Examples of how Alan Kelly's data was coded

1. **A straightforward example**

Kelly's data as provided by him in Microsoft Word

DEREK OLDHAM (organ)

Cc 20633   Semele: Where'er you walk (Handel)
   -1   21-10-30
   -2
   -3   32-2014 C2146

As automatically tagged in XML by CHARM's partners CCH

```
<head>
        <performer>DEREK OLDHAM (organ)</performer>
        <venue>Kingsway Hall, London</venue>
</head>
<record xml:id="MAT201_17367">
        <catNum></catNum>
        <matrix><num>Cc 20633-1</num></matrix>
        <date when="">21-10-30</date>
        <title>Semele: Where'er you walk<composer>Handel</composer></title>
</record>
<record xml:id="MAT201_17368">
        <catNum></catNum>
        <matrix><num>Cc 20633-2</num></matrix>
        <date when="">21-10-30</date>
        <title>Semele: Where'er you walk<composer>Handel</composer></title>
</record>
<record xml:id="MAT201_17369">
        <catNum></catNum>
        <matrix><num>Cc 20633-3</num></matrix>
        <date when="">21-10-30</date>
        <title>Semele: Where'er you walk<composer>Handel</composer></title>
        <note>32-2014 C2146</note>
</record>
```

As it appears via a search in the CHARM discography

Composer: Handel, Work: Semele: Where'er you walkHandel, Performer: DEREK OLDHAM (organ), Date: 21-10-30
Catalogue**: Kelly**

Performer: **DEREK OLDHAM (organ)**
Venue: **Kingsway Hall, London**
Num: **Cc 20633-1**
Date: **21-10-30**
Title: **Semele: Where'er you walk**

Composer: **Handel**

Composer: Handel, Work: Semele: Where'er you walkHandel, Performer: DEREK OLDHAM (organ), Date: 21-10-30
Catalogue**: Kelly**

Performer: **DEREK OLDHAM (organ)**
Venue: **Kingsway Hall, London**
Num: **Cc 20633-2**
Date: **21-10-30**
Title: **Semele: Where'er you walk**
Composer: **Handel**

Composer: Handel, Work: Semele: Where'er you walkHandel, Performer: DEREK OLDHAM (organ), Date: 21-10-30
Catalogue**: Kelly**

Performer: **DEREK OLDHAM (organ)**
Venue: **Kingsway Hall, London**
Num: **Cc 20633-3**
Date: **21-10-30**
Title: **Semele: Where'er you walk**
Composer: **Handel**
Note: **32-2014 C2146**

Notes to this example:
-Each of the three takes (differentiated by the take number -1, -2 and -3 at the end of the matrix number) has a separate entry in the discography. Only the third has issue numbers given in the 'Note' because only that take was issued.
-This entry cannot be found by searching on Work = 'Where'er you walk' because the title as tagged turns out to be 'Semele: Where'er you walk', which one might not have guessed. Work = 'Semele' will return it correctly. 'Where'er you walk' will find it, however, in a Free Text Search. More refined tagging (identifying both works and extracts from works) would be desirable, but see the end of the next example to understand why that's not been done yet.

2. A more faulty example

Kelly's data as provided by him in Microsoft Word

The Pirates of Penzance - Complete recording (continued)   (see 15906 above)

Cc 16262   SHEFFIELD, DAWSON, BAKER, GRIFFIN, GILL,
CHORUS
                    To gain a brief advantage
    -1 25-3-29   -1A
    -2   -2A 2-04719 D1688 (set 22)   -2A

(o MALCOLM SARGENT) **[see also 15906]**

Cc 16262 The Pirates of Penzance: To gain a brief advantage (Gilbert/Sullivan), with **PETER DAWSON, DOROTHY GILL, ELSIE GRIFFIN, GEORGE BAKER, LEO SHEFFIELD**

-1 25-3-29		[-1A]
-2	-2A 2-04719 D1688 (set 22)	[-2A]

Then automatically tagged by CCH, and checked by CHARM staff for major errors only, to produce

<head>
  <performer>(orchestra) <conductor>MALCOLM SARGENT</conductor> [see also 15906]</performer>
  <venue>C Studio, Small Queen's Hall, London</venue>
</head>
<record xml:id="MAT201_7802">
  <catNum/>
  <matrix>
    <num>Cc 16262-2</num>
  </matrix>
  <date when=""/>
  <title>The Pirates of Penzance: To gain a brief advantage<composer>Gilbert/Sullivan</composer>,
   with PETER DAWSON, DOROTHY GILL, ELSIE GRIFFIN, GEORGE BAKER, LEO SHEFFIELD</title>
  <note>-1 25-3-29 [-1A] -2A 2-04719 D1688 (set 22) [-2A]</note>
  </record>

Leading to the search results

Composer: Gilbert/Sullivan, Work: The Pirates of Penzance: To gain a brief advantageGilbert/Sullivan, with PETER DAWSON, DOROTHY GILL, ELSIE GRIFFIN, GEORGE BAKER, LEO SHEFFIELD, Performer: (orchestra) MALCOLM SARGENT [see also 15906]
Catalogue**: Kelly**

Performer: **(orchestra)**
Conductor: **MALCOLM SARGENT**
Performer: **[see also 15906]**
Venue: **C Studio, Small Queen's Hall, London**
Num: **Cc 16262-2**
Title: **The Pirates of Penzance: To gain a brief advantage**
Composer: **Gilbert/Sullivan**
Title: **, with PETER DAWSON, DOROTHY GILL, ELSIE GRIFFIN, GEORGE BAKER, LEO SHEFFIELD**
Note: **-1 25-3-29 [-1A] -2A 2-04719 D1688 (set 22) [-2A]**

The good news is that all the data has displayed. But you can see that <composer> has split the <title> data into two separate entries, and that the second is really performer data which should have been manually edited and put into <performer>. The cross-reference to matrix Cc 15906 should have been expanded and put into

<note>. The data for take -1 has mistakenly been tagged within <note>, along with the date, and both of those should have been manually edited to produce a complete preceding <record> for take -1. There should have been a <date> field in each entry containing the recording date (25-3-29). The duplicate recordings made simultaneously on a second set of recording machinery, producing matrices numbered -1A and -2A, should have been in <notes> for their respective <record>s. Clearly more editing of the XML would have produced a better result. But please bear in mind that this single XML file for Kelly's MAT 201 (the Bb/Cc list) runs to 414,181 lines of code, and that for the Kelly data alone there are 35 files in all (some considerably longer). Now you will see why not everything could be made perfect!