An Email a Day Could Give Your Health Data Away

Christof Lange, Thomas Chang, Maximilian Fiedler, and Ronald Petrlic

Nuremberg Institute of Technology, Nürnberg, Germany
ronald.petrlic@th-nuernberg.de

Abstract. Are doctors allowed to communicate with their patients via email? The GDPR sets the bar high for securing health data: either an end-to-end-encryption (E2EE) or a guaranteed transport encryption needs to be used. As E2EE (with PGP or S/MIME) is not widely used in practice, only a guaranteed transport encryption comes into question. But are doctors' email servers properly configured and provide such strong security guarantees? As we found out in a large-scale investigation of German medical institutions, this is not the case at all. Only a very small minority of email servers provides state-of-the-art security. In all other cases, communication between doctors and patients via email is not secure and, thus, not permitted with regards to the GDPR.

Keywords: Email Security · Opportunistic Transport Encryption · GDPR

1 Introduction

"Emails are like postcards; everybody can read them"—this analogy is often used by people when talking about email security. Looking at email security in more detail, one can see that this analogy is not very suitable, though. There are mechanisms (e.g., transport encryption) in place that guarantee confidentiality and authenticity to a certain extent, as we will point out in the next section. However, guarantees "to a certain extent" do not always suffice in practice. There are cases where we want true guarantees that the emails we send cannot be read or forged by any unauthorized third parties; one might think of communication with doctors, for example. End-to-end encryption (E2EE)—and signatures—with PGP or S/MIME can provide such guarantees. However, E2EE has still not found its breakthrough in practice for securing emails—leaving people with a certain feeling of unease when sending emails with sensitive content without any further protective measures (having the postcard analogy in mind). People can just not estimate whether their email will be sent in a secure way to the proper destination or not; chances are good[1], but there is no guarantee.

[1] Google regularly publishes numbers that show that around 90 % of all emails sent and received via Google Mail are protected with transport encryption: `https://transparencyreport.google.com/safer-email/overview`

Coming back to our communication with doctors[2] example from above, we should not only ask ourselves whether such communication via "unencrypted" email (i.e, without E2EE) is a good idea, but whether such communication is allowed at all. Since May 2018, the General Data Protection Regulation (GDPR) is applied in Europe, raising the bar in terms of *security* for controllers. According to *Article 32 GDPR*:

> Controllers shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk—taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons.

In our example, with respect to the nature, scope and context of the transmission of sensitive medical data via email and, thus, (potentially) high likelihood and severity for the rights and freedoms of patients (if their medical data gets disclosed), one can see that proper technical measures need to be set in place to ensure a level of security appropriate to the potential high risk. What would be such proper technical measures in our scenario? The solution seems to be clear: E2EE, being a state of the art measure with low costs of implementation. And yet, practically none of the medical institutions[3] have the ability to communicate with E2EE (with their patients). Even if medical institutions would have the ability, there would be few patients with whom they could communicate in a secure way, as most patients also do not have E2EE facility set up on their devices. [18] Despite all of that, email communication between medical institutions and patients is taking place (without E2EE), be it to set up appointments, send examination results and blood values, sick notes, covid test results, and so forth.

Not only since the application of the GDPR do data protection authorities (DPAs) argue with controllers (companies, medical institutions, public authorities,...) whether they need to employ E2EE in order to communicate with their customers, patients, citizens, etc. [25] In March 2020, the German DPAs[4] published a guideline for controllers[5], in which they state in detail which kind of technical measures need to be taken by the controllers (according to the risk of the communication—based on the sensitivity of the data) in order to be on the safe side (and not risk fines). Controllers that *intend to receive emails with high risk* or that *send emails with high risk* need to take (one of) the following technical measures:

- provide public key for E2EE; send emails with E2EE

[2] In the rest of the paper, we use the more general term medical institution, which also covers clinics and medical care centers.

[3] The same is true for other businesses and authorities as well.

[4] In Germany, other than in other European countries, there are 18 DPAs in toal.

[5] `https://www.datenschutzkonferenz-online.de/media/oh/20200526_orientierungshilfe_e_mail_verschluesselung.pdf`

– provide the facility for a "qualified" transport encryption

A qualified transport encryption according to the German DPAs means that TLS 1.2 (with secure cipher suites and perfect forward secrecy) or TLS 1.3 is used and that DANE is employed.

