

Hacking Organograms: Unlocking government data, 28 July 2017

This document holds links relevant to (and from!) [the hackday](#) hosted by the Institute for Government on government organograms on 28 July. Please feel free to add, edit and comment! (There's also a Twitter thread summarising the project [here](#).)

You can also reach us at whitehallmonitor@instituteforgovernment.org.uk.

Background

- [Hacking organograms: unlocking government data](#): or, why we hosted the day in the first place (Institute for Government, 5 July 2017)
 - [The data](#) for central government departments (collated by the Institute for Government from data.gov.uk and GOV.UK, as of 5 July 2017, .xlsx)
 - [Data for other government organisations](#) - use the menu and then 'Source data' (data.gov.uk)
- [Organogram data - tech notes](#) (David Read)

Other useful links

- [Whitehall Monitor 2017](#) – the Institute for Government's annual data-driven view of government, published January 2017.
- [ONS Public Sector Employment](#) – quarterly civil service staff numbers (latest IfG analysis [here](#))
 - We classify some public bodies as part of their parent department because they are line-managed by them – our classification system is available [here](#)
- [ONS Annual Civil Service Employment Survey](#) – annual data on civil service staff numbers, gender, grade, age, profession, pay, ethnicity, disability... (latest IfG analysis [here](#), previous analysis in individual comment posts from [here](#) back)
 - To help analyse professions, we've classified the many civil service professions into smaller groups – that will be available tomorrow should you need it

- [Civil Service People Survey](#) – annual survey of staff engagement (latest IfG analysis [here](#))
- Workforce Management Information – monthly releases by department of on- and off-payroll staffing, usually found on GOV.UK by department (e.g. [here](#) for the Home Office)
- Annual reports and accounts – published by individual departments on GOV.UK (e.g. here for [DWP](#)). The 2016/17 reports are starting to appear – departments can publish as late as January 2018
- [Accounting officer statements](#) – published by individual departments on GOV.UK, these give some sense of organisation structure
- [Public Bodies](#) – annual Cabinet Office publication with details of most types of public bodies (though [the government register](#) has a longer list which includes some high-profile units within departments)
- List of ministerial responsibilities - [PDF](#) currently out of date (data as of October 2016), although [the GOV.UK Ministers page should be current](#).

The hackday

- [#IfGhackday on Twitter](#).

Projects

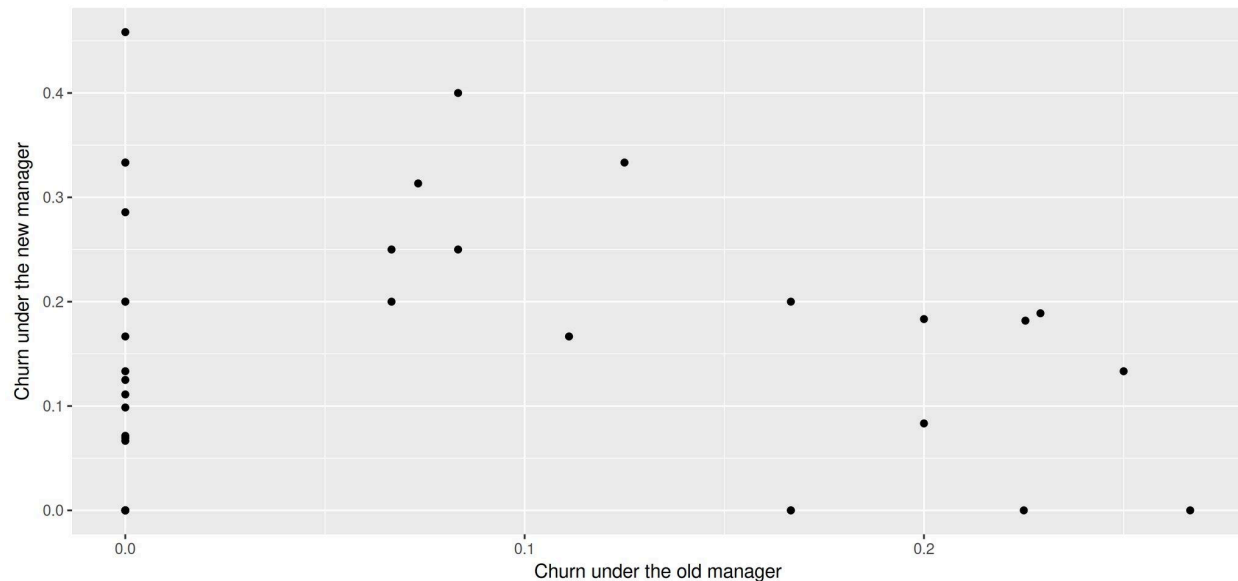
Please add any links, explanations of what you did, comments etc here:

