SUPPORT SMALL AND MEDIUM ENTERPRISES ON THE DATA PROTECTION REFORM II

Report on the statistics and efficiency of the hotline

Deliverable **D3.3** (Version 1.0)

Dr David Barnard-Wills, Dr Filippo Marchetti, Gábor Kulitsán & Renata Nagy
Budapest – Brussels – Waterford

July 2020

distribution level: **Public**

LSTS LAW, SCIENCE, TECHNOLOGY & SOCIETY STUDIES VRIJE UNIVERSITEIT BRUSSEL

TRILATERAL RESEARCH

Authors	
Name	Partner
Leanne Cochrane	TRI
David Barnard-Wills	TRI
Filippo Marchetti	TRI
Gábor Kulitsán	NAIH
Renata Nagy	NAIH

Internal Reviewers	
Name	Partner

Version Number	Author	Purpose/Change	Date
0.1	Barnard-Wills	Initial draft	

Institutional Members of the STAR Consortium		
Member	Role	Website
Nemzeti Adatvédelmi és Információszabadság Hatóság (NAIH)	Project Coordinator	naih.hu
Trilateral Research Ltd. (TRI)	Partner	trilateralresearch.com
Vrije Universiteit Brussel (VUB) Research Group on Law, Science, Technology and Society (LSTS)	Partner	https://lsts.research.vub.be/

This report has been prepared for the European Commission's Directorate-General for Justice and Consumers (DG JUST). The STAR II project (Support small and medium enterprises on the data protection reform II; 2018-2020) is co-funded by the European Union under the Rights, Equality and Citizenship Programme 2014-2020 (REC-RDAT-TRAI-AG-2017) under Grant Agreement No. 814775. The contents of this deliverable are the sole responsibility of the authors and can in no way be taken to reflect the views of the European Commission.

Permanent link: TBC

version
July 2020

Table of Contents

Background to the STAR II project

The STAR II project (Support small and medium enterprises on the data protection reform II) commenced in August 2018 and is intended to run for a two-year period. It is co-funded by the European Union under the Rights, Equality and Citizenship Programme 2014-2020 and is aimed at: (i) assisting European Union (EU) Data Protection Authorities (DPAs) raise awareness about the General Data Protection Regulation (GDPR) among small and medium enterprises (SMEs); and (ii) assisting SMEs to comply with the GDPR.

There are 22 million SMEs in the EU who form the core of the EU enterprise policy. These SMEs face distinctive challenges from data protection law and can often not afford professional legal advice. As such, they merit special support from public authorities as recognised by Recital 132 of the GDPR which specifies that when undertaking awareness-raising activities addressed to the public, data protection authorities should include specific measures directed towards, among others, SMEs.

The STAR II project outputs will include:

Reports on existing DPA approaches to engagement with SMEs and on SMEs experiences with the GDPR.

An email hotline run by the *Nemzeti Adatvédelmi és Információszabadság Hatóság* (NAIH) in both Hungarian and English;

A guidance document for DPAs on good practices in awareness-raising techniques among SMEs;

A handbook for SMEs to help them comply with the GDPR.

STAR projects, 2017-2019

The STAR II project follows on from the STAR project (Support training activities on the data protection reform), which is nearing completion and focused on providing support to the training activities of DPAs and data protection officers (DPOs) on the EU data protection reform, especially the GDPR. The STAR project was also co-funded by the EU under the Rights, Equality and Citizenship Programme 2014-2020.The outputs from the STAR project have included:

1) Training scenarios for each training category,
2) A Seminars' Topics List, based on the training scenarios,
3) Seminar Material for each one of the seminars,
4) Webinars (selected from the Seminars' Topics List),
5) A training Handbook,
6) A takeaway reference GDPR checklist,
7) A ten-point GDPR introductory list.

Executive summary

As part of the STAR II project, NAIH set up and operated a dedicated email-based "hotline" to receive and answer questions from small and medium enterprises (SMEs). The operation of the hotline is documented in a companion report (D3.2).

The purpose of this report (D3.3) is to conduct and report upon a statistical analysis of the hotline. Including identifying the most frequently asked questions, the number and nature of issues, frequency of contacts, response time, effects of awareness-raising campaigns and public appearances. The report first gives an overview of the data collected during the hotline operation. Then an analysis of the most commonly asked questions and topics, before conducting an operational analysis, primarily focused upon what can be learnt about the operation of the hotline based upon response times and the difficulty of various types of queries. The report concludes with implications for the STAR II project – largely than the data from the hotline supports the analysis of SME needs developed in the project's earlier reports, and with insight and recommendations for other supervisory authorities that may wish to replicate the email hotline or use the data examined here to support their own outreach activities.

The study primarily finds that operating an email hotline for SMEs can be quite time consuming, but that they are a good source of accurate intelligence on the issues with which SMEs are struggling with in a particular jurisdiction. Most of the effort and time involved comes from responding to concrete compliance questions addressed to very specific circumstances of individual SMEs, requiring the involvement of data protection experts. It suggests that a hotline should be operated whilst at the same time making the answers to the more simple and straightforward answers available in other formats.

List of Abbreviations

DPA	Data Protection Authority
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
EC	European Commission
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
ePrivacy Directive	Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector (ELI: data.europa.eu/eli/dir/2002/58/oj)
EU	European Union
GDPR	General Data Protection Regulation (Regulation EU 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, ELI: data.europa.eu/eli/reg/2016/679/oj
SME	Small and medium enterprise
STAR	Support training activities on the data protection reform
STAR II	Support small and medium enterprises on the data protection reform II
WP29	Working Party on the Protection of Individuals with regard to the Processing of Personal Data set up under Article 29 of Directive 95/46/EC (Article 29 Working Party). WP29 was replaced by the EDPB on 25 May 2018. The EDPB has endorsed many WP29 GDPR-related guidelines.

Introduction

As part of the STAR II project, *Nemzeti Adatvédelmi és Információszabadság Hatóság (*NAIH), the data protection supervisory authority of Hungary, set up and operated a dedicated email-based "hotline" to receive and answer questions from small and medium enterprises (SMEs). The operation of the hotline is documented in a companion report (D3.2). In summary, the hotline offered a dedicated email address that SMEs could use to send questions about data protection and particularly the GDPR to NAIH, and receive a response. This was in addition to other ways of contacting the authority. This hotline was new effort for NAIH, and in the context of the STAR II project we are interested in what can be learnt about the operation of this hotline, both for further SME engagement and support activity by NAIH, but also for other EU supervisory authorities.

The purpose of this report (D3.3) is therefore to conduct and report upon a statistical analysis of the hotline. This includes identifying the most frequently asked questions, the number and nature of issues, frequency of contacts, response time, effects of awareness-raising campaigns and public appearances, but also using this analysis to try and unpack some of the implications for hotline design and operation.

The report first gives an overview of the data collected during the hotline operation. Then an analysis of the most commonly asked questions and topics, before conducting an operational analysis, primarily focused upon what can be learnt about the operation of the hotline based upon response times and the difficulty of various types of queries.

Overview of the hotline data

In this section we provide an overview of the type of data collected during the operation of the hotline, and which serves as the basis for analysis in the subsequent sections.

Whilst the NAIH hotline was accessible in both English and Hungarian, and was potentially open to questions from anywhere in Europe (and was promoted as such by STAR II including through the SME associations we conducted research with in WP2, the hotline in practice only received enquiries from SMEs in Hungary.[1] This indicates the appetite expressed in the WP2 reports that SMEs primarily seek assurance and assistance from their national regulators, even given the harmonised ambition of the GDPR, and the potential for cross-border investigation and regulation offered by the one-stop-shop principle. As a result, the data here do not inherently generalise to SMEs across Europe, but do provide additional perspective on the operation of the hotline, and as shown below, combine with the more qualitative data from STAR II's surveys, interviews and workshops.

During the operation of the, NAIH collected in a self-report spreadsheet, data on:

- Date of receipt
- MS origin
- Language
- Nature of issues raised
- Issues raised
- Topic
- Difficulty of issue
- Keywords
- Level of difficulty of the answer
- Provisions of the GDPR concerned
- Date of reply
- Case worker
- Response status

We can provide some initial statistics on the use of the hotline for convenience of the reader, but for more detail on this can be found in the report on the operation of the hotline. Essentially, the hotline received 149 inquiries between March 2019 and March 2020 (both Marchs were half months).

Only eight emails were determined to be out of scope of the hotline. This is a relatively small proportion and suggests that the introductory or promotional material provided in relation to the hotline does a fairly good job out outlining what the hotline will cover.

[1] One email was jointly categorised as UK/HU.

Frequently asked questions

In this section of the report we focus upon the content of the questions asked of the NAIH hotline and what this can tell us about the data protection knowledge, needs, and concerns of SMEs in this context.