Summing up, we can conclude the following:

– Medical institutions have websites where they state their email addresses (and, thus, *intend* to receive emails with high risk—as emails containing sensitive medical data always impose a high risk),
– Medical institutions communicate with their patients via email (during covid pandemic time, covid test results were sent without any further protection mechanisms via email, for example),
– E2EE is generally not employed in email communication between medical institutions and patients.

Medical institutions would therefore need to make sure that their email servers are configured according to the requirements of the DPAs, i.e., support a qualified transport encryption. Otherwise, the medical institutions would put their patients' medical data at risk (when communicating via email) and furthermore risk high fines by DPAs.

In this paper, we investigate whether medical institutions in Germany adhere to the requirements by the DPAs with regard to secure email communication.

2 Background

This section describes relevant parts of the underlying principles of email communication.

2.1 SMTP Protocol

SMTP is a plaintext protocol originally defined in 1980 as part of a system for asynchronous email communication [7]. SMTP is used to transport the email from the sender's email client to the sender's email server and from the sender's email server to the receiver's email server. To finally collect the emails from the server, the receiver needs to request the mails using IMAP or POP.

Since it was first defined, the protocol has received multiple extensions, like CHUNKING for large messages [16], AUTH for user authentication [13] or STARTTLS for TLS transport encryption [14, 15] and many more.

2.2 Opportunistic Transport Encryption

For many years, data security and privacy protection was not a big concern and emails were sent completely without any security measures.

The first step in this direction was E2EE using tools like PGP [10]. PGP (and also S/MIME [3]) enables senders to encrypt their emails at their side and decryption is only possible for the legitimate receiver (with the proper private key); neither eavesdroppers nor the involved email servers can read the contents. This approach requires additional efforts for both sides of the communication (e.g., in our scenario: for patients and doctors), because they have to actively exchange their public keys.

The next step towards more privacy (more precisely, confidentiality) was transport encryption for the SMTP protocol using the previously mentioned STARTTLS extension [14]. Up until the STARTTLS command is issued by the sender, the communication is still sent over plaintext. The STARTTLS message then initiates a TLS Handshake where a common TLS version is negotiated and session keys are generated using public key cryptography. All of the following communication in this session is then encrypted.

This way of establishing an encrypted communication channel is easy to attack for a man in the middle (MITM), though. An eavesdropper with the ability to modify the communication could simply strip the STARTTLS capability from the EHLO response and therefore enforce plain text communication.

2.3 DNS-based Authentication of Named Entities

To mitigate this so-called downgrade attack, a technology called DANE can be used. It uses DNSSEC and a special DNS record type, called TLSA, to prove the authenticity of a mailserver. A domain owner needs to enable DNSSEC for their domain and create a TLSA entry, which stores either a signed TLS certificate, a public key or a hash value of one of these. When a mail transfer agent (MTA) wants to forward an email to this server, it firt checks if the contents of the TLSA record match with the contents of the certificate provided by the server during the STARTTLS initiation. This way, DNS is used as a trusted communication channel to prove the authenticity of the certificate used for the mail transport. DANE therefore makes it impossible for an MITM to just replace the certificate in transmission. This implicitly prevents a downgrade attack, since a mailserver with a valid TLSA record always has to support STARTTLS encryption. An additional security advantage of DANE is that the owner of the domain can specify a trust anchor of his choice and is not bound by official certification authorities.

3 Related Work

Durumeric, et al. [21] examined the coverage of STARTTLS in Gmail Inbound traffic between January 2014 and April 2015, recording an increase from 52% to 80% coverage within one year (mainly due to the start of support by Yahoo and Hotmail). Zhu, et al. [26] measured a low DANE support between 0.05% (for .com-domains) and 0.23% (for .net-domains) from March 2013 to October 2014.

Foster, et al. [22] compared the TLS support of the top 22 email providers in 2013 and 74 web services of the Alexa Top 100 list that send emails as part of their service. They found that the percentage of users that use email providers with TLS support increased from 52% in 2014 to 89% in 2015. The TLS support of email-sending web services is generally lower and varies between categories, being the highest for financial institutions (67%) and lowest for news and dating sites (0%).

This increasing trend can also be observed in a DANE study [23] where domains worldwide (50% from Germany) have been tracked between 2014 and 2020. The study concludes with a result of 96.4% secure domains regarding STARTTLS (ranging from 97.6% having at least one TLS-secure MX record and 94.3% exclusively having secure MX records). The authors also report a valid DANE support for 17.6% of domains.

