

Recommendation for a Consolidated Election Results Portal for the Scottish Local Elections Report Files

Based on: Scottish Local Government Elections May 2022

Des McNelis

10/09/24

Abbreviations

PR = Proportional Representation

STV = Single Transferable Vote

LGA = Local Government Authority

LG-STV-PR elections = Local Government STV-PR

RO = Returning Officer

BLT = Preference Profile files (file extension often used is .BLT – BaLlLoT)

H/W = Hardware

S/W = Software

There are three sections to this report:

1. A basic introduction and run-through of the election day processes and the aftermath.
2. The case for a Centralised Repository for the Publication of Scottish Local Elections Report Files.
3. The documentation of the many inefficiencies, errors and inconsistencies I have discovered during my research.

Section 1 – A basic introduction and run-through of the election day processes and the aftermath.

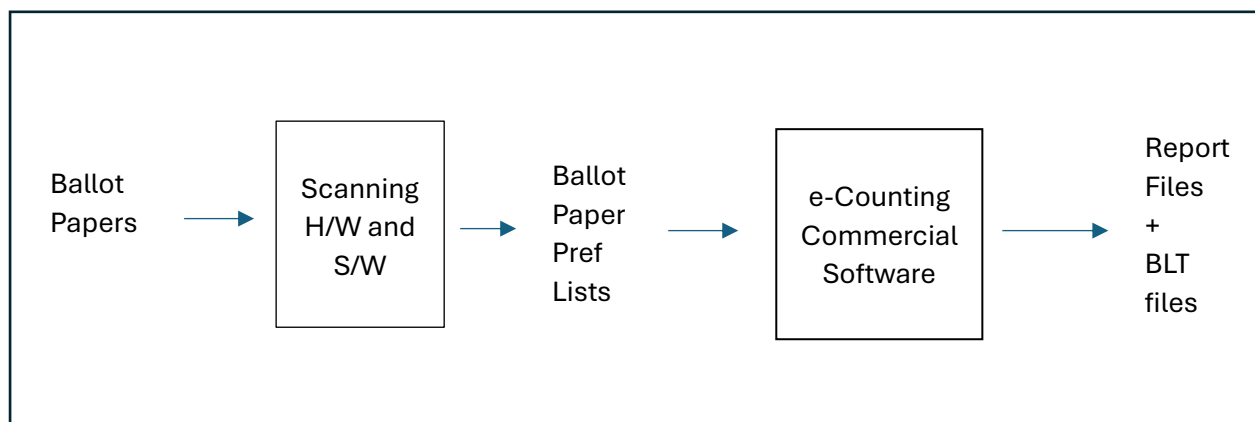
Introduction

As part of research I was performing on the Scottish Local Government Elections in May 2022, I required the Report files (especially, the BLT files) for all the wards for all the 32 Scottish Councils as generated by the idox eCounting software. It is the responsibility of each of the councils to publish (for both viewing and download) all of these files for that council. However, most of the councils only published a subset of these – and in a most inconsistent manner. I had expected to download all the files for a council as a ZIP container file – but only a handful of the councils offered this capability. EMBS did at least have all the files at one point in time – but retained them only for three months or so as per some agreement. Regardless, there are many files unpublished.

I was then forced to download all the files myself. This was an extremely laborious and error-prone exercise – compounded by the very poor naming of the various files making it very difficult to identify the files' contents. From a data analysis perspective, the Preference Profile (BLT) file is far-and-away the most important file. It is therefore very unfortunate that this file is published using different applications resulting in different file extensions (*.BLT, *.CSV, *.PDF, *.XLSX and *.TXT) – sometimes leading to corruption while attempting to convert to a basic text file.

The following diagram describes the journey ballot papers take through the Election e-Counting System. The BLT file structure is documented later.

Diagram 1: Election e-Counting System flow.



The set of 8 files to be published.

Each of the 32 councils has to offer for download the following eight (sometimes the last file is ignored) report files:

1. Declaration of Results Report
2. Candidate Votes per Stage Report
3. Transfers Report
4. First Preference Report
5. Preference Summary Report
6. Preferences by Ballot Box Report
7. Preference Profile file (the BLT file)
8. How to understand the Preference Profile

The first five report files are reasonably well processed.