GDS

First, when the manager changes, is there an exodus from their team? Or does the new manager stop the rot? Here we plotted churn, the percentage of people in the team who left after a new manager was appointed. Each point represents the average quarterly churn before and after the new manager. The point in the top-left shows that churn was zero under the old manager, but when the new manager arrived, people started leaving.

Churn before and after a change of manager

The column of points above zero on the x-axis shows teams where everything was just fine until the new manager arrived.
The row of points along zero on the y-axis shows teams where the new manager stopped the rot.



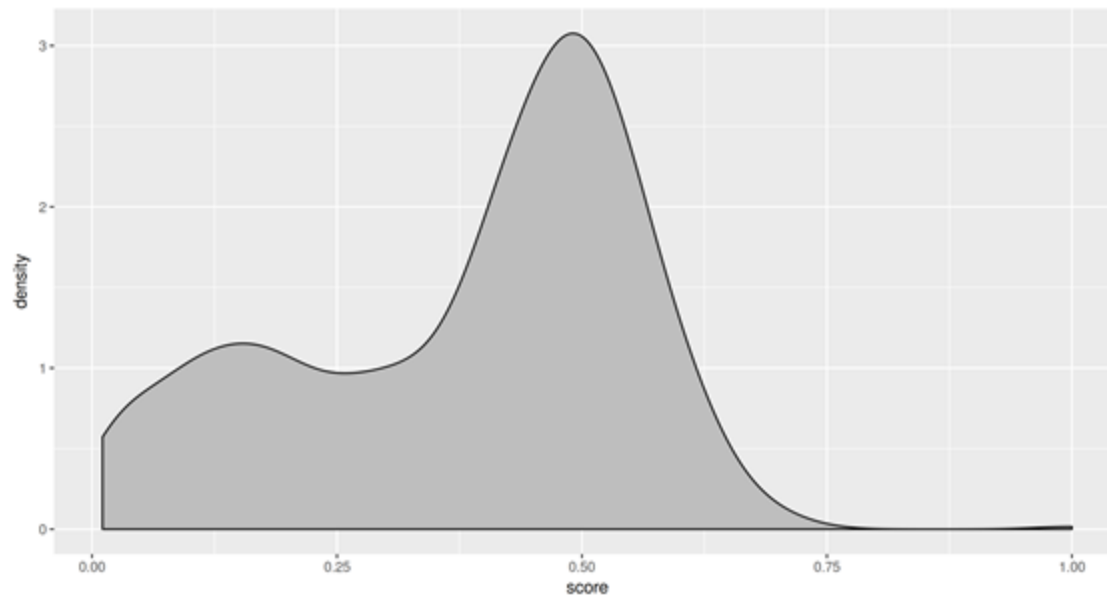
The next one takes a bit of explaining. To track individuals over time, we needed to recognise them by name. But the names could be different from quarter to quarter. For example:

John Smith
Smith, John
Smith, J
JOHN SMITH
John Smith, CBE

Our approach was this:

1. The standard tricks: convert everything to lowercase, remove punctuation.
2. Sort each name alphabetically within a full name, e.g. "smith john" becomes "john smith", and "tony stark" becomes "stark tony" consistently.
3. Find matches using the [Jaro-Winkler](#) algorithm.

That last step computes a score for each potential match, so we had to choose a threshold at which we would accept the match. To make that decision less arbitrary, I graphed the distribution of J-W scores, hoping that it was bimodal -- i.e. two clear peaks, one at a low (better) score for names that should match, and another at a high (worse) score for names that really didn't match.



