HOW PORTABLE IS PORTABLE?

The GDPR's right to data portability in practice

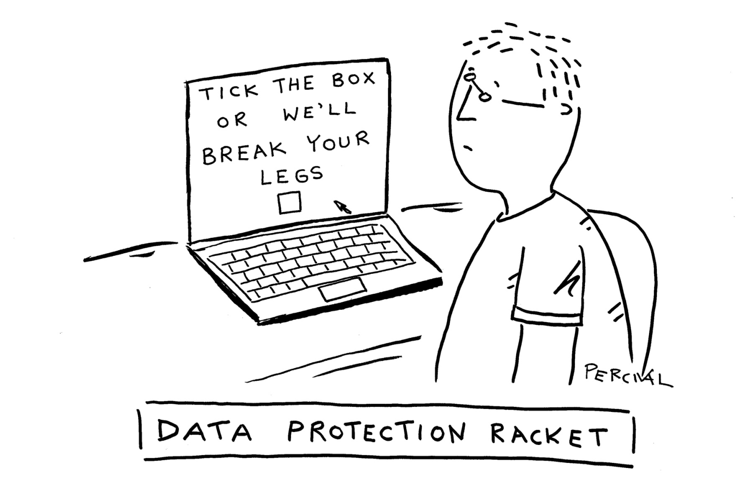
Janis Wong and Tristan Henderson

School of Computer Science
University of St Andrews
jccw@st-andrews.ac.uk

University of
St Andrews

FOUNDED
1413

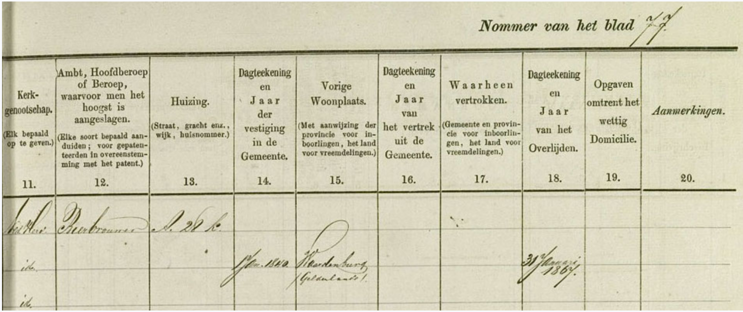
GDPR?



https://www.thephoenix.ie/2017/07/cartoon-3515-10/

DATA PROTECTION

- Dutch population registers from 1849
  - Name, Birth, Religion, Marital Status…
- Wide-scale misuse in WWII
- Do we have rights over our personal data?
  - Balancing our rights with potential misuse



Nommer van het blad 77

Kerkgenootschap. (Elk bepaald op te geven.)	Ambt, Hoofdberoep of Beroep, waarvoor men het hoogst is aangeslagen. (Elke soort bepaald aanduiden; voor gepatenteerden in overeenstemming met het patent.)	Huizing. (Straat, gracht enz., wijk, huisnommer.)	Dagteekening en Jaar der vestiging in de Gemeente.	Vorige Woonplaats. (Met aanwijzing der provincie voor inboorlingen, het land voor vreemdelingen.)	Dagteekening en Jaar van het vertrek uit de Gemeente.	Waarheen vertrokken. (Gemeente en provincie voor inboorlingen, het land voor vreemdelingen.)	Dagteekening en Jaar van het Overlijden.	Opgaven omtrent het wettig Domicilie.	*Aanmerkingen.*
11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
Ned. Herv.	Bierbrouwer	A. 28 k.							
id.			Jan. 1840	Waardenburg (Gelderland)			31 Januarij 1867		
id.									

https://www.dutchgenealogy.nl/population-registers/

GDPR

- New data protection regulation into force from May 2018
- Strengthens existing data subject rights (e.g. access, erasure, rectification)
- Introduces one new right: the right to data portability
- Data subject rights now free to exercise
- Aims to protect the processing of personal data of natural persons under a “technologically neutral” framework

DATA PROTECTION TERMS

- **Personal data:** "any information concerning an identified or identifiable natural person" (GDPR Recital 26)
- **Data subject:** those about whom personal data are collected and are about
  - The GDPR strengthens and introduces data subject rights
- **Data controller:** those who collect or determine what these data are used for
  - The GDPR enforces legal obligations on data controllers
- **Data processor:** those who process data on behalf of the data controller
- Defined by the ICO as the UK Data Protection Authority
  - https://ico.org.uk/