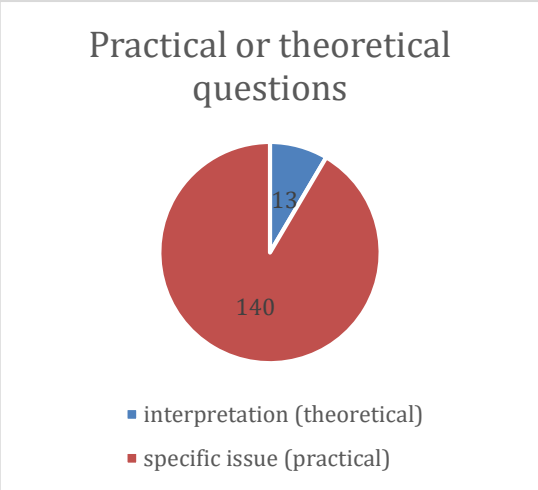

Figure 1 - Practical/theoretical questions

NAIH categorised each email as having a theoretical or practical orientation. The vast majority of emails received were practical questions, often relating to the concrete compliance requirements in a specific situation (see Figure 1). This supports our previous research suggesting that SMEs were seeking practical applied guidance, directly related to their activities.

NAIH categorised all email enquiries by the topic of the questions. This allows us to identify the number of questions for each topic, and therefore the identify the most common questions that were asked by SMEs, as shown in the following chart (Figure 2).

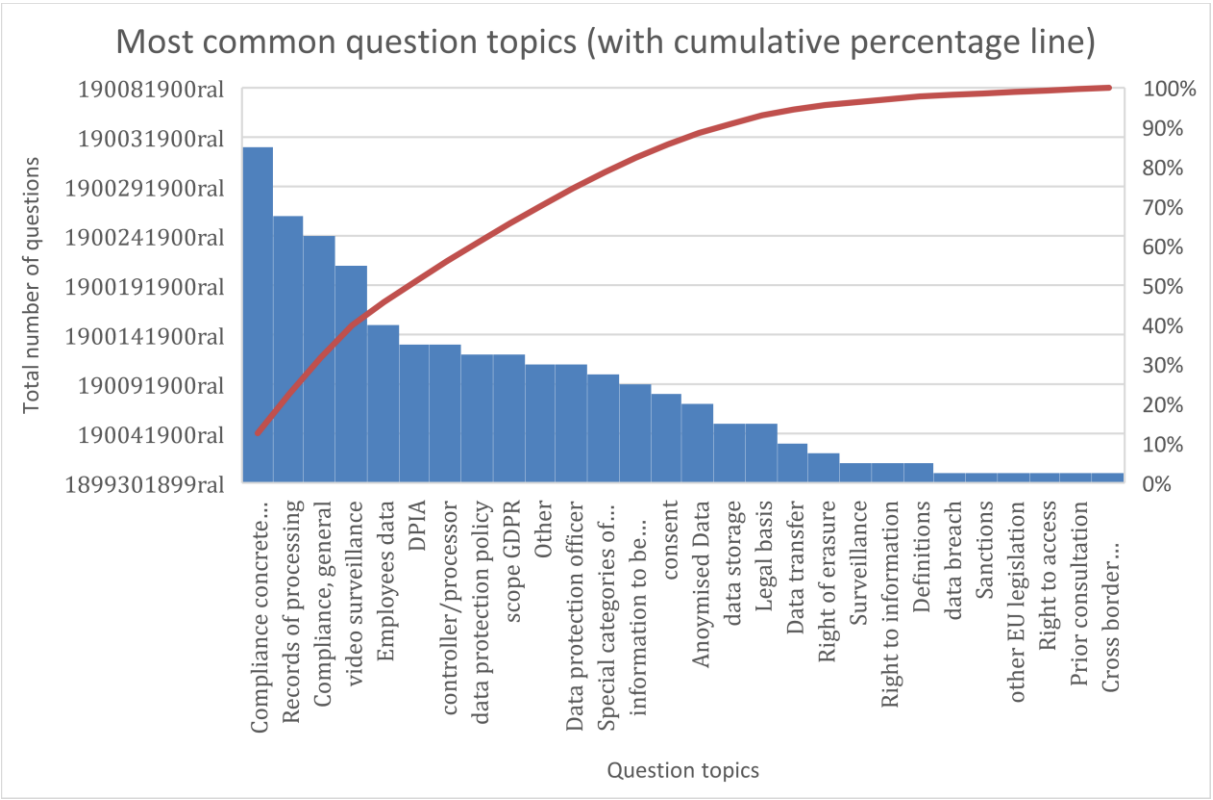

Figure 2 - Question topics

The line on the chart shows the cumulative percentage, so for example ten percent of questions were about concrete compliance issues and questions around records of processing, and 50% of questions are represented by the first five most common topics. The distribution of topics has quite a long tail, with a large number of topic categories with just one or two questions on that topic (even more so if the "other" category is broken down). The implication here is that a hotline benefits the most be ensuring that its knowledge base is strong in these areas, and that its staff are well trained on these areas. Video surveillance (the use and regulation around CCTV) remains a regular topic for SMEs.

It is worth considering to what extent there are missing topics, those that might be anticipated based upon either analytical, expert, or regulatory perspectives on the GDPR, but are not represented in the questions received.

- Certification and Codes of conduct – There were no questions about GDPR certification schemes. One question addressed the potential of ISO27001 certification as a way of demonstrating appropriate technical and organisational measures for data protection. This is an area that remains relatively low-profile.
- Other rights of the data subject – Whilst SMEs did ask questions about the rights to erasure and access, there were no questions received about the other rights of the data subject under the GDPR - the rights to portability, rectification, to object, in relation to automated decision making and profiling.

It is not possible to determine the reason for these absence from the data here, but it would be worth exploring if these issues are i) already understood by SMEs, ii) answered by information provided by NAIH or other sources, iii) SMEs are not aware of them or even thinking about them, iv) there is a lack of upstream pressure on SMEs to engage with these topics – for example, there may be little demand from data subjects to exercise these rights against SMEs.

NAIH reported the articles of the GDPR that were involved in drafting a response to an email question.[2] This allows us to examine the frequency with which these questions emerged, acting as a proxy for those areas of the GDPR which are creating the most questions for SMEs. This is visualised in Figure 3. It should be noted that these implications are attributed by the supervisory authority, and there this is affected by their own process of thinking, operational and legal philosophy and the way they have responded to the questions. All other articles should be understood has having received no questions on them to the hotline.

[2] Reported on a per-email, rather than a per question basis.

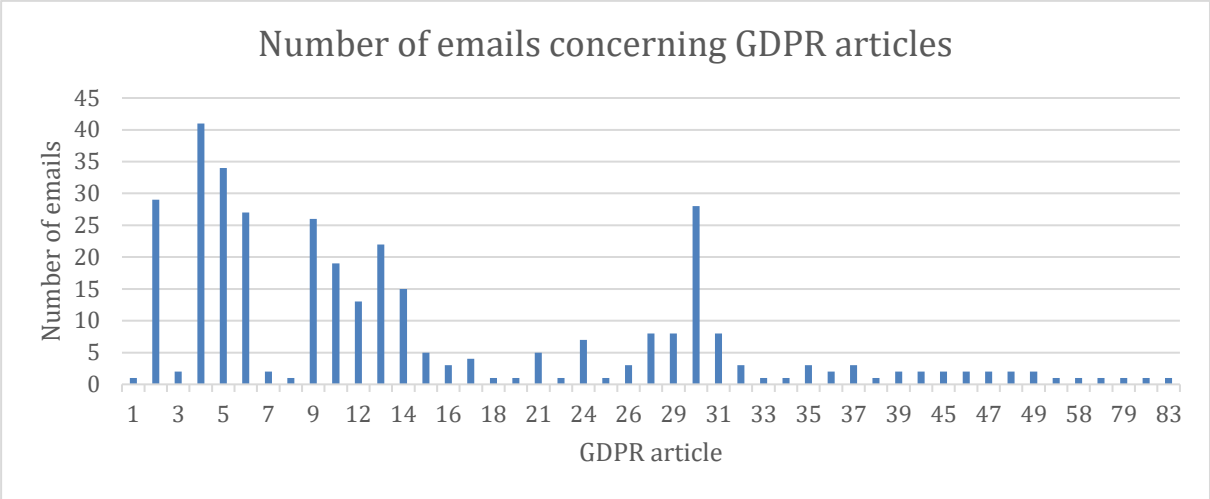

Figure 3 Frequency of GDPR articles involved in responding to emails

The most commonly addressed GDPR article is **Article 4 – Definitions**, most likely because this article is a significant touchstone for nearly all other areas of potential questions, e.g. "is this personal data or not?" Similarly, **Article 5 – Principles** is commonly referred to as a touchpoint for areas of more detailed or specific question, or to resolve questions that do not have a specific answer. **Article 2 – Scope** is similar, with many questions touching on the extent to which this question is affected by the GDPR or not. Fairly often SMEs wanted to know if a particular activity fell under the ambit of the GDPR or not, or counted as the processing of personal data.

