Hunting for Tigers

How to create a detailed regimental record from your archives

A Soldiers Database

Royal Hampshire Regiment Museum

archive@royalhampshireregiment.org

What, Why

- What: A database of soldiers who served in the Royal Hampshire regiment and its antecedents.

- Why: To help with management of the museum collection and biographical research.

How …

- It needs to cost as little as possible, preferably nothing at all:
  - In terms of financial cost of software.
  - In terms of skills and training needed to operate it. For us, this means volunteers with the right technical background.
- Start simple and leave the door open to further development over time:
  - Let volunteers work collaboratively, both within and (which is what presentation is about) between museums.
- Make it easy to use
  - Create utilities to help with loading and searching data.

… How

- The choice of Database Management System (DBMS)
  - We currently support two: **Microsoft Access** (because the Museum has it), and **Apache Derby** (because it's free).
  - It should be very little work (if any) to "port" the database to any other DBMS (such as PostgreSQL or MySQL)
- A standard XML format for loading or publishing soldiers' records.
- Software utilities.
  - These are Java and XSLT, but similar code could be developed in other programming languages.
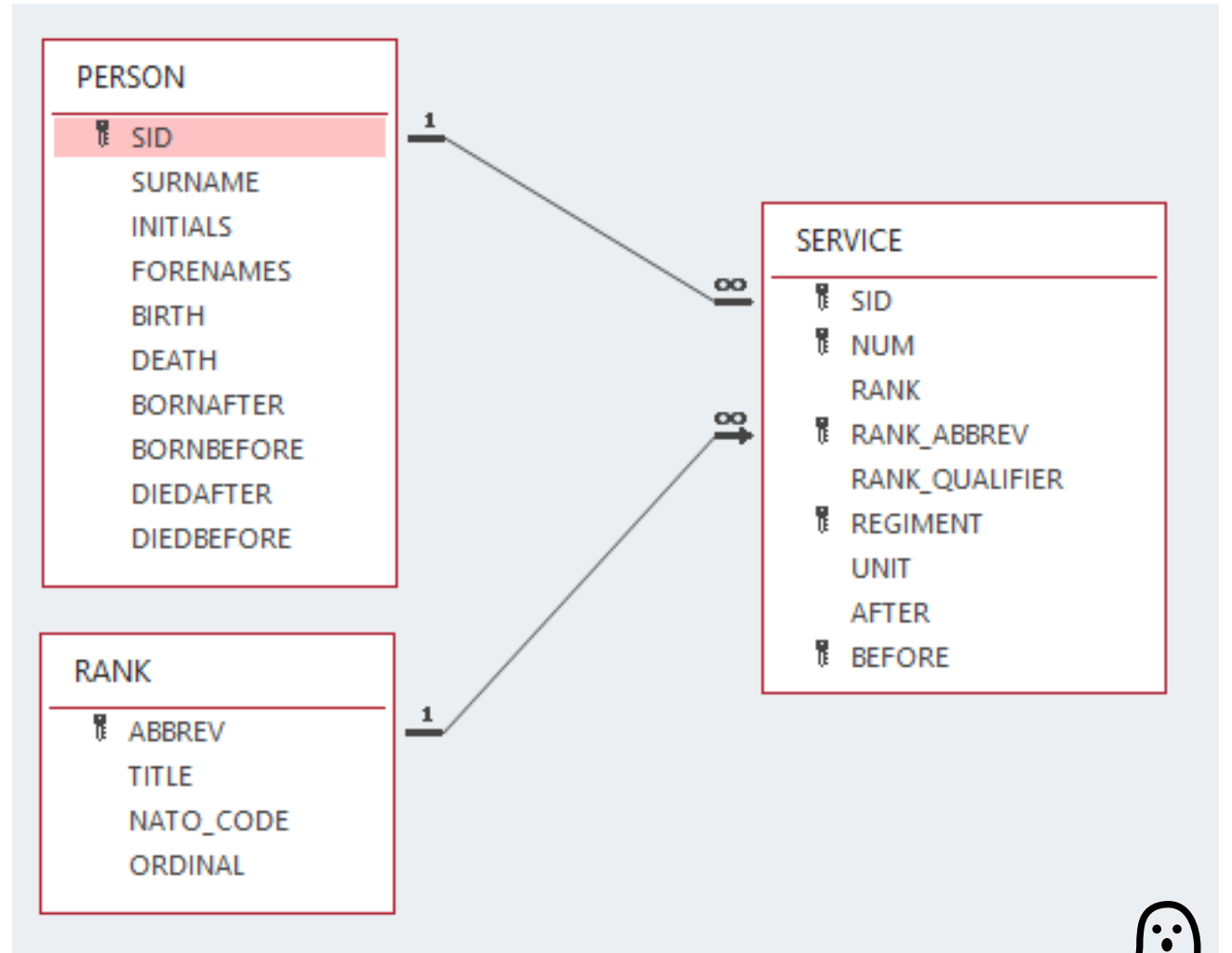- Publishing the above to GitHub