The Preferences by Ballot Box Report is often missing. A possible explanation for *some* of these omissions could be that there were too few voters for a particular ballot box (there is an election rule stating that, in these cases, the report should not be produced because of privacy concerns). However, there should still be a “nothing to report” file produced.

The “How to understand the Preference Profile” document is often poorly presented – it’s a standard “boilerplate” document with the council’s election team’s contact email and phone number – but in many cases that has simply been set to XXXXXX.

The Preference Profile (the BLT file) appears to have a very strange life. It is presented with several different file extensions (*.BLT, *.CSV, *.PDF, *.XLSX and *.TXT) – perhaps the file contents are scraped from some screen.

The BLT file.

The BLT file is a basic text file, although it is internally formatted in a particular style. It is generated by the e-Counting Software after the scanning of the ballot papers and the STV-PR computation has been completed.

There is one BLT file produced for each ward. The BLT file contains information on how voters ranked the candidates for the ward.

The BLT file contains *all* the data sufficient to (re)build the following election reports:

- Candidate Votes per Stage Report
- Transfers Report

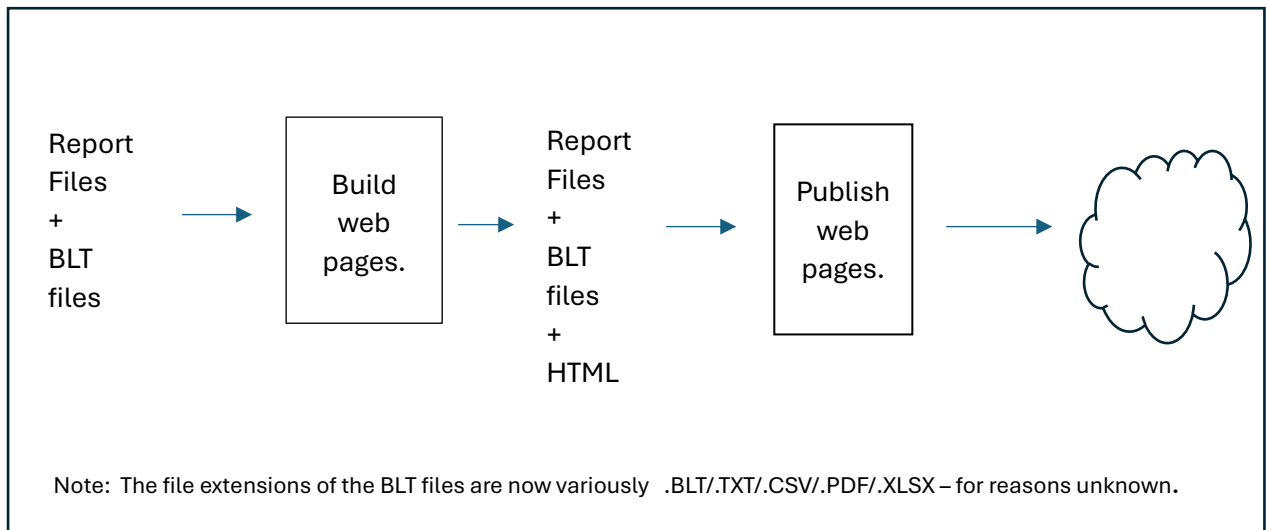
The BLT file is of particular import to the following stakeholders:

- Political party analysts.
- Academic researchers and analysts.

There follows a full description of the data contained within a BLT file.

The following diagram describes how the files output by the Election e-Counting software are ultimately published on each of the 32 councils' websites.

Diagram 2: Publication of Report and BLT files on 32 Council websites.



The data files for the 32 Scottish Councils can be viewed by visiting each of the 32 websites individually.