Lee, et al. [24] further examined the regional differences in DANE support by using data from OpenINTEL [9]. By comparing international top-level-domains (.com and .org) with country-top-level domains (.se and .nl) between October 2017 and October 2019, they found a higher DANE support for region-specific email servers (38.2% for .se and 9.8% for .nl) than for international ones (0.6% for .com and 0.73% for .org). Despite the generally low coverage of DANE support, they could observe an increasing trend in DANE support over these two years.

3.1 Our Contribution

In this paper, we present further analysis regarding email security as of June 2022. Particularly, we examine the security of email servers of medical institutions in Germany by measuring the used TLS versions and the coverage of DANE support among the used email providers.

4 Methodology

This section describes our methodology for performing a security analysis of email servers used by medical institutions.

4.1 Dataset Generation

Our analysis is based on domains of German medical institutions as of May 2022. We developed a python web crawler that extracts registered profile information of medical institutions from *jameda.de*, an independent medical appointment booking platform.

To extract these profiles, combinations of all available German cities and medical subject areas are queried. Many profiles contain a link to individual websites, which are further scanned for email addresses using regular expressions.

All information is compiled by the crawler into a dataset with the following structure:

- list of profiles containing:
  - name of responsible doctor
  - medical subject area
  - extracted email addresses
  - link to individual website and imprint section
- list of all email addresses
- list of all domains
- list of websites that could not be scanned automatically

Some websites may contain multiple email addresses, including contact details of medical associations, website building services and data protection officers. In order to eliminate non-medical email addresses, we performed a cleanup: each email address has to contain a part of the doctor's name or the medical institution's website URL in order to be seen as valid.

There were some cases in which no matches could be found and manual verification was necessary.

With this approach, we gathered a total of 3772 email addresses (4414 before cleanup) and 2938 domains (3382 before cleanup).

However, it is also possible that the crawler also collects email addresses that are not visible to humans due to styling and formatting. Because of this, the dataset could contain some outdated email addresses which are no longer used.

4.2 Security Analysis

The first step towards analyzing the email server's security configuration is to extract the domain from the email address and retrieve the respective `MX` and `A` records of the responsible SMTP server(s). Then a TLS version test and a check for DANE support is performed for each individual SMTP server.

DNS Records The developed software uses the tool dig for performing DNS queries [4]. We use the Google DNS as DNS server, which can be reached under the IP `8.8.8.8` [6]. It supports DNSSEC, which means that we can determine if it is a validated response by looking at the `AD` flag [5]. This is important for the DANE analysis. Through trial and error, we found that if too many requests are made, not all of them will be answered. For this reason we implemented a rate limit of 10 queries per second. The following steps are conducted to get the `MX` records and corresponding `A` records:

1. The domain name is extracted from the email address. Then all domain names are deduplicated which results in a list of distinct domains.
2. For each domain, all `MX` records are queried and saved. Note that we do not check whether implicit `MX` is used, which is why not all email servers may be covered [12]. In addition, it is stored whether the response was validated by DNSSEC (for DANE analysis).
3. For each `MX` record, all `A` records are queried and saved.

The result is a data structure in which all `MX` records are stored for each domain. Each `MX` record itself has one or more A record entries.

To increase reliability of mailing, there can be multiple SMTP servers for a domain. This can be due to multiple `MX` records or multihomed hosts (multiple `A` records for one `MX` record). The order in which the SMTP servers are contacted depends on the preference values of the `MX` records and the order of the `A` records. If there are multiple `MX` records, the sender tries to contact the host with the smallest preference value first. The `A` records are processed in the order in which the entries were provided by the DNS resolver.

One consideration was whether it is necessary to test each individual SMTP server or if it is sufficient to test the first SMTP server from the `MX` record with the lowest preference value (highest priority). If an SMTP server is not reachable, then the sender selects the next one in the ordered list (but there may be an installation-specific limit for connection attempts) [12]. When a domain offers multiple SMTP servers, but not all of them offer the same level of security, an attacker can take advantage of this fact. He blocks the SMTP connection attempts to the secure SMTP servers and, thus, the sender connects to the insecure email servers. Because of this, we decided to test each email server individually and verify if *all* SMTP servers from a domain implement a security feature.

TLS Version Scan Our goal is to determine whether an email server offers encryption via `STARTTLS` and if so, which TLS versions are supported.

For this purpose, the open-source tool testssl.sh was chosen because it is actively developed and has a high popularity [17]. The tool is supplied with each unique combination of `MX` record and IP address from the DNS lookup done before. With our configuration, it connects to the server on port 25 and performs the SMTP protocol. It is tested, whether `STARTTLS` is offered at all. If so, the availability of SSLv2, SSLv3, TLS 1.0, TLS 1.1, TLS 1.2 and TLS 1.3 is determined.