Relational Schema

Name, rank and number …

Each soldier is assigned a unique Soldier ID, must have at least a surname, and may have any number of associated service records.

"before" and "after" dates can be used if exact dates aren't known.

Sources

- Commonwealth War Graves Commission
- IWM Lives of the First Word War
- Medical Records (WOCL, SWB, Casualty ledger, Hospital Books)
- Medal Rolls
- Volunteer research
- Museum research and accessions
- …

Soldiers XML

```xml
<person>
  <surname>CARR</surname>
  <forenames>FREDERICK GEORGE</forenames>
  <initials>F G</initials>
  <service>
    <record number="20772" rank="Lance Corporal" regiment="Hampshire Regiment" before="1918-03-27"/>
  </service>
  <birth after="1917-02-27" before="1918-04-27"/>
  <death date="1918-03-27"/>
  <note source="Commonwealth War Graves Commission">
    <memorial reference="Bay 6." country="France">ARRAS MEMORIAL</memorial>
  </note>
</person>
<person>
  <surname>MOOR</surname>
  <forenames>GEORGE RAYMOND DALLAS</forenames>
  <initials>G R D</initials>
  <service>
    <record number="" rank="Lieutenant" regiment="Hampshire Regiment" before="1918-11-03"/>
  </service>
  <birth after="1895-10-03" before="1896-12-03"/>
  <death date="1918-11-03"/>
  <note source="Commonwealth War Graves Commission">
    <memorial reference="C. 86." country="France">Y FARM MILITARY CEMETERY, BOIS-GRENIER</memorial>
    <additional>Son of William Henry and Eva Helen Moor (nee Pender).</additional>
  </note>
</person>
```

Populating the database

- Insert records
  - Assuming the soldier isn't already in the database
  - Assuming the information is correct
  - Can do this in bulk from a file (Soldiers XML)
- Quality control
  - Is the database correct? Can we ever be sure? Do we have the same soldier twice? It is unlikely the database will ever be fully complete or correct, but we can work to test and refine it.

Using the database

Searching for a soldier

Sources are uncertain. We need to be able to cope with misspelt names and errors in service numbers. Forenames aren't usually given. Initials may be incomplete.

The approach is to find candidate matches and select the best. Start with specific queries, then follow-up with more general queries if those don't produce results: surname and number, number only, surname only. Need a threshold for accepting near matches.

Currently, automatic matching can cope if the name is misspelt or the number is not quite right, but not both at the same time.

GitHub

https://github.com/

A tool for shared development over the internet, providing *source code management* and *version control*.

GitHub is free to use if your work is public, i.e. *Open Source*.

You don't need an account to view or download what's there. You do need an account to create or modify content.

Soldiers DB can be found at:

https://github.com/tigersmuseum/soldiers