Here are the 32 direct hyperlinks:

<https://stv.prorep.org.uk/introduction/the-32-scottish-councils>

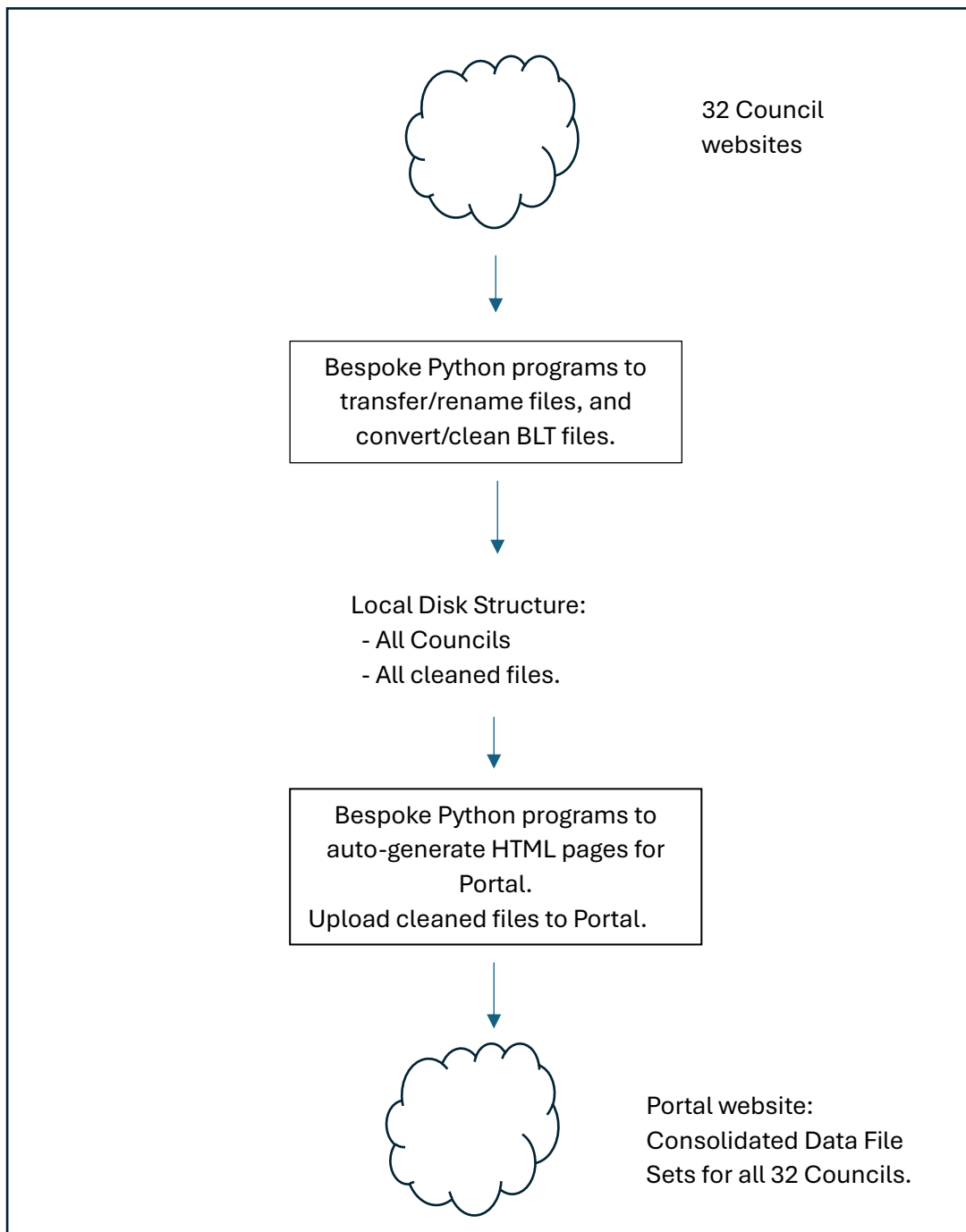
A brief glance, now, will demonstrate the inconsistency in look-and-feel and navigation.

I have downloaded all the available report and BLT files from all of the 355 wards. These files were then cleaned-up (both filenames and content) and then used to build a dedicated website which then allows a consistent and easy method of viewing and downloading the files (also in ZIP containers at both national and council level).