Article 28 – Processor is also frequently implicated by the email questions. This article sets out the responsibilities of an entity processing personal data, in particular, the appropriate nature of the relationship between controller and processor. As such it becomes relevant to those SMEs trying to 1) determine if they are a data controller or processor, 2) understand their responsibilities as a processor, or 3) engage a processor as a data controller and understand what they can expert or demand from the processor. In the hotline emails it is also associated by NAIH with general or exploratory "how can my business comply with the GDPR?" questions. A smaller number of emails were relevant to other adjacent articles on this topic such as controller-processor relations, joint controllers, and controllers/processors not established in the Union. This latter issue was a significant one from the perspective of our Validation Workshop participants in WP2 – for them, many SMEs were wrestling with their relationships with overseas service provides (e.g. cloud computing or other data services provided by large international companies).

Article 6 – Lawfulness of processing sets out the legal bases for the processing of personal data. Questions on this topic included questions about the legal basis for the processing of personal data of employees, including the applicability of consent, as well as questions explicitly asking in an SME could process certain personal data for a particular given purpose. NAIH also referred to this article in questions about the collection, management or

documentation of consent. **Article 9 – processing of special categories of data**, was implicated mostly in questions about processing health data, and questions requesting further clarity on the relationship between Article 9 and Article 6, but also in questions related to video surveillance and the incidental or unintended collection of special category data.

Articles 12 - Transparent information, communication and modalities for the exercise of the rights of the data subject, **13 - Information to be provided where personal data are collected from the data subject**, and **14 - Information to be provided where personal data have not been obtained from the data subject**, were frequently implicated in the same email, most typically in questions relation to what needed to be included in a privacy policy or notice. Article 13 was the most commonly invoked of the three suggestion that most SMEs with questions for the regulator were primarily thinking about their data collection directly from the data subject (customer databases, marketing efforts and employee data).

Somewhat surprisingly, there were relatively few emails that related to the role of data protection officers (Articles 37, 38, 39), potentially related to the relatively small proportion of SMEs that appoint one. There were a smaller number of emails that involved any of the potential mechanisms for the international transfer of data (Articles 44-48) There were single emails that addressed some of the Chapter IX provisions relating to specific processing situations. There was a single email related to Article 25 - data protection by default

Comparison with topics and issues identified by the STAR II survey

In STAR II D2.2 we provided the analysis of a survey of SMEs, including a question asking the respondents what questions they would like answers on from DPAs.[3] The message from this survey was that whilst there was interest in guidance across the topics covered by the GDPR, SMEs were particularly interested in guidance on technical and organisational measures for data protection. The chart below (figure four) reproduces the results from that report.

The data from the email hotline offers us the opportunity to compare these self-reported interests in guidance with the actual questions that SMEs asked of the regulator in the real-world, potentially giving us a better sense of the validity of the survey research. Whist the hotline data is quite localised, the survey data was collected from across Europe, meaning that overlaps suggests that the hotline experience may generalise more easily. Whilst the topic categories do not correlate on a one-to-one basis there is substantial overlap.

[3]David Barnard-Wills, Leanne Cochrane, Kai Matturi & Filippo Marchetti, *Report on SME Experience of the GDPR*, STAR II Deliverable D2.2, July 2019, p.41.

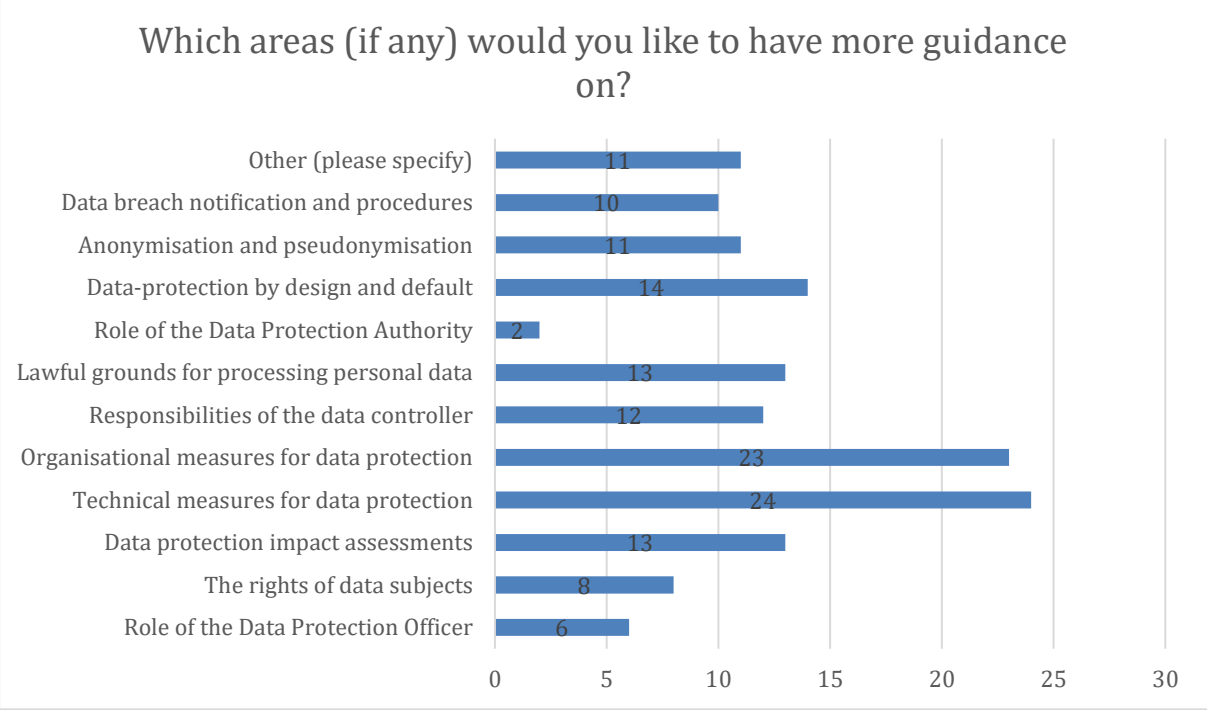

Figure 4 - Guidance requested by SMEs - STARII survey

Essentially, concrete compliance matters encompass organisational and technical measures for data protection. Whilst not all emails in this category were about technical and organisational measures, when the specific question was analysed, many of these were about technical and organisational measures. There is also some overlap between records of processing and the responsibilities of the data controller. Several topics occurred in similar proportion to what might be expected from the survey (for example, legal basis, controller-processor relations, and the role of the DPO).

Responsibilities for handling employee data was an area of demand that was not identified in the survey. The survey categories (excepting "other") are primarily conceptually generated from the GDPR, where there are no specific requirements for processing employee data (apart from acknowledging the potential differences in Member State law in Article 88). Handling employee data was an issue that arose in the WP2 validation workshop, where it was indicated, that apart from a customer contact database, employee data was the main form of personal data that many SMEs were involved in processing.[4]

One noticeable difference between the survey results and the SME hotline is that whilst additional guidance on data breach notification whilst a significant topic for survey respondents it was only the subject of a single enquiry to the hotline. This may signify that 1) SMEs responding in the NAIH context already know where to report a data breach, and/or 2)

[4] Leanne Cochrane, David Barnard-Wills, Kai Matturi & Filippo Marchetti, *Report on WP2 Validation Workshop*, STAR II Deliverable D2.3, July 2019,

SMES are not aware of their exposure to data breaches. We certainly anticipate that there is some reluctance to ask questions about data breaches to the regulator by identifiable email, but a greater willingness to express the desire for such information to an anonymous survey run by a third party.

Despite the apparent interest in data protection by design and default indicated in the survey, there were no questions to the hotline about this topic. There were questions were the answers from the supervisory authority could include elements of data protection by design, but this topic does not appear to have traction with SMEs in this context.

"Other" questions

The categories for topics were determined by NIAH at the start of the operation of the hotline. Therefore, the questions that were categorised under "other" represent questions that did not easily fit into these categories and were to a certain extent unanticipated. There were not many emails that could not be categorised in the anticipated categories. In brief, the questions categorised as other included questions about the cost of using the hotline, appropriate software for record keeping, and about obligations to retain employee timesheets.

Operational analytics

In this section of the report we attempt to use the operational data collected during the operation of the hotline to identify patterns, and understand some of the relationships between variables.

Number of contacts

March	April	May	June	July	August	September	October	November	December	January	February	March
15	22	16	15	13	6	6	11	11	11	10	8	5

Therefore, the average number of emails per month was 12.4. The most common (modal) number of emails was 11, as was the median number of emails. The highest number of emails received was 22 in April 2019, and the lowest was five in the March of 2020 (although this was a half month) or six otherwise in August and September 2019. The following graph shows the number of contacts to the hotline per month over time.