Perfect, so we chose a threshold of .35, somewhere in the dip between peaks. That seemed to be about right. In the table below, the 'score' is the difference between one name, and the name in the previous row. Elizabeth Gardiner's name changes in almost every quarter, but the score is always below .35 so we would treat all those names as the same person. But when 'adrian hogarth' is compared with 'elizabeth gardiner cb', the score is .49 -- clearly a different person.

```
# # A tibble: 12 x 5
# # Groups:   org, post [1]
#   org post quarter name score
#   <chr> <chr> <date> <chr> <dbl>
# 1 CO 186 2011-09-01 elizabeth gardiner NA
# 2 CO 186 2012-03-01 e gardiner 0.2814815
# 3 CO 186 2012-09-01 elizabeth gardiner 0.2814815
# 4 CO 186 2013-03-01 e gardiner 0.2814815
# 5 CO 186 2013-09-01 elizabeth gardiner 0.2814815
# 6 CO 186 2014-03-01 e gardiner 0.2814815
# 7 CO 186 2014-09-01 elizabeth gardiner cb 0.3079365
# 8 CO 186 2015-03-01 elizabeth gardiner cb 0.0000000
# 9 CO 186 2015-09-01 elizabeth gardiner cb 0.0000000
# 10 CO 186 2016-03-01 adrian hogarth 0.4950397
# 11 CO 186 2016-09-01 adrian hogarth 0.0000000
# 12 CO 186 2017-03-01 adrian hogarth 0.0000000
```

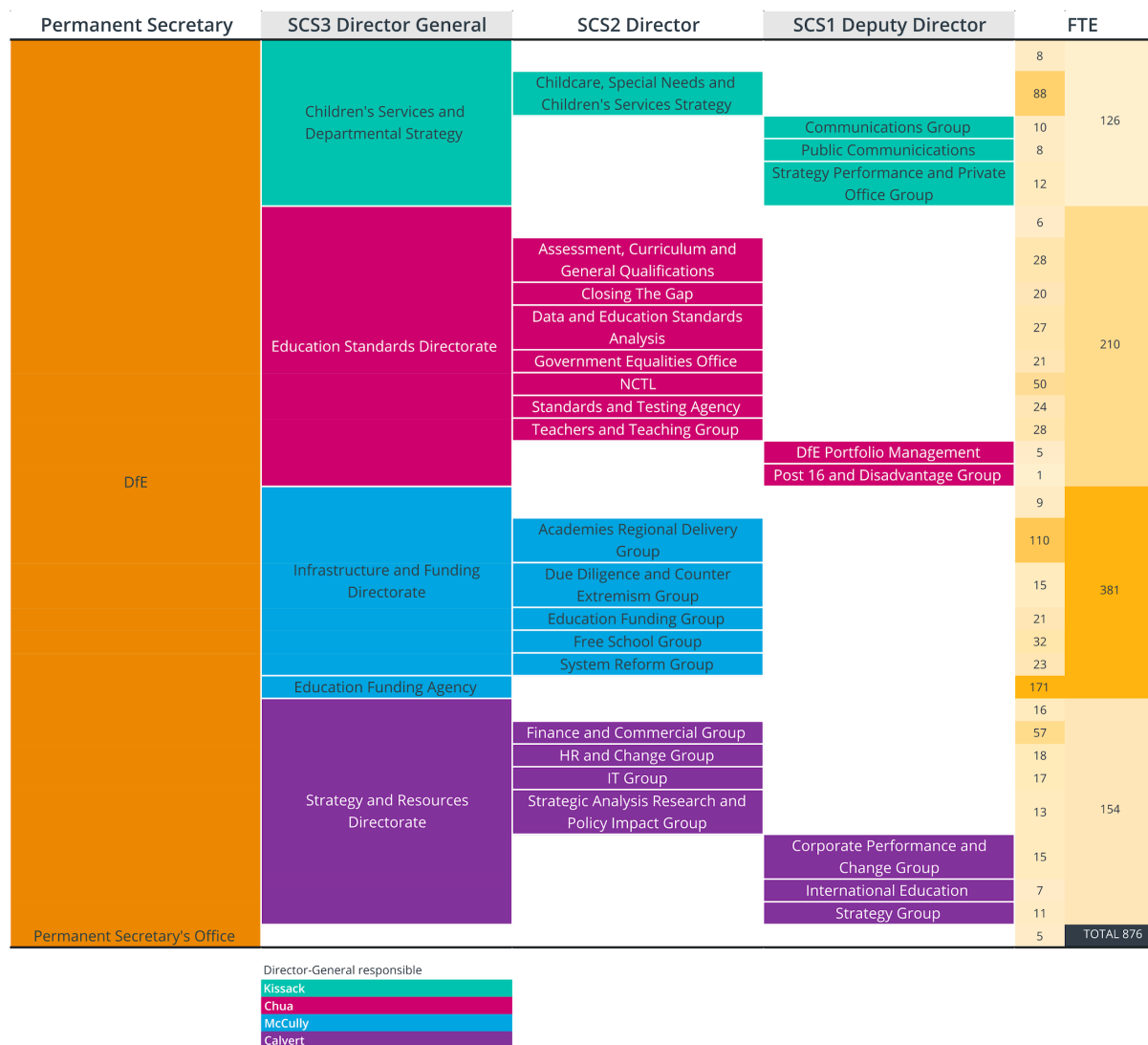
Also: code developed for the day (with comments), which focuses on structuring the data as a tree rather than a data frame: <https://github.com/mammykins/organogram>