Consolidated Election Results Portal: <https://stv-results.prorep.org.uk>

The following diagram illustrates this process.

Diagram 3: Build Portal website: Consolidated Data File Sets for all Councils.

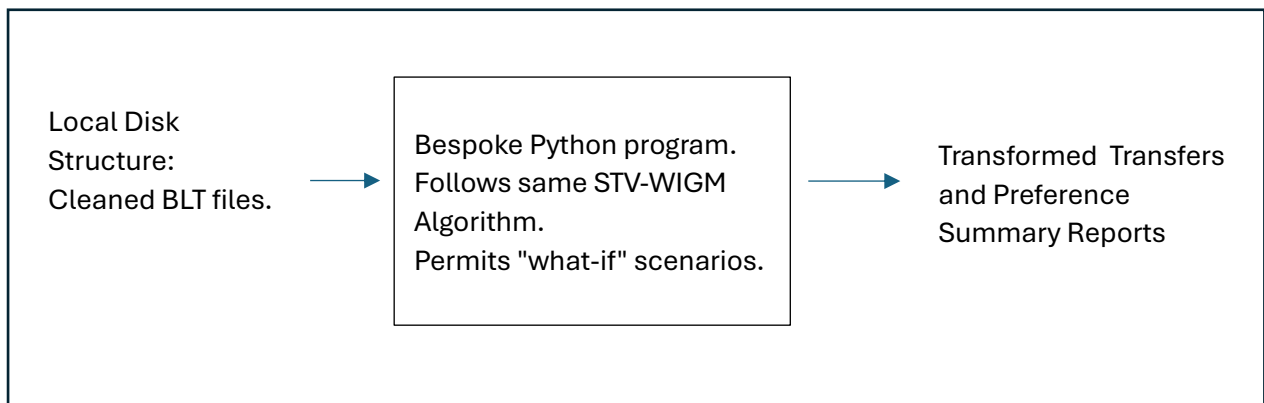


Here is an example of the power of this work.

I have written a Python program which emulates the same STV-WIGM Algorithm used in the Scottish Local Government Elections May 2022.

The Local Disk Structure is used to re-run the Scottish Local Government Elections May 2022 – but now with different input parameters to enable various “what-if” scenarios. Only the BLT files are required.

Diagram 4: Bespoke Python program to re-run Scottish Local Government Elections May 2022 with different parameters.



We now have the capability to re-run the Scottish Local Government Elections with a varying number of preferences. We can then compare the report outputs of the parameterised election re-run with the original reports to establish matching results. A match occurs when the same set of candidates is elected for a ward – re-run versus original.

I have now run my Python program for all the wards for all the 32 councils. This work is the subject of a separate report.

Section 2 – The case for a Centralised Repository for the Publication of Scottish Local Elections Report Files.

From diagrams 2 and 3 we can see that there is considerable scope for commonalising and standardising this work.

It is very clear that there has been no attempt at a standardised approach for the 32 councils. In particular, the BLT file is processed and distributed in many ways – for reasons unknown.

It is important to appreciate that my solution auto-generates the Centralised Repository website. This is achieved programmatically using the very popular and powerful programming language Python. The required HTML is generated programmatically.

Currently, the main tasks to be performed by each of the councils – for each ward – are:

1. Setup and (re)configure website election area for publishing report files.
2. (Re)design template web pages for the report files.
3. Gather all report files:
 - a. Probably involves special treatment for Preferences by Ballot Box Report and Preference Profile file (the BLT file).
 - b. Probably involves renaming of files.
4. Build a mapping table of the filenames versus files.
5. Build local disk structure to hold report files.
6. Update filenames.
7. Update HTML in web pages with correct report filenames.
8. Populate local disk structure with report files.
9. Transfer files from local disk structure to temporary website folder structure.
10. Test and correct temporary website.
11. Switch on new website.
12. Test new website.