DATA SUBJECT RIGHTS UNDER THE GDPR

- Right to be informed (Article 12)
- Right of access (Articles 12, 15)
- Right to rectification (Article 16)
- Right to erasure ‘right to be forgotten’ (Article 17)
- Right to restriction of processing (Article 18)
- Right to data portability (Article 20)
- Right to object (Article 21)
- Rights related to automated decision-making (Article 22)

DATA SUBJECT RIGHTS UNDER THE GDPR

- Right to be informed (Article 12)
- Right of access (Articles 12, 15)
- Right to rectification (Article 16)
- Right to erasure ‘right to be forgotten’ (Article 17)
- Right to restriction of processing (Article 18)
- Right to data portability (Article 20) (NEW!)
- Right to object (Article 21)
- Rights related to automated decision-making (Article 22)

RESEARCH QUESTION

- Now that the GDPR is here:
  - How do data controllers approach Article 20 (RtDP)?
  - Do they know how to meet their RtDP obligations?
- How do we answer this?
  - Make some requests and look at the responses

WHAT IS ARTICLE 20?

Right to data portability

1. The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided, where:

   1.1 the processing is based on consent pursuant to point (a) of Article 6(1) or point (a) of Article 9(2) or on a contract pursuant to point (b) of Article 6(1); and

   1.2 the processing is carried out by automated means.

2. In exercising his or her right to data portability pursuant to paragraph 1, the data subject shall have the right to have the personal data transmitted directly from one controller to another, where technically feasible.

HOW DO DATA CONTROLLERS KNOW HOW TO COMPLY?

- The ICO issued the Guide to the GDPR[1], including:
  - Definitions to key terms
  - Suggested data formats
  - Data controllers' obligations to data subjects
- The A29WP (now EDPB) published Guidelines on the right to data portability[2], clarifying:
  - What data should be included in responses to RtDP requests
  - How portable data must be provided
  - Interoperability and interoperable systems only as desired outcomes

[1] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/

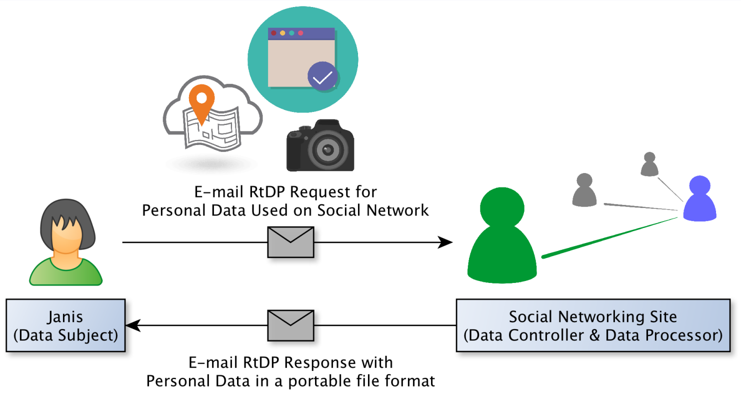
[2] ec.europa.eu/newsroom/document.cfm?doc_id=44099

METHOD

- Created a Python program to automatically generate and send RtDP requests
- Used e-mail since this is the only universally supported mechanism for making requests
- Sent 230 requests to a variety of data controllers
- Study ran from 25 May 2018 to 26 August 2018