Institute for Government

We took one department - DfE - and played around with visualising the structure and units of the departments. After some very basic charting (of FTE by unit), we looked at it through a Director General prism - looking at which directorates and, via the top level of line-management chains, which units, they were ultimately responsible for.

<http://bit.ly/2vm8Noc>

#IfGhackday: DfE units, by responsible DG, level of senior manager and employment (FTE), September 2016



Source: Institute for Government analysis of Department for Education, organogram published on data.gov.uk, September 2016.

Paybands for Junior Staff

We plotted the different names for grades and mapped them across the range of departments to produce a handy table showing what grades were equivalent to which:

[Version below published as part of [Institute for Government explainer on civil service grades](#)]

Departmental grade structures												
Rank	Short name	Full name	Other depts	CO	DCMS	DfID	DWP	FCO	HMRC	HMT	MoD	MoJ
Most senior	SCS	Senior Civil Service	SCS4	Permanent Secretary								
			SCS3	Second Permanent Secretary, Director General								
			SCS2	Director								
			SCS1	Deputy Director								
	G6&7	Grades 6 and 7	G6	Band A	AU	A1	G	D7	G6	E2	B1	A
			G7	Band A	A	A2	F	D6	G7	E	B2	A
	SEO/ HEO	Senior Executive Officer, Higher Executive Officer	SEO	B2	B	A2L	E	C5	SO	D	C1	B
			HEO	B2	C	B1	D	C4	HO	D	C2	C
	EO	Executive Officer	EO	B1	D	B2	C	B3	Officer	C	D	D
Most junior	AO/ AA	Administrative Officer, Administrative Assistant	AO	C2		C1	B	A2	AO	B	E1	E
			AA	C1			A	A1			E2	F

Source: Work emerging from Institute for Government hackday using departmental organograms, July 2017.

The pay band minimum and maximums on the far left were calculated by comparing all the salary ranges across departments and identifying the lowest minimum and the highest maximum. This is slightly skewed where departments have combined bands.

We took the data from the end of 2015 so the salary rates should have increased by 1% across the board and there was no data for DIT or BEIS and the column marked BEIS covered BIS. We then tried to produce some colourful graphs showing the ranges between grades across departments but ran out of time.

National Audit Office

We decided to focus our efforts on improving the current organogram explorer tool available on data.gov.uk which is clunky, slow and doesn't allow easy exploration of management structures within departments, given that you can only use it to drill down 1 person at a time and cannot search within it.