Recommendation: Build a Virtual Team to:

1. Take ownership and responsibility for all the tasks (defined above) currently performed by each of the 32 councils.
2. Refine the work I have completed so far.
3. Generalise it – as appropriate.
4. Expand on it – if appropriate e.g. bespoke reports. However, it is very important that any exuberance is tamed. Keep it Simple.

The Virtual Team.

The Virtual Team would need to only contain a handful of (mainly) technical IT people. It would start to become more active around one year before an impending election. This would include some involvement in the testing of the particular eCounting software selected. There should be the opportunity to have a full run-through of the system – in particular that the report files output by the eCounting software are sufficient for the Centralised Repository for the Publication of Scottish Local Elections Report Files. This latter work does not appear to have been performed previously.

The service could be funded by introducing a small annual charge to each of the 32 councils.

To be clear, once the main recommendation is accepted, what I have labelled as “Local Disk Structure” in diagrams 2 and 3 will now become the “Centralised Repository”. At election time, the various councils will only need to email their raw files to the Virtual Team.

There is a powerful and convincing case:

1. Gross inefficiency and high costs – there are 32 councils attempting to perform the same, significant workload. Mostly poorly.
2. Inconsistency – there are frequent examples of missing files, files with internal corruption, inconsistencies in file-naming, etc. There is no aspect of similar look-and-feel in the 32 council websites.
3. There is no Centralised Repository for the Publication of Scottish Local Elections Report Files – stakeholders have to visit several/all websites – all with very different presentation and navigation. Downloading files is often very tedious and error prone. Or simply not possible.
4. The Virtual Team will now be better skilled to quickly and efficiently resolve any issues during the critical period of updating and publishing results after the election.
5. Any related general or FOI data requests are far easier to satisfy.
6. Any future changes to the electoral system e.g. boundary changes, will now be far easier to consider and implement properly and efficiently.
7. The central website can be better protected against malicious actors. (The entire Comhairle nan Eilean Siar website has been unavailable due to hacking from November 2023 until at least March 2024).
8. It is now far easier to comply with any regulatory requirements – with confidence.
9. There are issues with National Language Support (NLS) and File Encoding. This is compounded by the 32 councils manipulating files and publishing the reports in different ways. See [Appendix 3](#) . A general solution is required.

I am happy to donate the high-quality Python code I have produced and tested – as referenced to in Diagram 3 (but not Diagram 4). I am also happy to donate some hours in a handover exercise.

Section 3 – The documentation of the many inefficiencies, errors and inconsistencies I have discovered during my research.

This section is necessarily more technical.

The main issue in the early stages of automation is the correct identification of a file's contents given its raw filename as exported from the eCounting software.

There are three “driver” files used by the Python programs:

1. List_Councils.txt – contains the static data concerning the 32 Councils.
2. List_Council_Wards.txt – contains the static data concerning the 355 Wards.
3. List_Report_File_Specs.txt – contains the mappings between the raw filenames and the various Report types.

A glance, now, at these files will help us to better understand the following processes.

The attached driver file List_Councils.txt contains many comments concerning errors discovered by me and necessary corrections made by me.

The attached driver file List_Report_File_Specs.txt contains some comments concerning issues discovered by me.

There is absolutely no convention or standardisation or consistency in the naming of the raw files published by the various councils.

Here are some examples (from 5 different councils) of the raw filenames for the "Declaration of Results" report:

```
DeclarationOfResultsReport_V0002_Dyce-Bucksburn-Danestone-Ward_06052022_112506.pdf
declarationofresultsreport_v0001_ward-1-banff-and-district_06052022_102710.pdf
Declaration of Results Report - Ward 1 Kirriemuir and Dean.pdf
DeclarationOfResults.pdf
file141890.pdf  !!!!
```

I find it very difficult to believe that the eCounting software does not have the capability to define a naming convention, perhaps something like:

```
<Report_name>__<Ward_ID>__YYYYMMDD_HHMMSS.pdf
e.g.
Declaration_of_Results_Report__Ward_1__Kirriemuir_and_Dean__20230506_023548.pdf
```

Embedded spaces in filenames can cause issues when working with filenames at the Operating System level. The underline character is often used to replace spaces in these circumstances.

