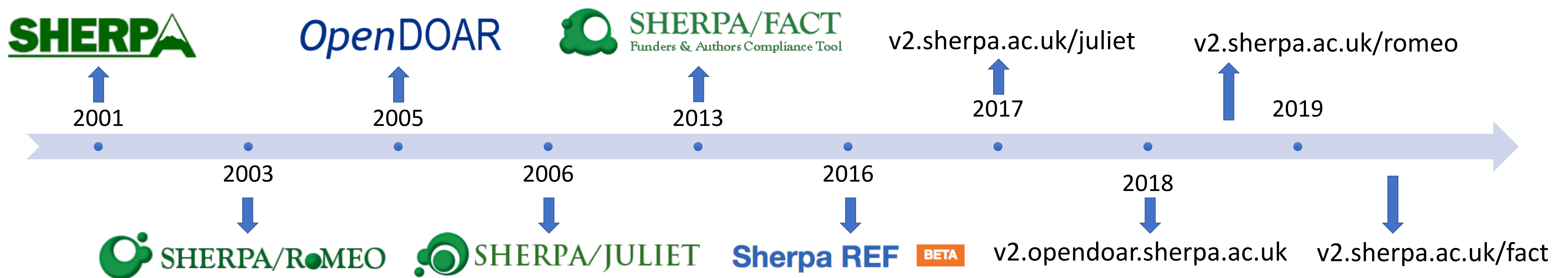


Introducing the New Sherpa Data Model and API

Adam Field, Jisc, adam.field@jisc.ac.uk



Key Challenges

Legacy Codebase: In 2016, our services had been active for up to 15 years. As our services evolved to meet the growing needs of the community, our codebase grew to support this. However, as best practices have moved on, parts of our systems became challenging to maintain and, critically, to extend.

Siloed Services: Our services were designed to provide human interfaces to our datasets, with APIs added later to support machine-to-machine transfer of data. This had led to a set of service silos with little crossover. These separate codebases created an additional technical overhead to running our services and added complexity to the joining up our datasets.

Organic Data Model: As the need for an API became apparent, we already had existing human interfaces and a de facto data model. Our APIs made this available as XML, but much of the data (e.g. embargoes) were stored as human-readable text for our human interfaces. This has resulted in our data not being as processable as it could be.

The v2.sherpa Project

We started coding v2.sherpa in January 2017. The overarching goal of the project was a like-for-like replacement of our services on **new infrastructure** based on a **single codebase** with a **common data model**. Built with an equal emphasis both **User and Machine Interfaces**, v2.sherpa would allow us to provide agile responses to changes in the open access landscape.

Over the past 18 months, the project has delivered rebuilds of **Juliet** and **OpenDOAR**. Development work on **RoMEO** is expected to complete later this year, which will be followed by a significant editorial effort to make our metadata more **machine-processable**.

Each service, at launch has a single API endpoint enabled, the **Object Retrieval API**. We have exciting plans for further improvement of our platform once we have completed this round of development.

The Object Retrieval API

[https://v2.sherpa.ac.uk/cgi/retrieve?item-type=funder&api-key=F1028D5C-3FC8-11E7-9056-BF1D1A23D8E2&format=Json&filter=\[\[{"id","equals","695"}\]\]](https://v2.sherpa.ac.uk/cgi/retrieve?item-type=funder&api-key=F1028D5C-3FC8-11E7-9056-BF1D1A23D8E2&format=Json&filter=[[{)

CGI Parameters

- item-type:** The type of the object to retrieve
- api-key:** Your API key. Every user account on v2.sherpa gets one
- format:** The format of the output
- order:** A comma separated list of fieldnames defining the order
- limit:** The number of records to retrieve, max 100
- offset:** The offset of the first record, for paging
- filter:** A JSON array specifying filter conditions

The Filter Parameter

Enables basic search functionality by applying up to three filters to a collection of records. Each filter is a JSON array of three values; a fieldname, an operator and a value. Examples:

```
[[{"id","equals","57"}]]  
[[{"name","contains word","cancer"}, {"country","equals","ca"}]]
```

Your API Key

To use our API, you will need to register for an API Key. Every user account gets one automatically. Accounts are created at:

<https://v2.sherpa.ac.uk/cgi/register>

API Documentation

<http://v2.sherpa.ac.uk/juliet/api.html>

Returned JSON

```
{  
  "items": [  
    {  
      "funder_metadata": {  
        "id": "695",  
        "name": [  
          {"language": "en", "name": "Wellcome Trust"}  
        ],  
        "country": "gb",  
        "identifiers": [  
          {  
            "type": "fundref",  
            "identifier": "http://dx.doi.org/10.13039/100004440"  
          }  
        ],  
        "url": [  
          {  
            "url": "http://www.wellcome.ac.uk/",  
            "language": "en"  
          }  
        ]  
      },  
      "open_access_publications_policy": {...},  
      "open_access_data_archiving_policy": {...},  
      "system_metadata": {  
        "date_modified": "2017-05-17 14:05:50",  
        "uri": "http://v2.sherpa.ac.uk/id/funder/695",  
        "publicly_visible": "yes",  
        "date_created": "2013-01-23 10:19:03"  
      }  
    }  
  ]  
}
```

Multilingual Fields

External Identifiers

Rich, Machine-Readable Data (not shown)

Standard System Metadata