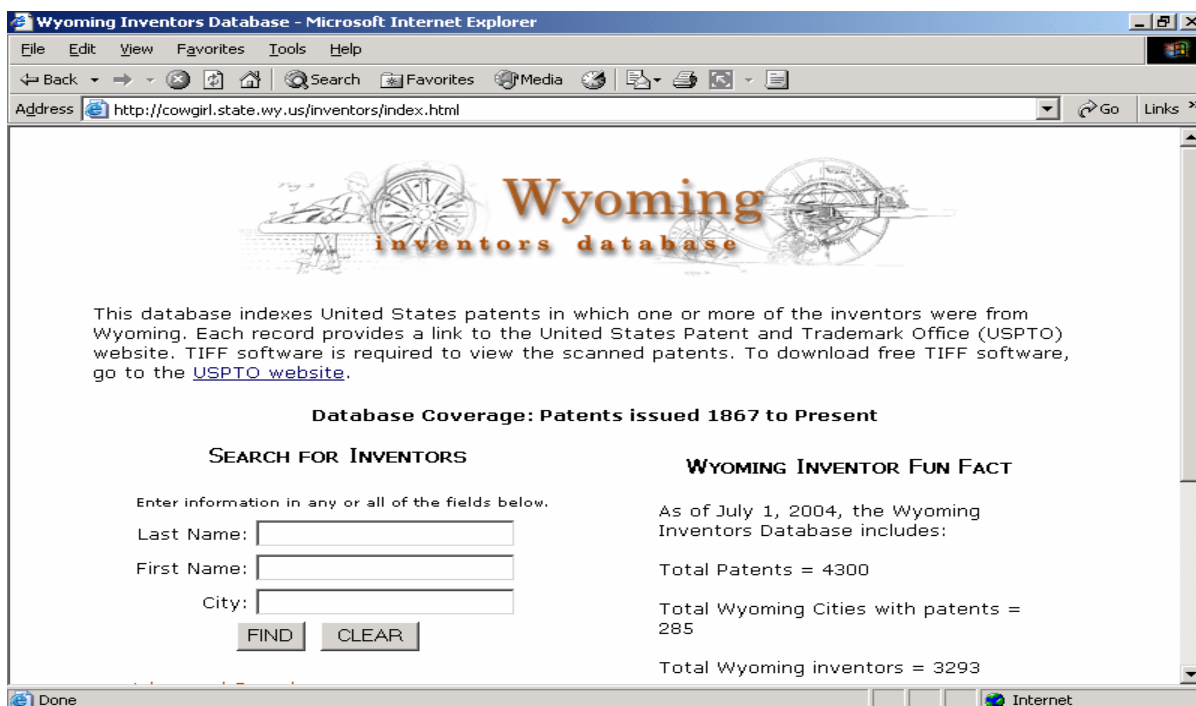
The creation of the WID has been an organic process. Because one of the systems librarians on staff, Desiree Saunders, has considerable experience turning local databases into web-based applications, she was consulted regarding initial data entry. This discussion led to a prototype of the database being built in Microsoft Access using the initial download of 1976 – present data from the USPTO's CASSIS2 database available at all PTDLs. This data served as the building blocks for the present database and introduced several issues for discussion regarding data normalization. In order to streamline the process of data entry, Desiree also built a web-based application for Dawn to enter her data so that she could forego learning Access initially and concentrate on the considerable task of gathering the data.

The process of entering data drove many changes in the application over a period of several months, including the addition of fields to facilitate data retrieval, sorting, and accuracy, as discussed in more detail later in this article.

Hand in hand with the creation of the data entry web forms, Desiree created Cold Fusion templates to handle SQL queries to the database from a publicly accessible website.

The WID's basic search screen supports searching of an inventor's last name, first name, and city or browsing by the first letter of the inventor's last name. The advanced search query allows an end-user to search all of the above fields as well as the patent number, assignee, or a range of dates. The Access database does not support complex queries, but attempts have been made to stretch the application's flexibility by supporting right hand truncation whenever possible. Search terms are passed to the database using standard SQL and the Boolean 'AND' operator.

After the initial mechanics of data retrieval had been worked out, the templates were turned over to Erin Kinney, WSL's Electronic Resources Librarian, to design page layout and consistent graphics. She is responsible for the clean, organized look of the application, see Figure 1.