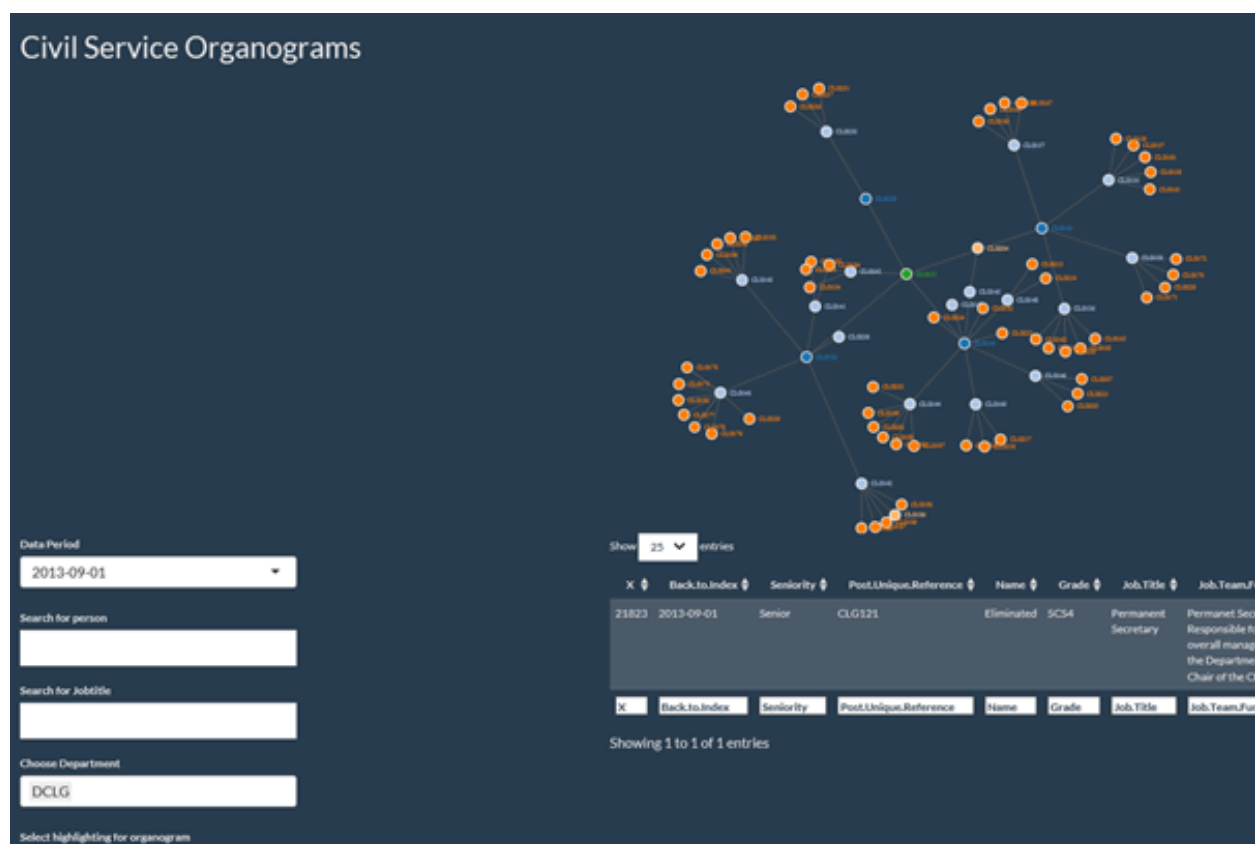
We therefore set out to develop an alternative, and hopefully much improved, organogram explorer tool. We used R for the data crunching and Shiny to create our output (Shiny is a great package that enables you to create powerful interactive web applications, straight from R, without having to use other programming languages such as javascript or HTML).

In particular, we decided to use an interactive network diagram to visualise the senior management structures within each department, as we had prior experience of using this to

show things like connections between data sets and breakdowns of departmental spending, so knew it would be effective.

A screenshot from our output (for DCLG) is shown below and, when accessed, the app allows users to explore the management structure within each department interactively by comparing departments side by side and over time while also identifying staff by grade, profession or job title. You can also search for individual senior civil servants by name to see where they sit within the structures, or just click on a dot to find out more details about the person filling that post.

Since the hack day, we've also posted the output up to the web at https://ben-coleman-dev.shinyapps.io/civil_service_organograms/ so anyone can have a play with it and see how it works.



The app obviously still needs a bit of work, and the data is a bit patchy in places, but we were pleased to get it done in the limited time available and think the results are quite impactful and informative. Since the hack, we've blogged about the event internally and are looking to develop our app further as we think there may be some real operational use for it within the business (eg. identifying who we might interview when conducting a VfM study on a certain project or looking at civil service capability more generally).