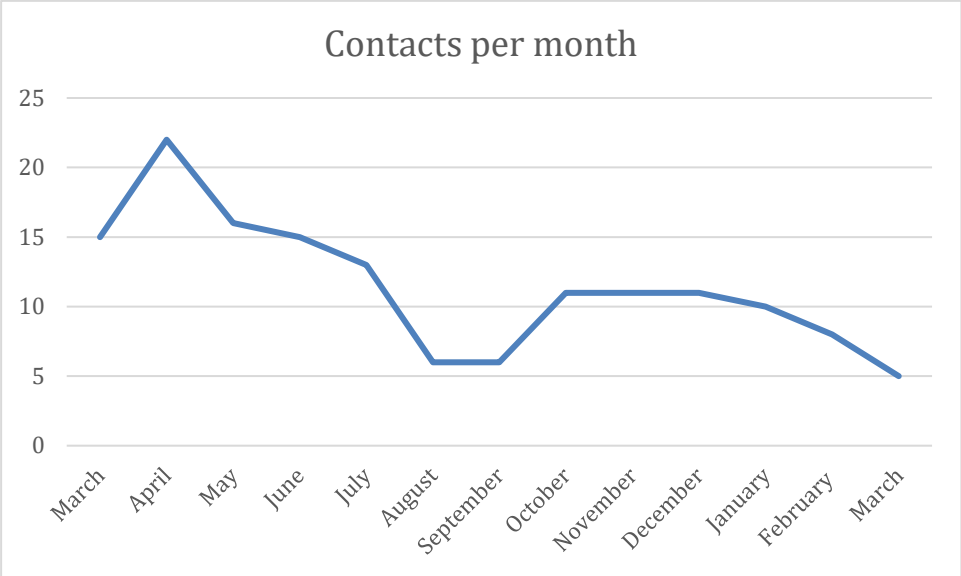

Figure 5 email contacts per month

Use of the service initially peaks after its launch and then there is a general downward trend after first peak. This may indicate that potential SME users who have a question they would like answered by NAIH essentially have these questions "in reserve", effectively unanswered, at the start of the hotline. Once this "reserve" of questions has been answered, the usage level reverts to what might be a more typical regular usage pattern, with some changes for seasonal variation and some reduction in awareness over time as the launch awareness campaigns cease.

There is an observable drop off in inquiries in the late summer months. This is likely a combination of a drop-off in awareness of the hotline from the initial launch publicity efforts, but also effects from summer vacations. If March is "corrected" to make up for the half-month, then the graph appears to stabilise at around ten emails per months.

There is some tracking of awareness raising activities about the hotline by NAIH with the use of the hotline. For example, awareness campaigns (including the radio advert) started at the same time as the service launch. In May 2019, the Budapest Commercial Chamber added details of the hotline on its GDPR information webpage.[5] NAIH Event – practical experiences of the GDPR at BKIK 29 October 2019[6] featuring the President of NAIH and reporting on the hotline aligns with the increased use of the hotline towards the end of October/early November.

[5] https://www.bkik.hu/hu/szolgaltatasok/gdpr

[6] https://www.bkik.hu/hu/hirek/gdpr-gyakorlati-tapasztalatok

Number of questions

Emails to the service could contain more than one question. About half the received emails had a single question. 93% of emails has three or fewer questions, A very small number had more than three. A single email had seven questions. As shown in the chart below.

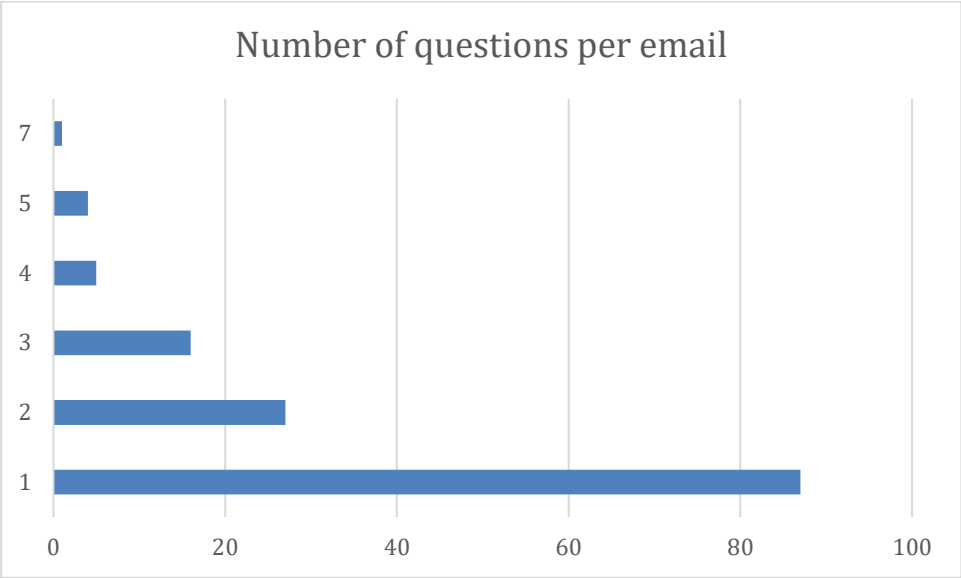

Figure 6 - number of questions per email

Difficulty categories

NAIH assigned emails to the hotline to one of three difficulty categories, based upon the routes those questions would have to take through NAIH. The difficulty categories are the following, taken from NAIH's internal memoranda on the operation on the hotline. Essentially, the hotline email address was staffed by a member of staff who could respond to straightforward queries themselves (having their response checked by a specialist) or reply to out of scope questions, or otherwise escalate these emails to a data protection expert.

Difficulty 1:	A question answerable on the basis of the Knowledge Base, OR, a rejected question (e.g. on the basis that it was too specific and amounts to a request for the lawfulness of a specific piece of data process	27 emails
Difficulty 2:	A question not answerable on the basis of the Knowledge Base but answerable unequivocally by the Hotline Expert on the basis of the law-enforcement practice of the Authority and the supervisory authorities under the General Data Protection Regulation, the relevant case law of the courts, and the documents of the European Data Protection Board assisting the application of law	60 emails
Difficulty 3:	hotline request pose a new or especially complicated question which cannot be answered unequivocally on the basis of the law-	49 emails

	enforcement practice of the Authority and its associate authorities, the relevant case law of the courts, and the documents of the European Data Protection Board assisting the application of law, the Hotline Expert shall involve other experts of the Authority in accordance with executive guidance where necessary.	

The following chart (figure seven) shows the breakdown of the total emails by difficulty level.

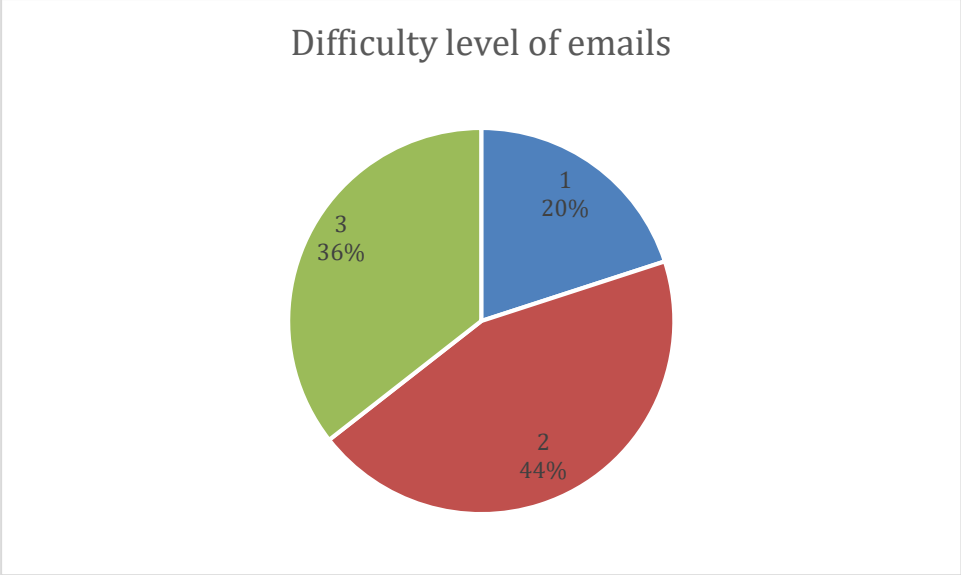

Figure 7 - Proportional difficulty of emails

There is a relatively smaller proportion of simple queries than might be expected. We could have easily expected that most queries could be answered simply, with a proportional number of harder queries – for example, an 80/20 heuristic – where 20% of queries take up 80% of the effort. What may be occurring here instead is that simpler questions SMEs might have can be answered through other "self-service" sources that provide a quicker response (a web-search for example or looking at existing guidance on the NAIH website). SMEs with questions in these categories might be self-selecting out of using the hotline. The implication is that the hotline's knowledge-base becomes either potentially less useful than might be expected, or that it needs to incorporate more complexity.

Contacts were categorised based upon the complexity of the whole request (so, for example, a query containing four separate questions including three easy to answer questions, and one complicated question, would be counted as a level 3 query for this process). If we compare the number of questions in an email with the difficulty assigned to the answer, we get the following charts (figures Eight and Nine).