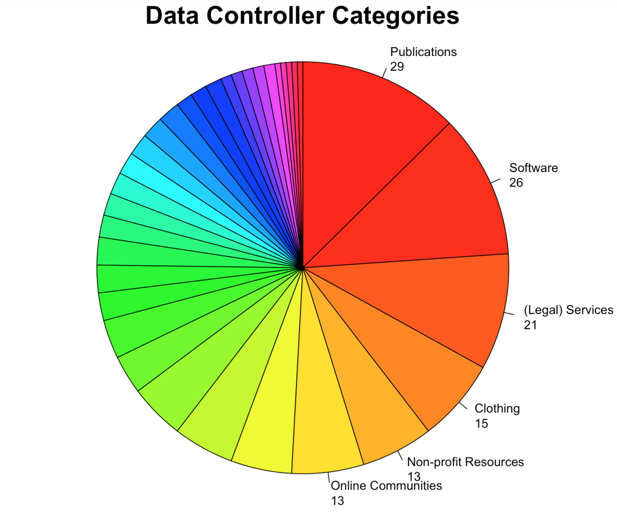
EXERCISING THE RTDP

The RtDP process simplified


E-mail RtDP Request for
Personal Data Used on Social Network
Janis
(Data Subject)
Social Networking Site
(Data Controller & Data Processor)
E-mail RtDP Response with
Personal Data in a portable file format

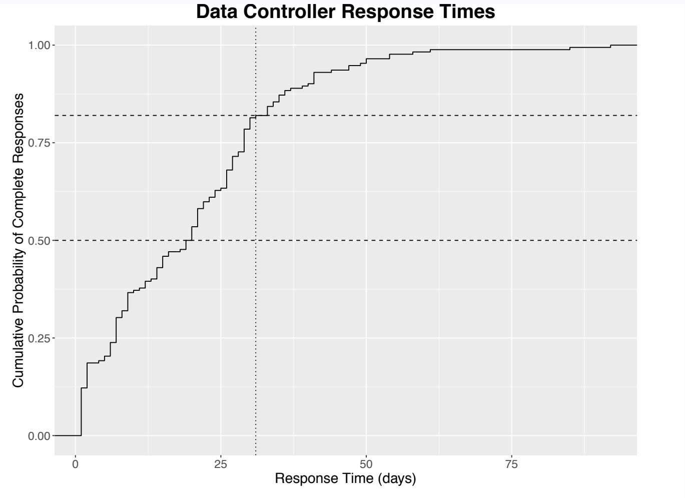
WHAT KIND OF DATA CONTROLLERS?

A wide variety of data controllers — not just SNSes!



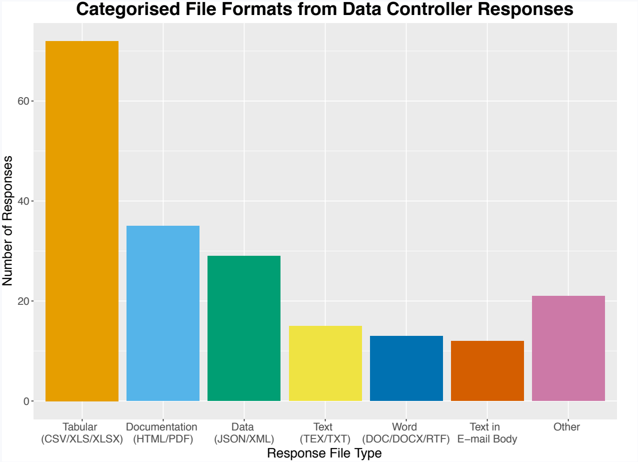
HOW LONG DID CONTROLLERS TAKE TO RESPOND?

172 of 230 data controllers responded, with 85% responding within a month



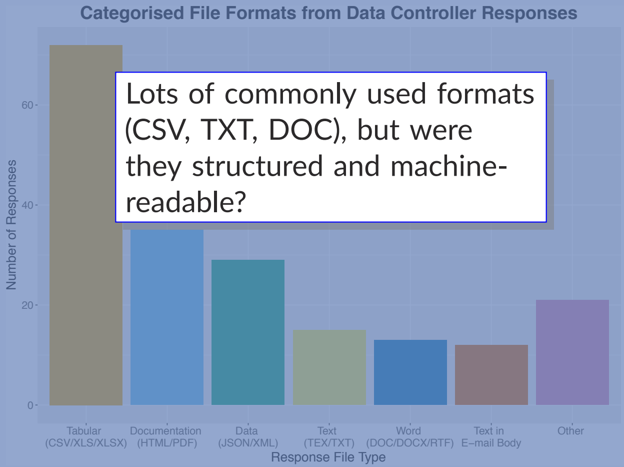
WHAT FILE FORMATS DID CONTROLLERS USE?

The wide range of response formats represent different data types, with no consensus on what is most appropriate



WHAT FILE FORMATS DID CONTROLLERS USE?