From an extract from List_Report_File_Specs.txt we can see how the mappings between the raw filenames and the various Report types are enforced:

```
# Before name-string comparison the following editing is performed:
# All SPACES, underscores and hyphens are removed.
# The remaining characters are converted to UPPER-case.
# So, "00QA_Dyce-Bucksburn-Danestone_Ward_PREFERENCES-BY-BALLOT-
BOX_V2.pdf" will match "PreferencesByBallotBox", and be classified as
a "Preferences by Ballot Box Report".
#
F01,"DeclarationOfResults","Declaration of Results"
F02,"CandidateVotesPerStage","Candidate Votes per Stage Report"
F03,"Transfers","Transfers Report"
F04,"FirstPreference","First Preference Report"
F05,"PreferenceSummary","Preference Summary Report"
F06,"PreferencesByBallotBox","Preferences by Ballot Box Report"
F07,"PreferenceProfile","Preference Profile"
```

After a raw filename has been identified as a particular report, then a prefix is added by a Python program to fully identify the file easily and consistently by later programs:

```
DeclarationOfResultsReport_V0003_Ward-1-Linn_06052022_122617.pdf
-->
C16_W01_DECLARATIONOFRESULTS__DeclarationOfResultsReport_V0003_Ward-1-Linn_06052022_122617.pdf
```

C16 is the code for “Glasgow City”; W01 is the code for Ward 1.

"DECLARATIONOFRESULTS" now consistently defines a "Declaration of Results Report".

The Preference Profile (the BLT file) appears to have a very strange life. It is presented with several different file extensions (*.BLT, *.CSV, *.PDF, *.XLSX and *.TXT) – perhaps the file contents are scraped from some screen. There again, it could be that the eCounting software offers a drop-down list of file types. This all needs to be resolved by the time of the next election.

Relational Database Management System (RDBMS).

I intentionally used a flat-file text “driver” methodology for reasons of simplicity and because of the many inconsistencies and missing data encountered in gathering the required data files from the various council websites. Going forward, I’m not convinced that adopting an RDBMS (e.g. MySQL, MariaDB, MS SQL Server, Oracle) would be sensible unless significant improvements are made to the generation and consistent naming of the underlying data files.

References

My Consolidated Election Results Portal (All Councils): <https://stv-results.prorep.org.uk>

My shortcuts to the Election results areas of the 32 Scottish Councils:

<https://stv.prorep.org.uk/introduction/the-32-scottish-councils>

Election office contacts – Electoral Management Board for Scotland

<https://www.emb.scot/reports/election-office-contacts>

Acknowledgements

I would like to thank Dr James Gilmour for his advice and support over the last 12 months or so.

Attachments

- 3 List_*.txt driver files.
- Sample output: Transfers Report - My text style - Glasgow.txt . View this file with a fixed-pitch font like Courier New. Or open with a basic text editor like Notepad or Notepad++.

Appendix 1 Bespoke PDF – How to understand the Preference Profile file.

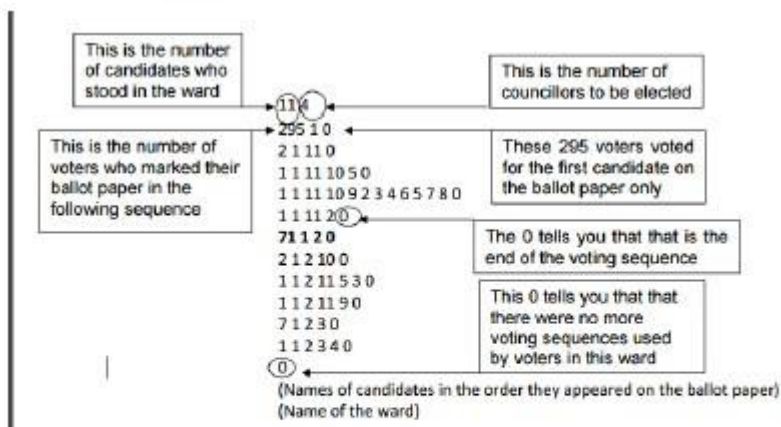
This file is listed for each ward for each council. Its contents are mostly this:

HOW TO UNDERSTAND THE VOTING INFORMATION

The Returning Officer has made voting information from the 2022 council elections available to anyone who wishes to find out more about voters ranked the candidates in order of preference. There is one report for each ward.