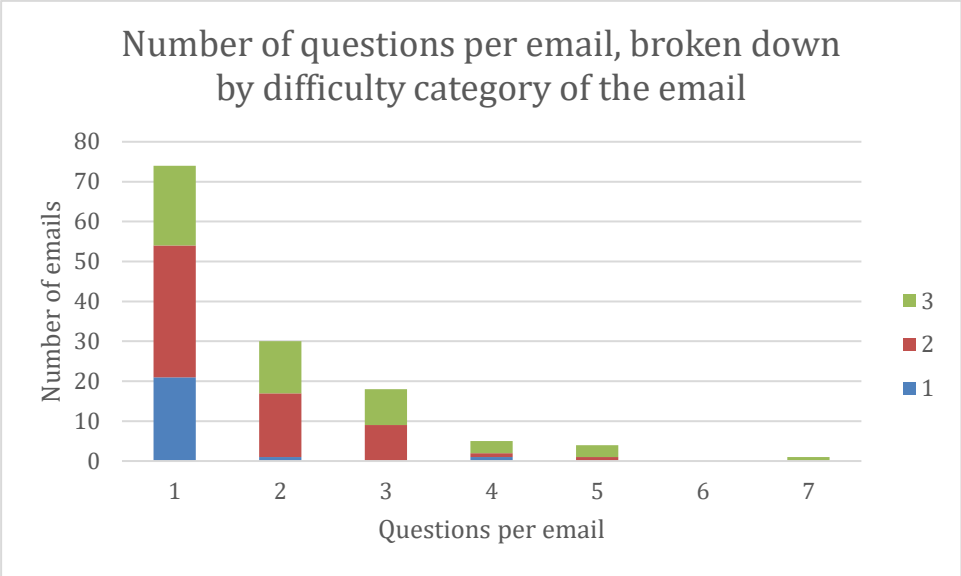

Figure 8 - Questions per email per difficulty category

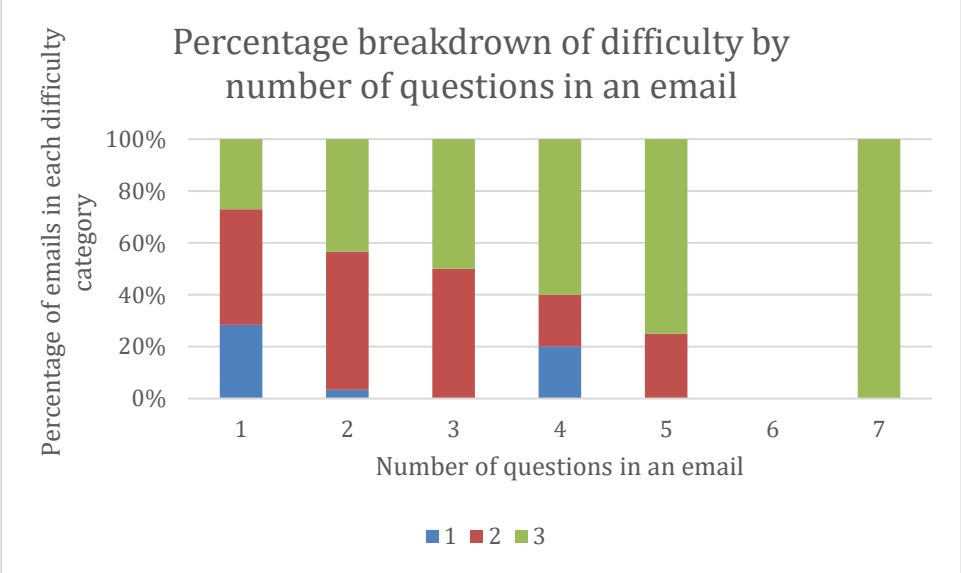

Figure 9 - Difficulty category as a percentage of number of emails

As might be expected, the total number of questions in an email does appear to make a difference to the experienced difficulty of that email to process. As the number of questions in an email increases, the more likely the email (as a whole) is be treated as a category three difficulty, requiring intervention from a data protection expert. Understanding that there are fewer data points at high numbers of questions (e.g. there was only one email with seven questions, and none with six), the change in likelihood is approximately close to what might be expected with random chance – simply, the more questions in an email, the more likely it is that one of them will be very difficult to deal with. There is no real evidence here that emails with more questions in them have inherently more difficult component questions.

Comparing the distribution of difficulty within emails in each number category, the emails with just one question are most similar to the overall distribution of difficulty. (28%/45%/27% compared to 20%/44%/36%). This suggests again that the increase in assigned difficulty come primarily from the larger number of questions.

This does suggest that the NAIH process, of having hotline attendant create answers to category 1 questions in a multiple question email (dealing with "low hanging fruit"), before escalating the enquiry to a data protection expert for the more difficult component parts is an appropriate way of handling multiple question emails. We do not see any particular increase in inherent difficulty for the rare or infrequently asked questions.

The assigned difficulty does appear to vary by the topic of the questions. The most common category of questions are concrete compliance questions. None of these was considered a difficulty 1 question, and the average difficulty level for this topic was 2.55. This likely relates to the need to interpret the meaning of the concrete situation being discussed, and apply the legal requirements to this situation. This likely cannot be done directly from a knowledge base, even a detailed one, and would require additional data protection expertise or perspective. This feature alone could explain the lower number of category 1 difficulty email than expected – because a large proportion on the emails received are about this type of topic, which is not amenable to using a static knowledge base. Compare this to the average difficulty of some other topic: questions about anonymous data have a similar average difficulty of 2.5, whilst the easiest questions to answer come on the topic of records of processing, with an average difficulty of 1.4. General compliance questions are also likely to be in category 1, with an average difficulty of 1.5 (1.4 if we remove the emails with multiple questions on different topics).

How long did it take to respond to emails?

NAIH collected the dates on which an email was received and the date on which a response was issued. From this we were able to calculate response rates in number of days[7]. These are strict calendar days rather than formal working days. In general, most emails were responded to in 5-20 days.[8] The following chart (figure 10) provides the frequency distribution for the response time.

[7] Some of the data included very high response times. During data cleaning, it became apparent that his was a result of incorrectly inputted dates (e.g. putting the date as 2018, rather than 2019). These obvious input errors have been corrected. Smaller, but still high response times (e.g. over a month) have been checked with NAIH for accuracy. These are strict calendar days rather than formal working days, and for these purposes we have not attempted to identify which response times fall over weekends or on national/public holidays.

[8] 95% were answered in < 22 days, 50% answered in < 11 days

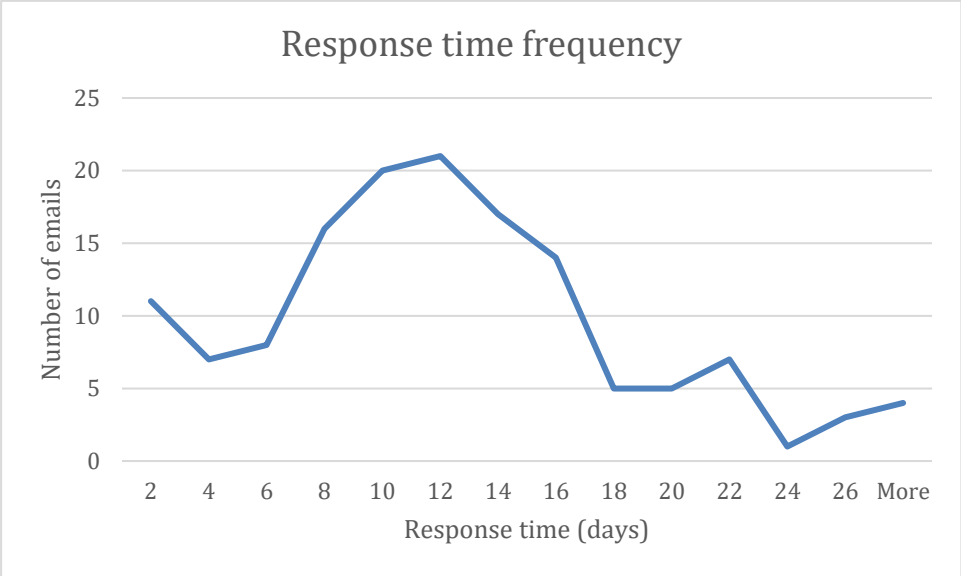

Figure 10 - Frequency of response times

We can further analyse response time by some of the factors which may influence it: the difficulty of the questions, the number of questions in an email, and over time.[9]

The difficulty of the questions received does seem to make a difference to response times, but mostly in the difference between the easiest emails, and the other two more complicated categories. It is not linear, and for example, there are some category one emails that took more time to respond to than category three emails. The following chart (Figure 11) breaks down the response times by the three difficulty level categories.

