Resolving DS-3602

Problem Summary

Solr statistics records created before DSpace 6.x are still indexed with their legacy id.

The following PR will migrate legacy statistics records to use DSpace 6x UUID values: https://github.com/DSpace/DSpace/pull/1774

Because this PR requires a SOLR schema change, we have deferred this fix until DSpace 7x.

Iterim Solution for DSpace 6x

DSpace 6.0 had some code in place that will search for either a UUID or a legacy id when displaying usage records. This logic was incomplete.

The following PR addresses that issue: https://github.com/DSpace/DSpace/pull/1782

Unfortunately, when statistics records are retrieved by facet query separate counts are returned for records indexed by legacyId and records indexed by UUID.

Item Page

In the following record note that only one bitstream is present.

A-Guide to the Historical Bioethics Newsletter Collection
Norton, Charles N; Dellasanta, Nicholas

URI: http://hdl.handle.net/10822/711246
Date: 2014-11-06

Description:
You can request use of this collection by visiting https://bioethics.georgetown.edu/using-the-library/item-hold-request-form/ and filling out the form.

Show full item record

Files in this item

Name: A Guide to the ...
Size: 162.4Kb
Format: PDF

View/Open

Usage Statistics

The resulting usage information lists the item twice and the bitstream twice. A modification has been made to annotate the legacy statistics records.

Statistics

Total Visits

	Views
A-Guide ...	6
A-Guide ...(legacy)	1

Total Visits Per Month

	January 2017	February 2017	March 2017	April 2017	May 2017	June 2017	July 2017
A-Guide ...	0	0	0	0	0	0	6
A-Guide ... (legacy)	0	0	0	0	0	0	1

File Visits

	Views
A Guide to the Historical Bioethics Newsletter Collection.pdf(legacy)	2
A Guide to the Historical Bioethics Newsletter Collection.pdf	1

Top country views

Views

Top cities views

Views

Item Statistics

```
{
  "responseHeader": {
    "status": 0,
    "QTime": 1,
    "params": {
      "fl": "id",
      "indent": "true",
      "q": "type:2 AND (id:fdb92c32-5312-406a-a1c2-ebe75c743e8b OR id:10)",
      "_": "1501007655254",
      "wt": "json"
    }
  },
  "response": {
    "numFound": 7,
    "start": 0,
    "docs": [
      {
        "id": 10
      },
      {
        "id": "fdb92c32-5312-406a-a1c2-ebe75c743e8b"
      },
      {
        "id": "fdb92c32-5312-406a-a1c2-ebe75c743e8b"
      },
      {
        "id": "fdb92c32-5312-406a-a1c2-ebe75c743e8b"
      },
      {
        "id": "fdb92c32-5312-406a-a1c2-ebe75c743e8b"
      },
      {
        "id": "fdb92c32-5312-406a-a1c2-ebe75c743e8b"
      },
      {
        "id": "fdb92c32-5312-406a-a1c2-ebe75c743e8b"
      }
    ]
  }
}
```

Bitstream Statistics

```
{
  "responseHeader": {
    "status": 0,
    "QTime": 1,
    "params": {
      "fl": "id,owningItem",
      "indent": "true",
      "q": "type:0 AND bundleName:ORIGINAL AND (owningItem:fdb92c32-5312-406a-a1c2-ebe75c743e8b OR owningItem:10)",
      "_": "1500503770019",
      "wt": "json"
    }
  },
  "response": {
    "numFound": 3,
    "start": 0,
    "docs": [
      {
        "id": 41,
        "owningItem": [
          10
        ]
      },
      {
        "id": 41,
        "owningItem": [
          10
        ]
      },
      {
        "id": "cb850cc2-6124-43bb-ad3f-dc3ed2f740e7",
        "owningItem": [
          "fdb92c32-5312-406a-a1c2-ebe75c743e8b"
        ]
      }
    ]
  }
}
```

A fix for this bug would require changes to the following code. Much of this code is quite old and it appears to need significant refactoring.

https://github.com/DSpace/DSpace/blob/master/dspace-api/src/main/java/org/dspace/statistics/content/StatisticsDataVisits.java#L204-L399

I recommend that we create a new ticket to refactor this code in DSpace 7.x. Perhaps the changes for DS-3602 should be merged in the meantime.