The wide range of response formats represent different data types, with no consensus on what is most appropriate



Categorised File Formats from Data Controller Responses

Lots of commonly used formats (CSV, TXT, DOC), but were they structured and machine-readable?

WERE FILE FORMATS COMPLIANT?

File Format	Structured?	Commonly Used?	Machine-Readable?
Email body	✗	✓	✗
CSV	✓	✓	✓
DOC/DOCX	✗	✓	✗
EML	✓	✓	✓
HTML	?	✓	?
ICS	✓	✓	✓
JPEG	✗	✓	✗
JSON	✓	✓	✓
KMZ	✓	✗	✓
MBOX	✓	✓	✓
MP4	✗	✓	✗
PDF	?	✓	?
PNG	✗	✓	✗
RTF	✗	✓	✗
TEX	✓	✓	✓
TXT	?	✓	?
VCS	✓	✓	✓
WAV	✗	✓	✗
XLS/XLSX	✓	✓	?
XML	✓	✓	✓

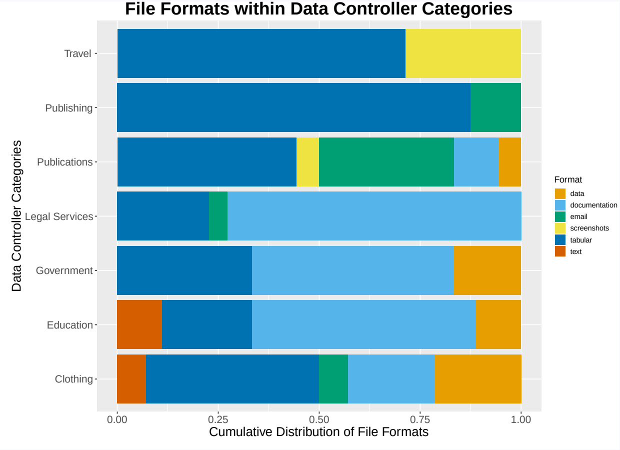
WERE FILE FORMATS COMPLIANT?

File Format	Structured?	Commonly Used?	Machine-Readable?
Email body	✗	✓	✗
CSV	✓	✓	✓
DOC/DOCX	✗	✓	✗
EML	✓	✓	✓
HTML	?	✓	?
ICS	✓		
JPEG	✗		
JSON	✓		
KMZ	✓		
MBOX	✓		
MP4	✗		
PDF	?	✓	?
PNG	✗	✓	✗
RTF	✗	✓	✗
TEX	✓	✓	✓
TXT	?	✓	?
VCS	✓	✓	✓
WAV	✗	✓	✗
XLS/XLSX	✓	✓	?
XML	✓	✓	✓

Most formats either comply (✓) or do not comply (✗) with the ICO definitions but some are ambiguous (?).

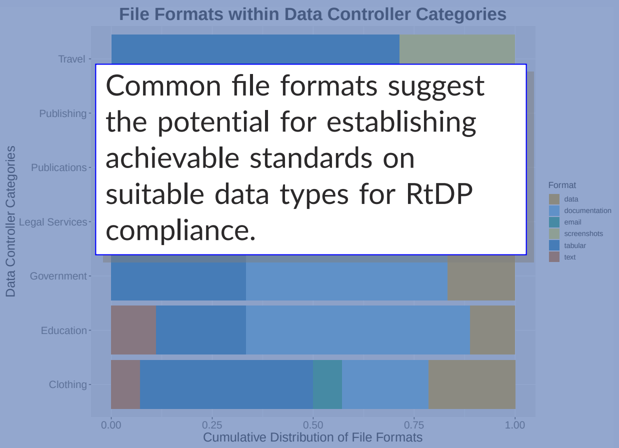
FILE FORMAT PATTERNS FOR CONTROLLER CATEGORIES?

Some data controller categories shown have at least one file format, with over 0.4 distribution represented in the responses



FILE FORMAT PATTERNS FOR CONTROLLER CATEGORIES?