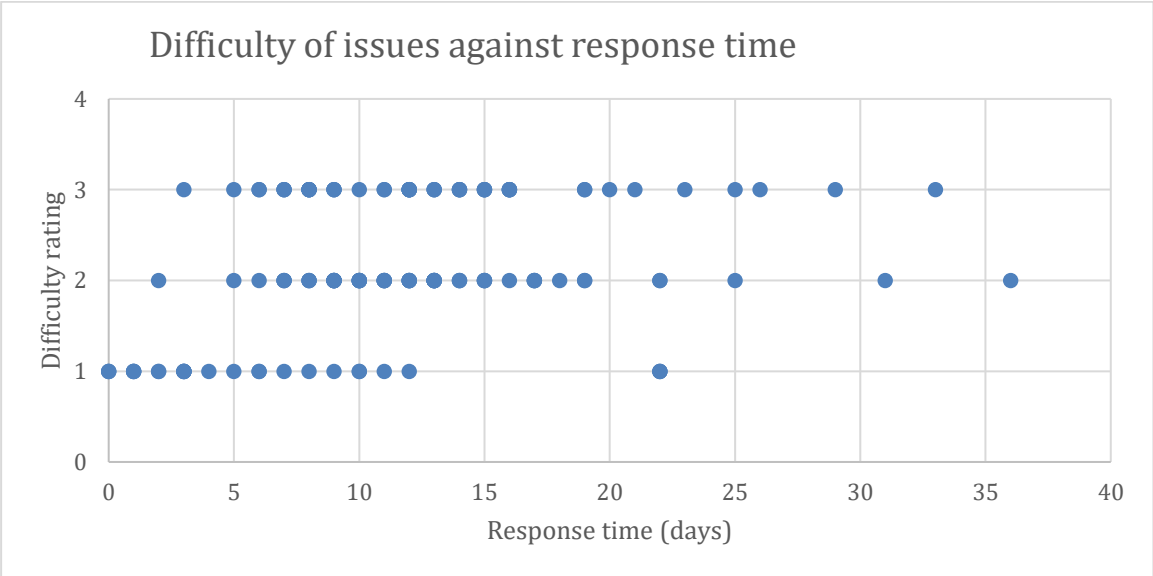

Figure 11 - difficult of issues against response time

[9] Whilst NAIH did assign case workers to each email, we do not consider it appropriate within a research context to analyse differences in response time between case workers, and as such we discarded that data.

As might be anticipated, difficulty 1 emails tend to get answered quicker. Just a little under half are answered within five days, and nearly all within 12 days. A quarter of emails at this level of difficulty were answered on the day they were received or on the following day. Difficulty 2 emails take somewhat longer to answer. Very few (3%) received a response in under 5 days. 78% of these queries got a response in between 5-15 days. The difficulty three emails have a similar distribution to Difficulty 2 emails.

The number of questions raised in an email does not itself appear to make a significant difference to the response time.[10] Figure 12 plots the response time against the number of questions in an email.

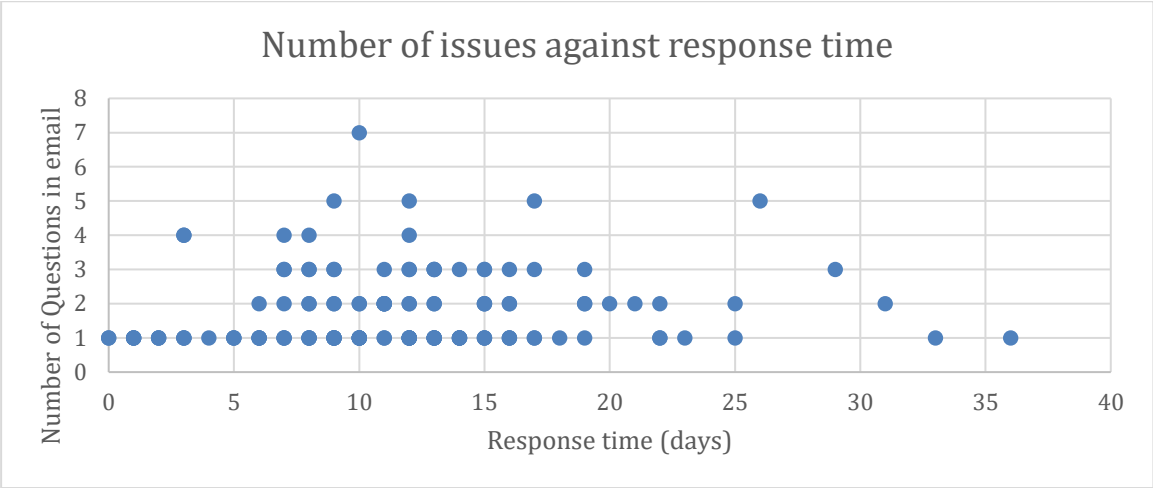

Figure 12 - Number of issues against response time

This may indicate that the response time is determined more by the difficulty of the most difficult question than by the raw number of questions. For example, there are single questions where the response time varies between 0 and 36 days – although the majority are under 20 days. The biggest difference here is that emails with more than one question are very unlikely to be answered in under a week.

We are also interested in any changes in response rate over time. This might indicate if responses were getting quicker as the hotline become more established. The following chart (figure 13) plots response times against the date the email was received, broken down by the three difficulty categories. This shows the distribution of the response times. Figure 14 simplifies the same data (at the expense of representing the variation) by using the average response time each month for the each of the three categories

[10] A linear regression of response rate against number of issues in an email suggests a very low level of correlation between these two variables.