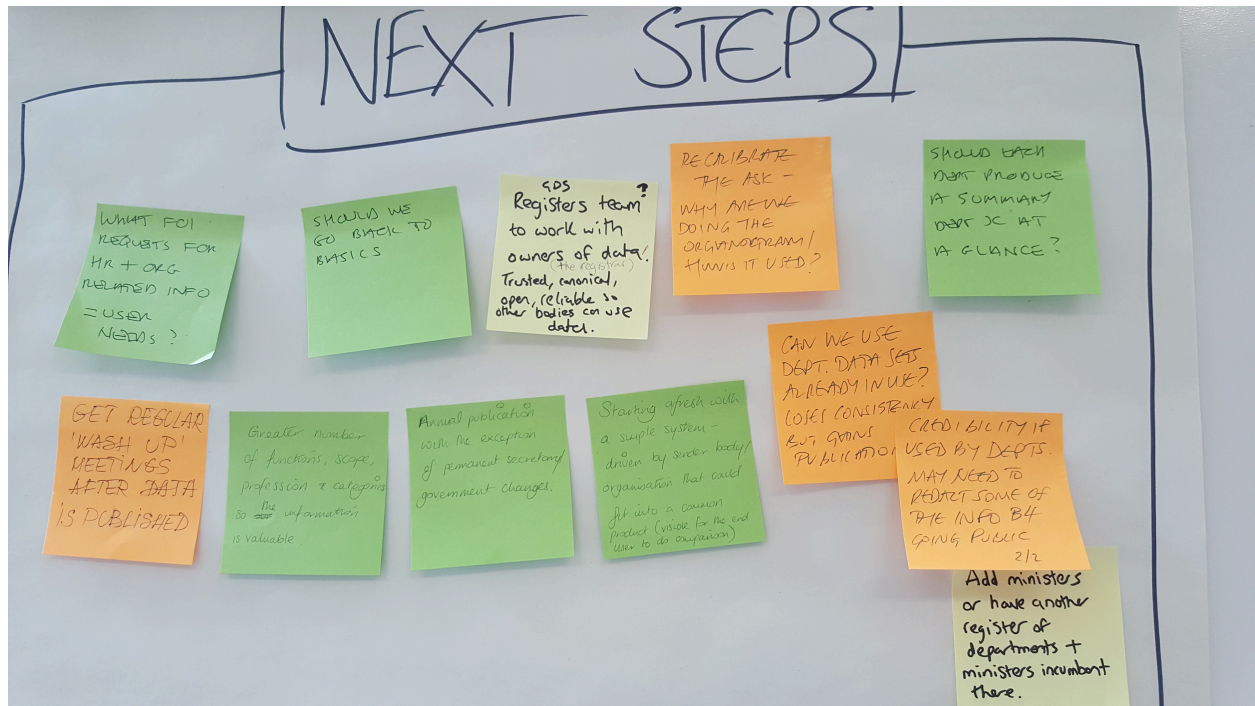
Flipboards

We had four flipboards throughout the day - please add any further comments:

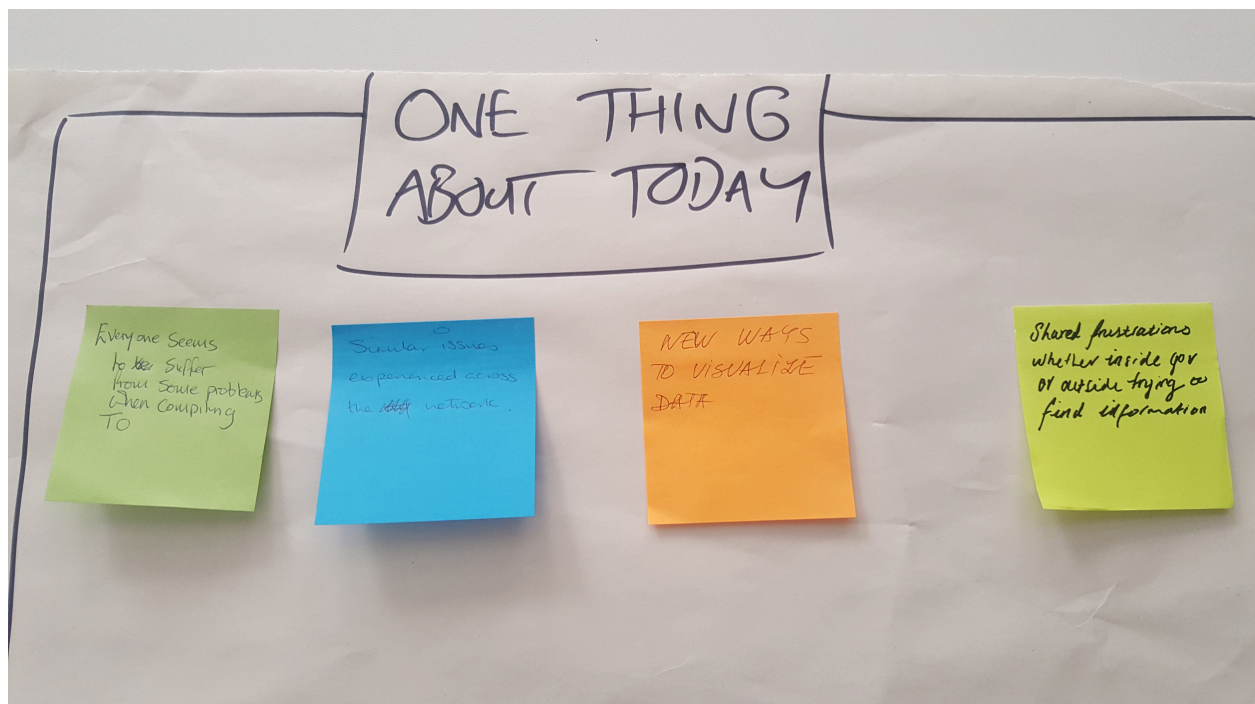
THE DATA

- DATA TOO ONE-SIDED TO PRODUCE
- CONTACT'S LIST FOR VERIFICATION OF STAFF ON SECOND-MENI
- MAP OF GRAMES IN DEPT. MANAGED TO CME OFFICE GRAMES
- Link to Hospitality data.
- PEOPLE DON'T FIT INTO CATEGORIES eg. Secondaries / Relinquishes
- Start date
- Exit data
- End date of role (with dates and end date)
- Eds across 2 departments not displayed. (eg. in DH + DUP. on secondment?)
- Contact details (Clots give generic email address)
- Absence of ministers.
- VARIOUS RELEASES MISSING
- Link to Hospitality data.
- REGISTER OF DEPT'L UNITS
- Do we need a "REGISTER" OF SENIOR CIVIL SERVANTS? (MISS TO LINK TO REGISTER)
- NOT CLEAR WHETHER CHANGE OVER OR JOB SHARE
- Department's own organisation charts (useful for establishing structure eg. DELG, DFT)
- Inconsistent name formatting: SURNAME, FIRST FIRST SURNAME SURNAME, TITLE FIRST NAME STAFF NUMBER WOULD BE BETTER
- Lack of reporting lines matters, given clarity of accountability
- Unlikely for Job team questions to list as "Undefined"
- ALION/PAY BAND ACROSS PUBLIC SECTOR (NDAB)
- IMPROVE VISUALISATION OF SHARED POSTS WHEN HAND-OVER
- No data from some departments provided
- No data on response letters of officials for projects eg. who is the relevant person to contact for project
- No clear SRO or accountability for Organisation + accuracy of info for monthly
- Need to have Real Guidance & commitment from CO on Professions & for CO to agree with ASGC
- Inconsistent grade data for junior grades - add CS equivalent column
- Often Database are FINANCIAL not Human Resources based
- Need to have MAJOR PROGRESS I.P.A. WHO IS SRO ORGANISMS
- SHOULD ALLOW DEPT. TO TALK TO RIGHT PERSON IN ANOTHER DEPT.
- LACK OF TRANSPARENCY FOR SC51 → NAMES → EMAILS
- Need to LINK MAJOR PROGRESS I.P.A. WHO IS SRO ORGANISMS
- SHOULD ALLOW DEPT. TO TALK TO RIGHT PERSON IN ANOTHER DEPT.
- LACK OF TRANSPARENCY FOR SC51 → NAMES → EMAILS
- Department's own organisation charts (useful for establishing structure eg. DELG, DFT)
- Inconsistent name formatting: SURNAME, FIRST FIRST SURNAME SURNAME, TITLE FIRST NAME STAFF NUMBER WOULD BE BETTER
- Lack of reporting lines matters, given clarity of accountability
- Unlikely for Job team questions to list as "Undefined"
- ALION/PAY BAND ACROSS PUBLIC SECTOR (NDAB)
- IMPROVE VISUALISATION OF SHARED POSTS WHEN HAND-OVER
- No data from some departments provided
- No data on response letters of officials for projects eg. who is the relevant person to contact for project
- No clear SRO or accountability for Organisation + accuracy of info for monthly
- Need to have Real Guidance & commitment from CO on Professions & for CO to agree with ASGC

What the next steps could be



One takeaway from the day



What people thought worked well (yellow) and thought could be better (pink) about the day