The first line of the report shows how many candidates there were and how many councillors were to be elected. The subsequent lines begin with the number of voters who ranked the candidates in the same order. There then follows the order that those voters used.

For example, the row in bold below shows that 71 voters gave their first preference to the candidate who appeared first on the ballot paper, and their second preference to the candidate who appeared second. The names of the candidates in the order they appeared on the ballot paper is given at the end of the report. The figure '0' means that there were no more preferences given to any candidate by voters who used this sequence. The final '0' on the last line shows that the report has finished and no more sequences were used.



If you have any questions about the reports, please contact the Elections Office on XXXXX or email XXXXXX.

My understanding is that this was a template produced by EMBS. Clearly, it was supposed to have been updated by/for each council. I think a more helpful name for the document would be "How to understand the Preference Profile file".

The associated filename is often understand_the_preference_profile.pdf .

My recommendation would be to have two national-level documents:

- An updated document called "How to understand the Preference Profile file" based on the above document, but without the Elections Office contact details addendum.
- A new document called "Scottish Council Election Offices Contacts". This single document would contain the contact details for all the 32 councils: perhaps simply Contact Number and Email Address.

There would need to be only two such files on the website at a top-level. Each ward would simply have pointers to these two files.

The data required for the new document is already available at:
Election office contacts – Electoral Management Board for Scotland
<https://www.emb.scot/reports/election-office-contacts>

Appendix 2 A Case Study – An example of how not to do it.

I apologise most profusely if I have misunderstood any of this.

Some of the councils decided to publish extra content in their area allocated to election results. My opinion is that this work – often quite significant in effort – added little or nothing to the functionality of the website.

There is at least one extreme case which I will use as an example of wasted effort: [here](#)

Comments:

1. Only one file is available directly for download – the Preference Profile (the BLT file) – and even then, only as an .XLSX file (which it is not).
2. The BLT has the required double-quotes missing from the Candidates' names – and the ward's name – thus preventing any parsing.
3. The other reports appear to have been manually transcribed from the original report files into manually created web pages. This has been done very poorly e.g. the Transfers Report – Ward 1 Milngavie:
 - a. Columnisation has been lost – numbers are left-aligned (like text).
 - b. Important data has been lost – there are no figures for Stage 2 – it is completely blank.
 - c. No listing of the two candidates elected (MATHIESON and POLSON) at the end of the final stage 5. That appears to be a consistent issue across all wards.
 - d. There are several other errors.
4. There is a link to “Transfer Animation [opens in a new window]”. This links to Gerry Mulvenna's website <https://election.indylive.radio> – which is itself based on another website: “Credit to @electionsNI : This site took its starting point from the work carried out by Bob Harper (@bobdata on twitter) for the #AE17 assembly election in Northern Ireland, which is available at <http://electionsni.org> “. The base data for this website has been derived from many well-meaning contributors transcribing figures from many sources (e.g. from displays in election counting rooms) – and is thus of debatable quality. However, this is all well-intentioned work.

Appendix 3 National Language Support (NLS) and File Encoding.

This is a very complex and involved area. Here's an example.

Let's assume that this is the correct rendition of Mr Massie's middle? name (I think it's Gaelic for "the island of Lewis") <https://www.glasgow.gov.uk/LeodhasMassie>



Ward: Newlands/Auldburn(2)
Party: Scottish Green Party

Biography

However, when we inspect the Declaration of Results report (Glasgow Ward 2 - Newlands-Auldburn [here](#)) we see a different, incorrect rendition:

Glasgow City Council

Declaration of Results

This report is the Declaration of Results script for Returning Officer.

Contest Name Ward 2 Newlands-Auldburn

I will now declare the result of the election in the Ward 2 Newlands-Auldburn. There are 3 Councillors to be elected.