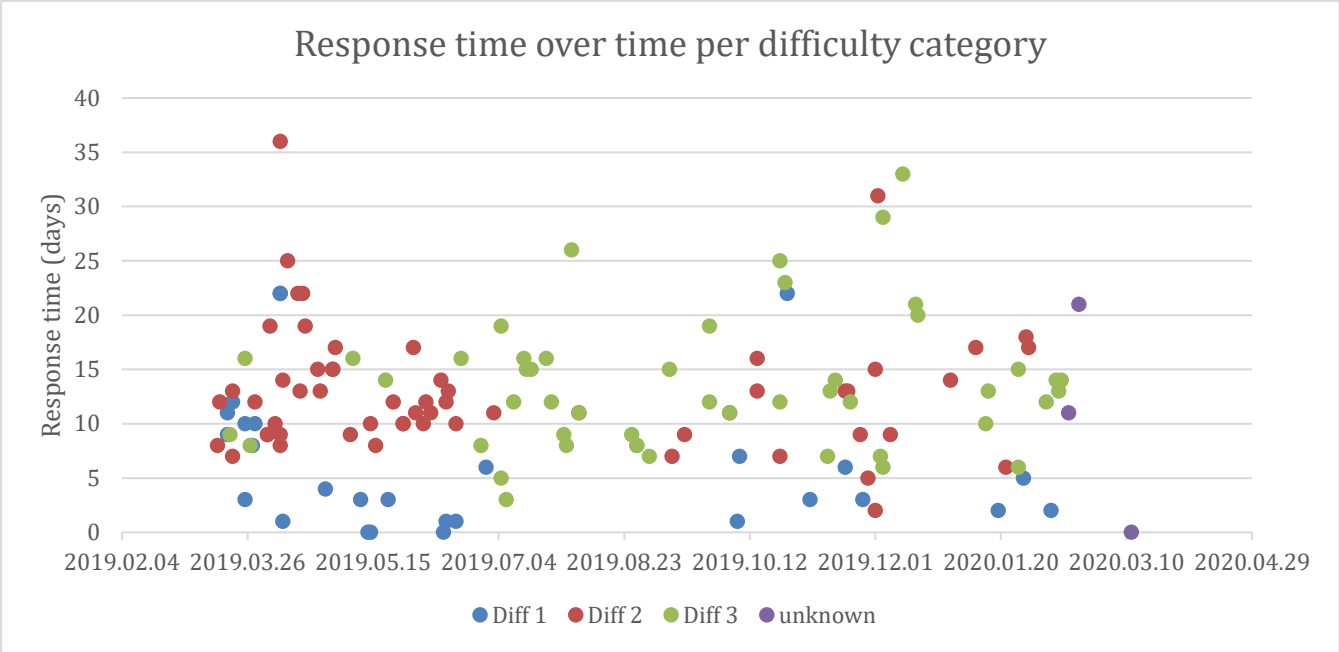

Figure 13 - Response times per difficulty category

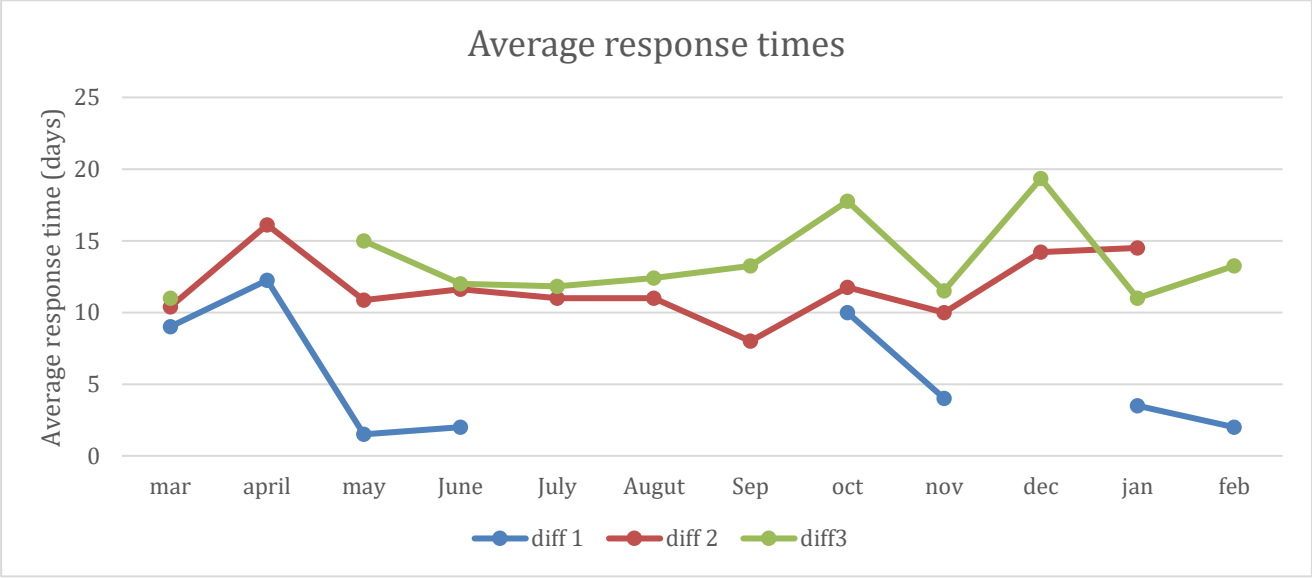

Figure 14 - Average response times

There are two peaks of the slowest response times. The first of these coincides with the start of the hotline, and the second peak of slower responses clusters around the Christmas/New Year vacation period in late December. Plotting questions against time also shows that the summer break does not disrupt all questions equally. The emails received in this period were nearly all categorised as difficulty three, requiring the intervention of additional data protection expertise.

This graph shows us that the response time for the easiest emails does reduce somewhat over the course of the operation of the hotline.[11] After month two, the response time for a category one email is down under a week (apart from one potential outlier) and is typically around three days. This could represent increased familiarity with the project's knowledge base, getting past any initial teething problems, and staff getting used to new processes and roles.

The response time for difficulty two emails does not radically change over the period of operation. Interestingly, this graph suggests that the response time for difficulty category three emails increases over time, as does its variability. We should be aware that as the team involved in the hotline become more experienced, an email that might have been categorized as category three at the start of the project, might only be classed as a category two later in the project. This would leave in category three emails that were difficult to deal with. There are fewer category two emails in the later part of the project.

Response time does not appear to be primarily affected by the volume of emails. If we plot the number of emails per month, against the average response time for all categories, we can see that response time peaks with the initial peak in emails, but doesn't fall over the summer, before having a peak in December, that is not correlated with an increase in emails. This suggests that apart from at periods of particularly high demand, response time is more determined by the working environment and process (e.g. Christmas vacations).[12]

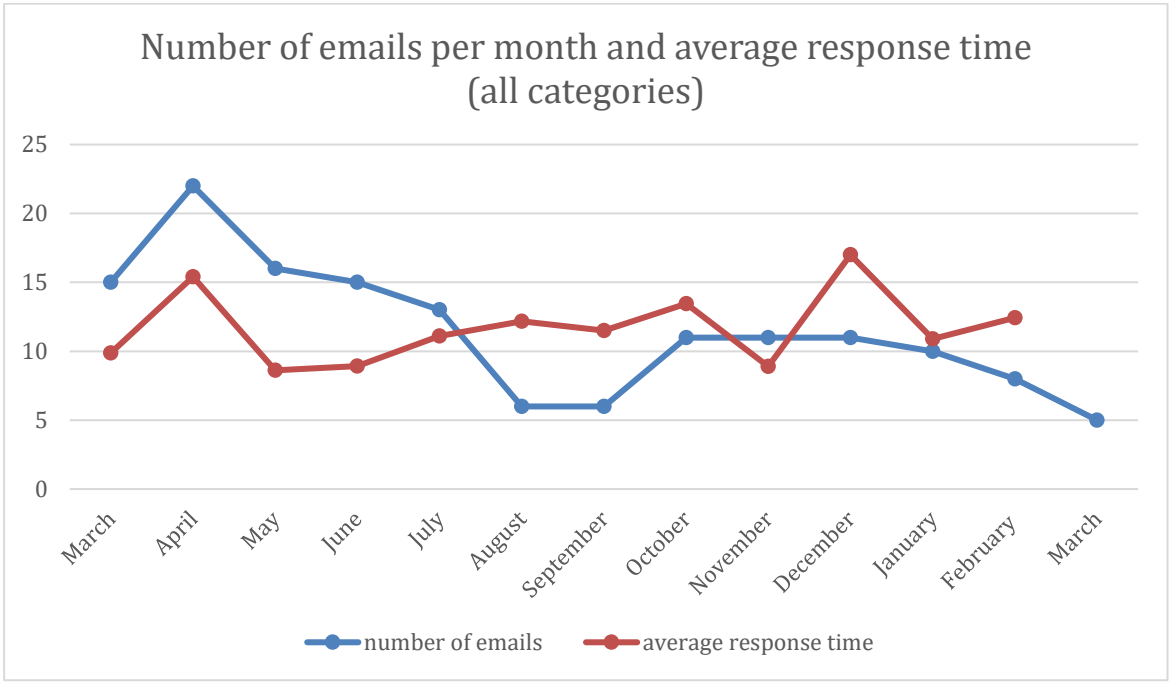

Figure 15 - comparison of number of emails per month and average response time

[11] The line is disconnected because as previously stated, there were no category 1 emails received in the summer months.

[12] Correlation coefficient of +0.09. – 0 = no correlation, 1=perfect correlation

We noted earlier that the most common topic (concrete compliance) questions had a higher than average assigned difficulty level. Looking at the average response time for emails on this topic gave an average response time of 14.7 days (likely in practice to be three working weeks). Contrast this with questions about records of processing. This was a category that was typically assigned to difficulty 1 (could be answered from the knowledge base). The average response time for questions on this topic was 7.6 days, but if emails containing multiple questions were removed[13] they were, on average responded to in 2.8 days. This really suggests that a very significant part of the effort of running the hotline is answering concrete compliance questions.

To try and further understand the variation in response time, we plotted the category 2 and 3 difficulty emails on chronological plot akin to a Gantt chart, showing their start and end date. This gave a visual demonstration of how many emails were being processed at any given time, as well as a way of assessing any comparative changes in response time. The figure below is just an illustration of a section of this large chart. Plotted along the top are the days of operation, and along the vertical axis are the individual emails. The coloured-in section represents the number of days the email was with NAIH before a response was issued.

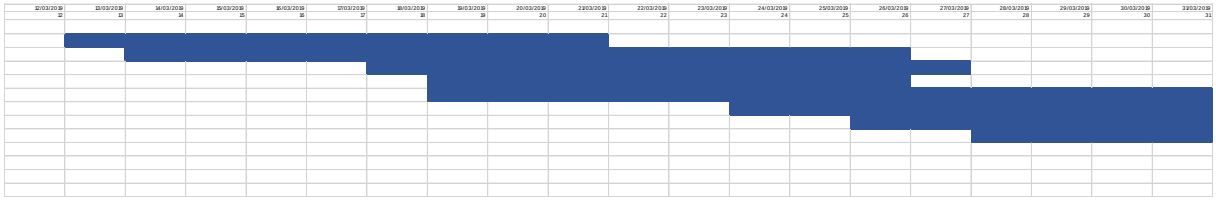

Figure 16 - extract from visual plot of response times

The overall chart is too large to reproduce here, but the pattern it demonstrates is from the first peak in April, there were around ten emails in process until May. Then typically between 6 to 12 emails in process until a drop to 2 to 4 from June. In august there were typically 2-3 emails in process, rising to between 6 to 8 emails from November 2019 through to the end of the hotline, with a comparatively quiet patch around the new year. This analysis suggest that response times are fairly staggered (in a similar manner to receipts), and that there are not many days when a large number of pending responses were cleared. This suggests that there is now artificial bottleneck slowing down these responses. There were some sustained patches were no emails were responded to for some time around Christmas, and a week in the middle of summer, which may well represent staff holidays. Looking at only category three emails, there is a clump in February 2020, where there were six difficult emails in process, but this does not appear to have made a significant difference to either category three or overall response time. This might well indicate that the procedures for responding to category 3

[13] Along with one likely outlier at 36 days – most likely the result of data entry error with the month.

difficulty emails had improved over time. This method also allows us to pull out individual emails with unusual response time (in comparison with other emails occurring at the same time) and investigate these further. For example, this method gives greater certainty that the category 1 email on processing reports (which typically have the lowest response times) recorded as having a 36-day response time is most likely a data entry error! On the other hand, another email that took 26 days to respond to in August, including when it was the only email being processed at the time, did include five different topics and was categorized as difficulty three.