**Figure 1.** Wyoming Inventors Database, Basic Search Screen  
<http://cowgirl.state.wy.us/inventors/index.html>

## Planning and Recommendations

Below are some planning steps recommended to other libraries based on our experience of dealing with problems both anticipated and unforeseen.

### 1. Decide What Data you Want to Include

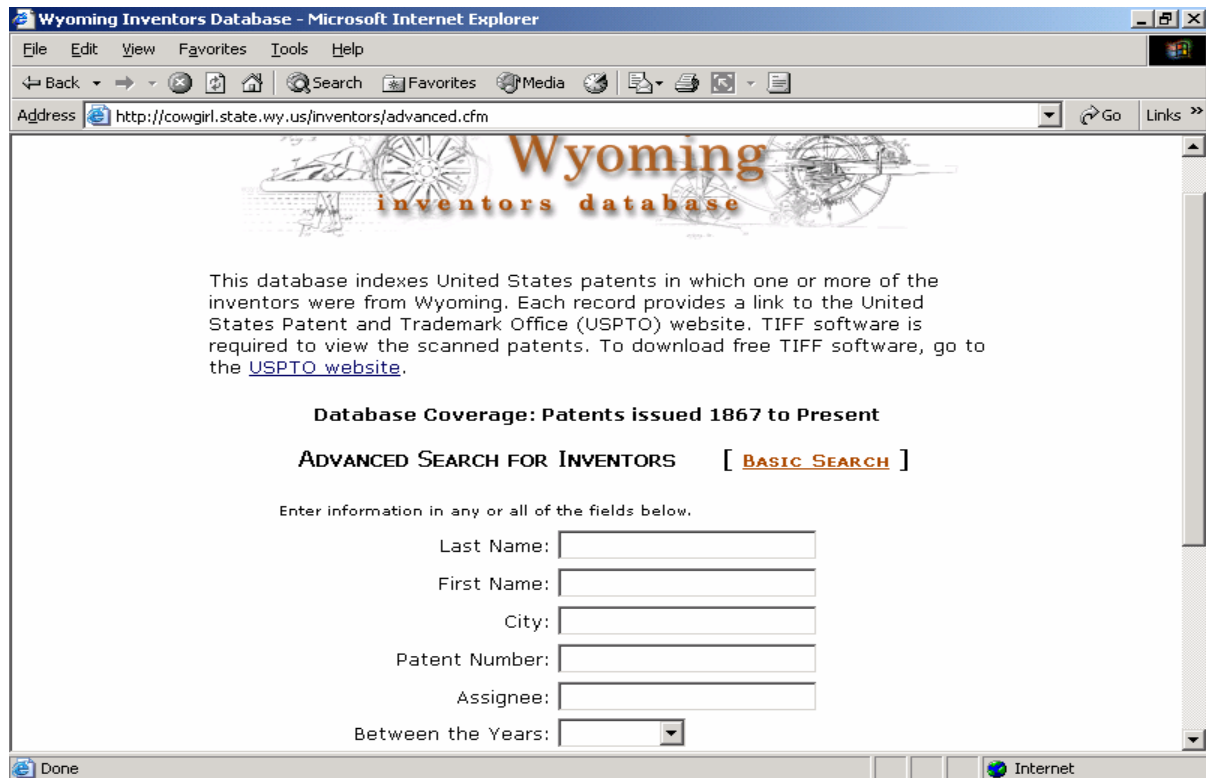
Our database included the following publicly viewable fields: Title, First Name, Middle Name, Last Name, Suffix (e.g. Jr.), City, State or Province, Assignee, Patent Number, Issue Date (searchable by 5 year periods), and Notes, see Figure 2. The following fields are not searchable: Title, Middle Name, Suffix, State or Province, and Notes.

We originally decided we did not need a state field since this was after all the *Wyoming* Inventors Database. However, we soon discovered that this affected the non-Wyoming co-inventors in the database. The initial web templates displayed Wyoming by default as the output for the state information. Once it became apparent that the state field was necessary, it was added to the database and the output of the search results template was modified to include that data so that an inventor from Denver, CO would not appear as an inventor from Denver, WY.

A similar yet originally unanticipated issue was documenting the country of residence for each inventor to accommodate patents with international co-inventors. Revisions to the database fields and output templates were necessary to resolve this problem.

We did not consider including counties as a search field because Wyoming only has 23. There are so few cities that if you mention one to a Wyomingite there is a good chance they will automatically know the county.

We decided to make the Title field non-searchable due to the vagueness of patent titles. **Figure 2** below shows the resulting database and search fields.