The electorate was 18,963 and the number of ballot papers received was 7,779 giving a percentage turnout of 41.0%. There were 214 rejected ballot papers. The quota of votes for a candidate to be elected was 1,892.

The number of first preference votes for each of the candidates was:

Candidate Name	Affiliation	Number of First Preference Votes
Stephen CURRAN	Glasgow Labour	1,928
Linda DEVLIN	Glasgow Labour	554
Sean FERGUSON	Scottish National Party (SNP)	2,074
Leòdhas Iain MASSIE	Scottish Greens - Delivering For Our Community	638
Rage RAGE	Scottish National Party (SNP)	428

Lastly, if we look at the corresponding Preference Profile (BLT) we see ([here](#))

```
1 7 3 5 4 10 9 1 2 6 0
1 3 5 4 6 9 0
0
"Stephen CURRAN" "Glasgow Labour"
"Linda DEVLIN" "Glasgow Labour"
"Sean FERGUSON" "Scottish National Party (SNP)"
"LeÃfÃ³dhas Iain MASSIE" "Scottish Greens - Delivering For Our Community"
"Rage RAGE" "Scottish National Party (SNP)"
..
```

This rendition is different again – and certainly incorrect. This disparity is due to issues related to National Language Support (NLS) and File Encoding.

File encodings map from bytes to character renditions e.g. decimal 65 maps onto “A”.

Over the years, there has been an evolution in NLS – from ASCII through Code Pages through UNICODE through UTF. These are attempts to offer a general and extensible solution to the problem.

For all of this to work, it is critical that the whole life journey of the data – from input into the earliest data consumer – is honoured throughout.

Complicating this further, whichever data viewer a user is utilising can impact the experience. Typical viewers are web browsers (e.g. Google Chrome, Safari, MS Edge); applications (e.g. MS Word/Excel); raw text editors (e.g. Notepad++).

At the Operating System level – when working with filenames – I believe there is no need to attempt to embrace NLS – it just causes unnecessary problems. It’s more sensible to use A-Z and a-z. So, something like the sub-string Ward_10_Steornabagh_a_Deas (where the “o” had a grave accent) is appropriate.

It’s much more difficult at a higher level e.g. reports and web page text.

I have adopted UTF-8 (no BOM) throughout in the creation of all text and HTML files – which should also be the case for the BLT files generated by whichever eCounting software is used.

Technical: Here’s a hexadecimal dump of the foot of this particular BLT file:

```
000088f0 4e 22 20 22 53 63 6f 74 74 69 73 68 20 4e 61 74 N" "Scottish Nat
00008900 69 6f 6e 61 6c 20 50 61 72 74 79 20 28 53 4e 50 ional Party (SNP
00008910 29 22 0d 0a 22 4c 65 c3 83 c2 b3 64 68 61 73 20 )".. "LeÃfÃ³dhas
00008920 49 61 69 6e 20 4d 41 53 53 49 45 22 20 22 53 63 Iain MASSIE" "Sc
```

This error could have occurred *anytime* on or after the capturing of the candidate’s registration data.

All I’m doing at this stage is highlighting the issue and providing some background. In the interim, I may investigate further and come up with further insight.

Appendix 4 Software used.

Here is a list of the software I have used in this project.

Software used:

- My Workstation Operating System: Windows 11.
- Basic File Downloading/Renaming/Correction
 - Directory List & Print Pro.
 - Bulk Rename Utility.
 - UltraEdit.
 - Excel.
- Basic File Content Analysis
 - UltraEdit.
 - Excel.
 - EncodeAnt
- Programming
 - Python (IDE: Anaconda/Spyder).
 - R (IDE: RStudio) (in the early days only).
- My Main Website
 - WordPress on Linux.
 - FileZilla.
 - PuTTY.
 - LibreOffice Draw (Graphics).
- My Simple Website A working template for the Centralised Repository.
 - Linux.
 - CoffeeCup.
 - FileZilla.
 - PuTTY.
- Websites hosted by IONOS (1&1).

--- End ---