GDPR articles

We can also analyse the average response time for emails involving particular GDPR articles. The following table plots the average response time against the GDPR articles relevant to an email (in this case for just those emails with a single question).

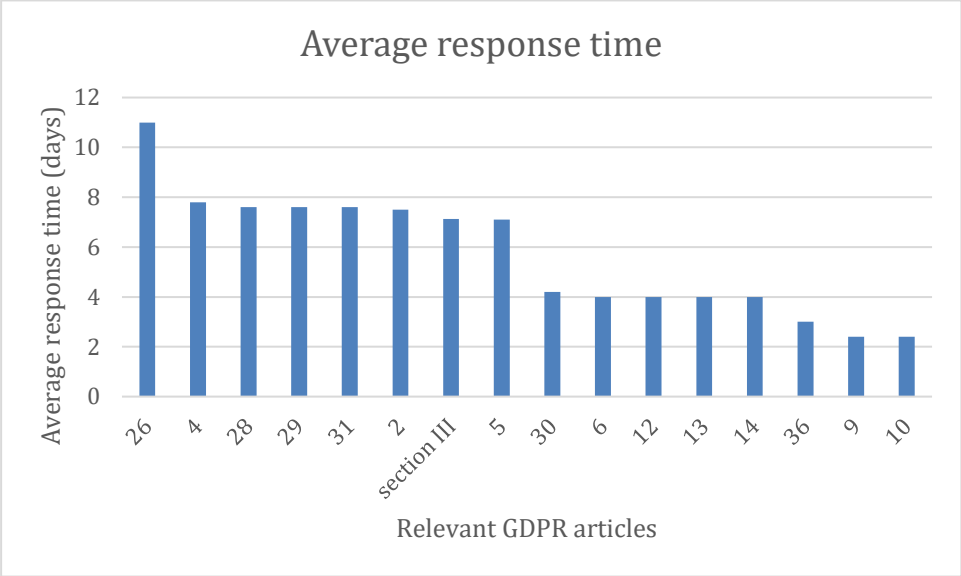

Figure 17 - Average response times for questions on different GDPR articles

First, Article 26 (Joint controllers) only came up in one question, so the average is less reliable, and this may be an outlier. Excepting this, there are three main clusters of response times at around seven days, four days and two-three days.

Cluster 1 – Average response time of around Seven days	
Article 2	Material scope
Article 4	Definitions (article on “personal data”, “controller” and “processor”)
Article 5	Principles relating to processing of personal data
Article 28	Processor
Article 29	Processing under the authority of the controller or processor
Article 31	Cooperation with the supervisory authority
Section III	Data protection impact assessment and prior consultation
Cluster 2 – Average response time of around four days	
Article 6	Lawfulness of processing
Article 12	Transparent information, communication and modalities for the exercise of the rights of the data subject
Article 13	Information to be provided where personal data are collected from the data subject
Article 14	Information to be provided where personal data have not been obtained from the data subject
Cluster 3 – Average response time of two-three days	
Article 9	Processing of special categories of personal data
Article 10	Processing of personal data relating to criminal convictions and offenses
Article 36	Prior consultation

Comparison with expected response times

Understanding actual response times is important because it forms part of managing the expectations of service users. For example, it is harmful to tell people they will receive a response in a week, if it then takes two. Part of the experimentative purpose of STARII is to explore what it actually takes to operate such a hotline, and if initial staffing decisions are These actual response times might be considered acceptable, but this would be for each other authority to determine.

NAIH’s memorandum of operation for the hotline anticipated a response time of 8 working days. Looking at the average response times for the three different categories of difficulty, category 1 difficulty emails were, on average under the upper time limit. However, it appears to have taken more time for each of the other two categories than expected. NAIH may therefore wish to either revise their expected response times, or identify how they can increase their response time in practice. It is not possible based upon the data collected to identify any particular bottlenecks or limitations in the NAIH the process (but given that response time did not vary with the volume of email received this would be a place to investigate further).

Difficulty	**Actual average response time**
Category 1	6.3
Category 2	12.9
Category 3	13.5

Conclusions

Implications for STAR II

This analysis in general provides more confidence in the conclusions already drawn in WP2. These conclusions about the priorities of SMES do not seem to be invalidated by the analysis of the experience of operating the SME hotline by NAIH. This means those conclusions can be used to support the development of the guidance for data protection authorities and SMEs as is intended in STAR II.

The list of common question topics, and the articles that emerge with greater frequency provides some indications of where such guidance material (particularly the SME handbook) should prioritise its effort.

Implications for DPA awareness raising in relation to SMEs

Communicate the scope of the hotline and making clear what type of questions will be answered in order to reduce the number of questions that are out of scope. This should also contribute to increasing satisfaction by managing user expectations. It seems acceptable to continue to allow hotline users to send multiple questions in one email. At the response times observed in this study, it would potentially be appropriate to send status updates (e.g. when an email is escalated to an expert) to the questioner.

Easy questions can be self-serviced. The hotline experienced fewer than expect "easy" questions. This seems to indicate that for these type of questions SMEs will seek to "self-service" the answers rather than take the potential delay of using an email service. An ideal outcome would be to reduce the number of such questions as low as possible by providing common answers to these prominent on the authority's website, or in promotional material targeted at SMEs. As a rule of thumb, if a question can be answered by a staff member copy-and-pasting an answer from a document, that information should be made available to the SME already. There is some indication that a wide knowledge base is more useful than an in-depth one for the purpose of initial triaging of emails.

Time and effort - The majority of staff effort and time in such a service will be taken up by responding to concrete questions where the GDPR needs to be applied to concrete operational context and activities of an SME. These are the most common questions, and tended to have a higher than average difficulty. This means data protection expertise remains important in handling these emails, and they to a large extent cannot be handed off to a knowledge base. In the hotline pilot, the response times were larger than NIAH initially anticipated which should be taken into account in further planning in this area. Dealing with these complicated, practical questions can be time-consuming.

SMEs are not particularly interested in asking theoretical questions. Whilst it is tempted for people with a full and historical picture of the GDPR to respond on the basis of definitions and

principles, this may not be what SMEs are looking for. It may be more valuable to them to focus on more applied areas such as obligations of processors and controllers, or the provision of information to data subjects.

Experience and Intelligence - These services can learn, staff can gain more experience and improve their own knowledge base. Even in a comparatively short run of a year, there is some evidence of learning and improvement (a reduction in the number of middle difficulty emails, an absence of a spike in response time when the team were faced with a higher than usual number of questions towards the end of the project). Operators of an SME hotline should ensure from the start that they are capturing the knowledge they are gaining about the SME experience of the GDPR and the particular problems being encountered. Such hotlines serve as a real-world indicator of what SMEs want information on. This can then serve as a guide for prioritizing training. If 50% of questions are on five topics, then those are a clear training priority, or an area where a supervisory authority might want to develop specific guidance resources. An exception to this is breach notification and handling. There is some evidence here that SMES may be reluctant to bring such questions to the regulator.

Origin of guidance - SMEs are primarily interested in information from "their" regulator, even if this information is already available elsewhere from another regulator. Significant ground could be gained for SMEs by either direct adoption and translation, or active endorsement of support material created by their peers.

Consistency - If a hotline is going to effectively shut down over a holiday period then this should be communicated to users. Alternatively, vacation cover should be arranged in order to keep response times in line with other parts of the year. Over time, services such as this will get a better sense of their peak and ebb times.

Data collection – understanding the differences in operation of a hotline can potentially be enhanced with some changes to the way that data is collected during the process. It has proved difficult with this second-order analysis to determine to what extent staff on the hotline were dedicated to the hotline, and to determine what other pressures may be impacting upon the response time of the hotline. Whilst we can work out a response time based upon when emails were received and responded to, it is impossible to determine from the data collected the actual person-hours spent on each email. Hotline operators would be advised to consider the time and actual cost of their current methods of engagement. Having a good number of categories for topic means that relatively few topics are placed in the more difficult to analyse "other" category. It could be possible to have clearer data by disaggregating response times and difficulty per question, rather than per email, but this would add an additional burden upon hotline operators which is likely not warranted. For NAIH's process analysis it would be beneficial to track those difficulty 1 questions which were not initially approved by the expert, but needed revision.

Some areas of the GDPR are not raising questions for SMEs. We can't tell from this data if this is because those topics do not impact upon SMES, or that SMEs are not even aware they might need to asking questions about these topics. However, when these are topics that are 1) novel and 2) still being developed at a theoretical and research level (e.g., data protection by design and default), or have quite high legal uncertainty (e.g., international transfers) we can suspect they are not "on the radar" for SMES (at least in the Hungarian context, this may well differ across Member States).