**Figure 2.** Wyoming Inventors Database, Advanced Search Screen  
<http://cowgirl.state.wy.us/inventors/index.html>

## 2. Define the Boundaries of your State

In defining the scope of the database, we decided to include all patents issued to one or more inventors who listed an address located in the present day boundaries of Wyoming. To do this, Dawn had to research the many territorial designations of Wyoming.

Wyoming Territory was established in 1868 and Wyoming became a state in 1890. Prior to this, sections of Wyoming were part of Dakota Territory, Idaho Territory, Utah

Territory, Territory of Lincoln, Oregon Country, District of Louisiana, Territory of Louisiana, Territory of Missouri, Indian Country, and Territory of Nebraska. Dawn searched each of these designations for patents issued to present-day Wyoming locales.

### **3. Be Prepared to Look at Thousands of Patents.**

Dawn has had to study thousands of patents to decide which ones to include and has eliminated approximately 50% of those considered initially.

## **Identification of Data Sources**

The sources used to retrieve data depended primarily on the State Library's collection. The following are the sources used to find the patents to add to the Wyoming database. They may not be the only sources other states use, especially states that have a longer history of patenting activity. As the explanations below illustrate, it is advisable to use a wide variety of sources to try to find the most information possible as well as to verify the data.

### **1. USPTO's CASSIS2 Database**

We decided to maintain a current database so the first set of data included was a download from CASSIS2 of the 1976 to Present data. Be aware that the decision to keep the database current requires extra effort on the part of the PTDL. The work on the database is never done. Since Dawn already compiles a monthly listing of patents issued to Wyomingites for distribution to newspapers, she just combined the two tasks.

### **2. Micropatent's 1836 OCR Backfile**

This database does not allow field searching which is how we discovered virtually every state has a county or city of Wyoming. We also had a lot of misreads. After a while, Dawn found consistent misreads and searched them, for example Wyoming could be spelled as wyoxing or wyoniing. Micropatent has not been perfect but well worth the money at a subscription price of \$495/year.

Micropatent did find patents we would not have found in the print indexes. For example, some inventors identified their location as "U.S. Army" which the print indexes noted. However, inside the document they may have identified where they were stationed, for example, "Fort Laramie, Wyoming Territory."

### **3. Geographical Index in the *Official Gazette*: Patents**

The *Official Gazette* has a Geographical Index that starts in 1965 and identifies patent numbers by state. We found many errors in this index, as well as with the patent count by state in the *Annual Report* below. Do not rely on them completely.

#### **4. USPTO's OCR database**

During Jim Miller's tenure as Fellowship Librarian at the Patent and Trademark Depository Library Program, he extracted a list of Wyoming patent numbers from the USPTO's OCR database that begins with 1920. This provided a backup for the Micropatent list. Like Micropatent, this list does not limit to a state field, but each OCR database found patents that the other did not.

#### **5. *Annual Report of the Commissioner***

This index lists patentees alphabetically and includes a patent count by state in the front. We often found more patents than listed in the count. Do not stop looking when you reach the number listed for your state.

This was, honestly, the most boring, tedious, and lengthy part of the database creation. This task needs to be broken up to avoid going into a daze looking at the tiny print.

#### **6. Technology Assessment and Forecast Geographical Profile Report**

This report listed all the patents issued to a specific state (identified by the first-named inventor) from 1969 to the mid-1990s. There were 33 patents listed in error in the Wyoming report.

#### **7. Get a Little Help from your Friends**

While creating the WID Dawn was contacted by retired PTDL librarian Donna Hanson who was doing some research of her own on Western patents. She was working forward through the *Annual Report* and Dawn was working backward. Their research results merged in the year 1909. Hanson's contribution filled in Micropatent's holes and cut short the tedious print index search.

### **Unanticipated Problems**

Changes to the data entry application as well as the publicly accessible website have been ongoing as we encounter unforeseen issues. Be prepared to be flexible and change the database as you go. The following are examples of problems that we have encountered:

#### **1. Location Misspellings**

Dawn entered data in the database exactly as it appeared on the front page of the patent but there were numerous city misspellings by the inventor or the USPTO. To facilitate data retrieval, we created an additional field in the database called 'Other WY Location'. This location does not appear in the output of search results, but is used by the 'City'

search field. The SQL query to the database looks in both the standard City location as well as Other WY Location when a City search is input.