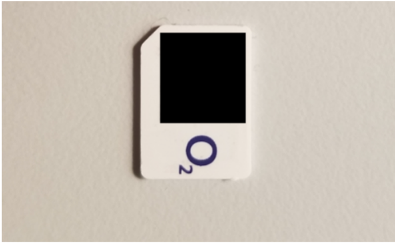
Some data controller categories shown have at least one file format, with over 0.4 distribution represented in the responses



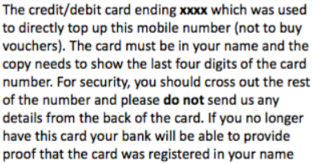
Common file formats suggest the potential for establishing achievable standards on suitable data types for RtDP compliance.

OTHER PROBLEMS

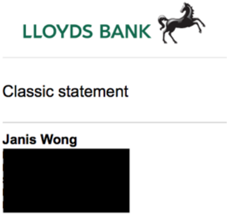
Additional personal data required for identifying data subject: Photographic national ID, proof of address, recent bank transactions… – new privacy risks?



[redacted]



The credit/debit card ending **xxxx** which was used to directly top up this mobile number (not to buy vouchers). The card must be in your name and the copy needs to show the last four digits of the card number. For security, you should cross out the rest of the number and please **do not** send us any details from the back of the card. If you no longer have this card your bank will be able to provide proof that the card was registered in your name

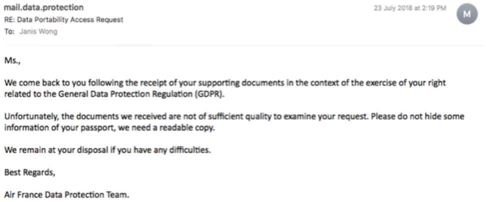


LLOYDS BANK

Classic statement

Janis Wong

[redacted]



mail.data.protection
RE: Data Portability Access Request
To: Janis Wong

23 July 2018 at 2:19 PM

Ms.,

We come back to you following the receipt of your supporting documents in the context of the exercise of your right related to the General Data Protection Regulation (GDPR).

Unfortunately, the documents we received are not of sufficient quality to examine your request. Please do not hide some information of your passport, we need a readable copy.

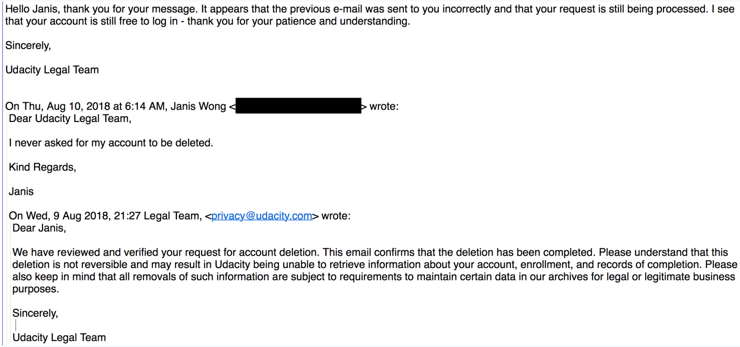
We remain at your disposal if you have any difficulties.

Best Regards,

Air France Data Protection Team.

OTHER PROBLEMS

Confusion between different rights: Some data controllers thought we were making Article 15 (access) or Article 17 (erasure) requests



Hello Janis, thank you for your message. It appears that the previous e-mail was sent to you incorrectly and that your request is still being processed. I see that your account is still free to log in - thank you for your patience and understanding.

Sincerely,

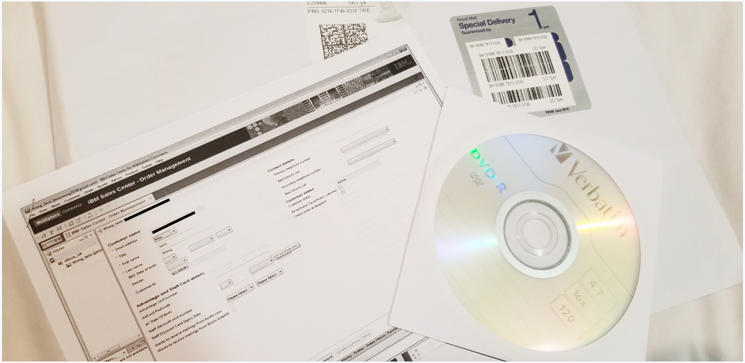
Udacity Legal Team

On Thu, Aug 10, 2018 at 6:14 AM, Janis Wong <[redacted]> wrote:

> Dear Udacity Legal Team,
>
> I never asked for my account to be deleted.
>
> Kind Regards,
>
> Janis
>
> On Wed, 9 Aug 2018, 21:27 Legal Team, <privacy@udacity.com> wrote:
>
> > Dear Janis,
> >
> > We have reviewed and verified your request for account deletion. This email confirms that the deletion has been completed. Please understand that this deletion is not reversible and may result in Udacity being unable to retrieve information about your account, enrollment, and records of completion. Please also keep in mind that all removals of such information are subject to requirements to maintain certain data in our archives for legal or legitimate business purposes.
> >
> > Sincerely,
> >
> > Udacity Legal Team

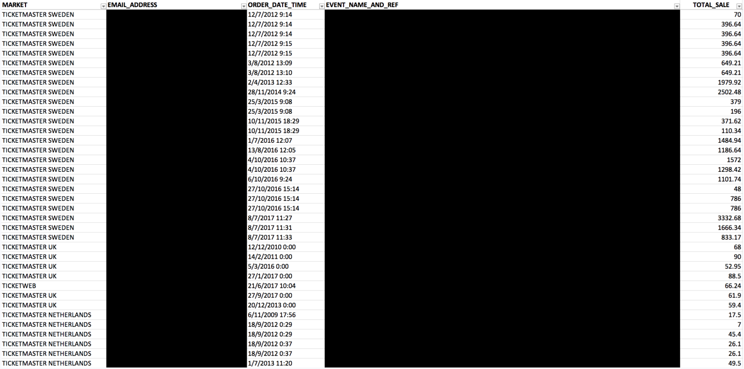
OTHER PROBLEMS

Security approaches varied: Data sent by e-mail (some of which were lost), passwords by post, CD-ROMs by post, download portals…



OTHER PROBLEMS

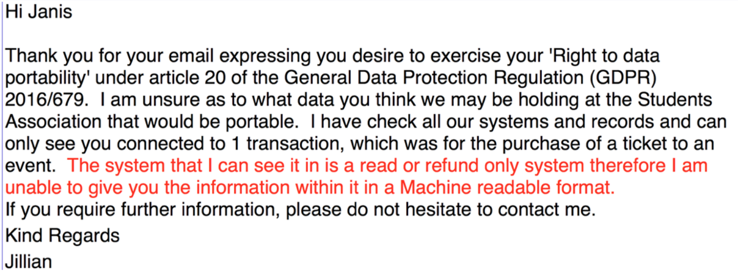
Data breach! One controller sent back data belonging to other requesters