Initially, the correct spelling went into Other WY Location and the misspelling that appeared on the patent went into the City field. Dawn felt strongly that the data should be entered exactly as it appeared on the patent. However, further discussions regarding data normalization became necessary when we began compiling statistics. Dawn decided to handle this by using the City field to display the correct spelling and the hidden 'Other WY Location' field to store the misspelled city name. Thus, the city of Greybull always appears as Greybull in the City field, but may include the misspelling of 'Graybull' in the Other WY Location field. Where these issues are encountered, Dawn has added a note to the public record, "City appears on patent as ...". This does put the responsibility on the data entry person to recognize city misspellings.

Also, we did not anticipate inventors living in Yellowstone, which went through many designations before it became a National Park. For example, we have a patent for Mammoth Hot Springs, National Park Reservation, Wyoming Territory. Nowhere does it say Yellowstone on the patent but Mammoth Hot Springs is its current headquarters.

This record received the most tweaking from Dawn, which was reluctantly done. Dawn contacted Yellowstone experts at the University of Wyoming library who confirmed that this location would have been called Yellowstone at the time the patent was applied for and issued. To keep the statistics on Yellowstone patents consistent, the city on this patent became Yellowstone Park. A note was added about the location as it appears on the patent and the Other WY Location field was Mammoth Hot Springs.

If it turned out that Mammoth was not called Yellowstone at that time, Dawn would have left the city as Mammoth Hot Springs. She did not want to change locations based on current designations. Part of the fun of the older patents is the history they evoke. That is why there are patents in the WID with a location of Shoshone Agency. It was not until later that it became designated the Wind River Indian Reservation. Dawn did add the Wind River Indian Reservation name to the Other WY Location, in case the searcher was unaware of its earlier name.

## **2. Name Misspellings and Changes**

Another issue similar to the city misspellings is obvious name misspellings or name changes. The way the database is set up makes creating an "Other Name Location" field more difficult than the Other WY Location, which handles city misspellings, since each part of the name (first name, middle, surname, etc.) is stored in a separate field in the database.

By obvious name misspellings, we refer to inventors who have multiple patents and unique names that are a letter or two off. Another case is when the patent identifies a

legal name change, for example, “Grace Ballard, now by marriage Grace Ballard Howell.”

To avoid slowing down the database by creating a searchable Other Names field, we added a “See also” note to the record and included advice in the FAQ on expanding searches. Some records received no changes, for example, Bessie M. Smith Walsh is still only searchable by Walsh and no note has been added since there is no extra record under Smith to view.

### **3. Patent Sorting**

The initial load of patents was merely sorted by patent number in the public application’s results screens. Since many of the numbers were sequential, it was not immediately apparent that we would need to sort by date. Once we started entering patents with five and six digit numbers, it became necessary to add a Sort Date field in the data entry application. The sort date field ensures that results screens will consistently sort patents by correct date.

### **4. Requests to be a part of the WID**

After the State Library issued a press release about the Wyoming Inventors Database, which was published in newspapers around the state, Dawn received an email from a Wyomingite who wished to be listed in the database even though his patent was issued while he was out of state. Although we had decided to limit patents included in the database to those that stated a Wyoming address on the front page, we decided to take his request as a compliment and include his patent in the database. His information was added with an explanatory note that he was a Wyoming native and resident.

### **5. Notes**

What notes were added to the Notes field were arbitrary and up to Dawn. She would have liked to have included more but did not want to bog down the database. The list below includes the types of notes that were used the most:

#### **Notes on the Inventor:**

- Citizen of / Subject of ... (e.g. Subject of the Czar of Russia)
- Deceased; “Joe Smith,” administrator/trix or executor/trix

#### **Notes on the Patent:**

- Assignee held x percentage
- City appears on patent as ...
- Reissue of Patent no. ...



- Patent reissued as Patent no. RE...
- Title on patent misspelled as ...
- First Wyoming patent, First Territory of Wyoming patent, etc.

## **Conclusions**

The *Wyoming Inventors Database* was a year and a half in the making but our plans for its use have expanded since its completion. We are analyzing the data for statistical information of use to economic development officials, as well as, planning a museum exhibit at the Wyoming State Museum for 2004-5, outreach to genealogy groups to demonstrate its potential for research, and publications on the historical data discovered about inventors in Wyoming.

## **References**

1. New from State Library: Index to Iowa Inventors (2001, June/July). *Footnotes: the Official Newsletter of the State Library of Iowa*, 25 (6/7), p. 5.