MARKET	EMAIL_ADDRESS	ORDER_DATE_TIME	EVENT_NAME_AND_REF	TOTAL_SALE
TICKETMASTER SWEDEN	[redacted]	12/7/2012 9:14	[redacted]	70
TICKETMASTER SWEDEN	[redacted]	12/7/2012 9:14	[redacted]	396.64
TICKETMASTER SWEDEN	[redacted]	12/7/2012 9:14	[redacted]	396.64
TICKETMASTER SWEDEN	[redacted]	12/7/2012 9:15	[redacted]	396.64
TICKETMASTER SWEDEN	[redacted]	12/7/2012 9:15	[redacted]	396.64
TICKETMASTER SWEDEN	[redacted]	3/8/2012 13:09	[redacted]	649.21
TICKETMASTER SWEDEN	[redacted]	3/8/2012 13:10	[redacted]	649.21
TICKETMASTER SWEDEN	[redacted]	2/4/2013 12:33	[redacted]	1979.92
TICKETMASTER SWEDEN	[redacted]	28/11/2014 9:24	[redacted]	2502.48
TICKETMASTER SWEDEN	[redacted]	25/3/2015 9:08	[redacted]	379
TICKETMASTER SWEDEN	[redacted]	25/3/2015 9:08	[redacted]	196
TICKETMASTER SWEDEN	[redacted]	10/11/2015 18:29	[redacted]	371.62
TICKETMASTER SWEDEN	[redacted]	10/11/2015 18:29	[redacted]	110.34
TICKETMASTER SWEDEN	[redacted]	1/7/2016 12:07	[redacted]	1484.94
TICKETMASTER SWEDEN	[redacted]	13/8/2016 12:05	[redacted]	1186.64
TICKETMASTER SWEDEN	[redacted]	4/10/2016 10:37	[redacted]	1572
TICKETMASTER SWEDEN	[redacted]	4/10/2016 10:37	[redacted]	1298.42
TICKETMASTER SWEDEN	[redacted]	6/10/2016 9:24	[redacted]	1101.74
TICKETMASTER SWEDEN	[redacted]	27/10/2016 15:14	[redacted]	48
TICKETMASTER SWEDEN	[redacted]	27/10/2016 15:14	[redacted]	786
TICKETMASTER SWEDEN	[redacted]	27/10/2016 15:14	[redacted]	786
TICKETMASTER SWEDEN	[redacted]	8/7/2017 11:27	[redacted]	3332.68
TICKETMASTER SWEDEN	[redacted]	8/7/2017 11:31	[redacted]	1666.34
TICKETMASTER SWEDEN	[redacted]	8/7/2017 11:33	[redacted]	833.17
TICKETMASTER UK	[redacted]	12/12/2010 0:00	[redacted]	68
TICKETMASTER UK	[redacted]	14/2/2011 0:00	[redacted]	90
TICKETMASTER UK	[redacted]	5/3/2016 0:00	[redacted]	52.95
TICKETMASTER UK	[redacted]	27/1/2017 0:00	[redacted]	88.5
TICKETWEB	[redacted]	21/6/2017 10:04	[redacted]	66.24
TICKETMASTER UK	[redacted]	27/9/2017 0:00	[redacted]	61.9
TICKETMASTER UK	[redacted]	20/12/2013 0:00	[redacted]	59.4
TICKETMASTER NETHERLANDS	[redacted]	6/11/2009 17:56	[redacted]	17.5
TICKETMASTER NETHERLANDS	[redacted]	18/9/2012 0:29	[redacted]	7
TICKETMASTER NETHERLANDS	[redacted]	18/9/2012 0:29	[redacted]	45.4
TICKETMASTER NETHERLANDS	[redacted]	18/9/2012 0:37	[redacted]	26.1
TICKETMASTER NETHERLANDS	[redacted]	18/9/2012 0:37	[redacted]	26.1
TICKETMASTER NETHERLANDS	[redacted]	1/7/2013 11:20	[redacted]	49.5

OTHER PROBLEMS

Machine-readable? One controller claimed a machine-readable response was not possible; two sent paper scans as allegedly machine-readable



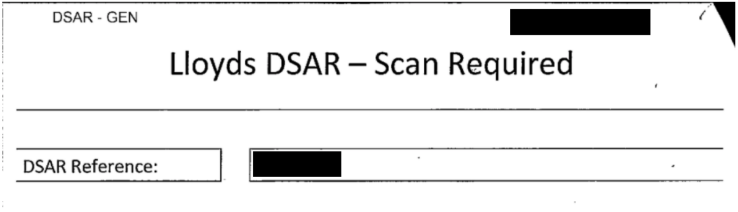
Hi Janis

Thank you for your email expressing you desire to exercise your 'Right to data portability' under article 20 of the General Data Protection Regulation (GDPR) 2016/679. I am unsure as to what data you think we may be holding at the Students Association that would be portable. I have check all our systems and records and can only see you connected to 1 transaction, which was for the purchase of a ticket to an event. The system that I can see it in is a read or refund only system therefore I am unable to give you the information within it in a Machine readable format.

If you require further information, please do not hesitate to contact me.

Kind Regards

Jillian



DSAR - GEN [redacted]

Lloyds DSAR – Scan Required

DSAR Reference: [redacted]

WHAT NEXT?

- Looking at content of responses to determine compliance
- We need better definitions of structured, commonly used, and machine-readable
- Looking at the right to transmit (perhaps too early given lack of preparedness for even the easier bits of Art 20)
- Is interoperability feasible?
- How can technology help?

CONCLUSIONS

- GDPR makes it easier to conduct empirical data protection studies
- Data controllers vary in their approach to RtDP: file formats, security, interpretation of the requirements
- Better guidance needed
- More empirical work is needed

🏠	tinyurl.com/janiswong	tnhh.org
✉	jccw@st-andrews.ac.uk	tnhh@st-andrews.ac.uk
🐦	@janiswong_	@tnhh