DANE Scan The next step is to analyze if an email server supports DANE.

For a secure implementation of DANE, (1) DNSSEC must be enabled (for the retrieval of the `MX`, `A` and `TLSA` records), (2) a `TLSA` entry must be published, and (3) `STARTTLS` must be offered with a correct certificate chain [19, 24].

Whether DNSSEC is available for querying the `MX` records is included in the data from the DNS queries. To check the remaining conditions, we employ the open source program gotls [11]. The tool receives the `MX` record and DNS server as input. In this case the Google DNS server is passed. The domain of our server is used as the hostname for the EHLO command.

First, all `TLSA` entries are queried. Through the use of DNSSEC, it must be ensured that the response is authentic. Then the `A` entries are queried and a connection with the individual SMTP server(s) is established. The tool checks if the presented certificate matches the data from the `TLSA` entry. From the console

log it is possible to extract whether DANE is available and correctly implemented for a particular SMTP server.

4.3 Execution

The runtime of our program[6] on a server with 2 vCores is approximately 2 hours for the dataset described in section 4.1. Parallelization allows multiple servers to be tested simultaneously.

To avoid spam, some email servers have blocked requests from dynamic IP addresses. Therefore we run our program on a server with a static IP address. Furthermore, an rDNS record is entered in the DNS, which points to our server.

5 Evaluation

In this section, we present the results of our email server analysis. We analyzed 2938 different domains. Recall that there may be multiple SMTP servers for one domain. For 2836 domains (which are used by 2806 medical institutions) there is at least one valid test result for an SMTP server. For the further analysis, only those domains are used.

A detailed breakdown of the dataset is shown in Table 1.

Table 1. For the majority of domains we were able to obtain test results

Status	Domains
No `MX` Records	77 (2.6%)
No valid test results (e.g. connection problems, blocklist)	25 (0.9%)
Valid test results (for at least one SMTP server)	2836 (96.5%)

5.1 TLS Deployment - Domains

A test is performed to determine whether STARTTLS is supported and if so, which TLS versions are offered.

Highest Possible TLS Version First we check the maximum TLS version supported by *at least one* SMTP server under a domain (the maximum TLS version of the strongest member of the group). However, there is no guarantee that a patient can reach the SMTP server(s) with the high TLS version (for example, because an attacker blocks the connection). Therefore, we also examine the maximum TLS version that *all* SMTP servers of a domain support (the maximum TLS version of the weakest member of the group).

[6] The sourcecode for the project is available at `https://git.informatik.fh-nuernberg.de/email-research/medical-institution-email-check`

2712 domains (95.6%) have the same highest TLS version for each SMTP server. There are differences for 124 domains (4.4%): Here the maximum TLS version for at least one SMTP server is different.

The analysis for *all* SMTP servers shows that 59.6% of the domains use TLS version 1.3 and 39.8% use TLS version 1.2. The TLS versions SSLv2 to TLS 1.1, which are considered insecure and are not state-of-the-art anymore, are used by 0.1% of the domains. No encryption is possible for 0.5% of the domains. Table 2 shows detailed results.

Table 2. The majority of domains (99.5%) offers encryption with STARTTLS

Highest possible TLS Version	Domains (at least one SMTP server)	Domains (all SMTP servers)
No STARTTLS	9 (0.3%)	13 (0.5%)
SSLv2	0 (0%)	0 (0%)
SSLv3	0 (0%)	0 (0%)
TLS 1	0 (0%)	2 (0.1%)
TLS 1.1	0 (0%)	0 (0%)
TLS 1.2	1012 (35.7%)	1130 (39.8%)
TLS 1.3	1815 (64%)	1691 (59.6%)

Lowest Possible TLS Version The lowest TLS version of a domain is the minimum TLS version that *at least one* SMTP server under that domain supports (from the weakest member of the group, the minimum TLS version). This is the worst case (from an encryption point of view) where the email is still accepted by the receiving end with transport encryption.

2796 domains (98.6%) have the same minimum TLS version for all SMTP servers. For 40 domains (1.4%), at least one SMTP server offers a different minimum TLS version.

We find that 72.3% of domains accept emails with the deprecated versions SSLv2 to TLS 1.1. 27.7% of domains require at least TLS version 1.2 for email transport. Table 3 shows detailed results.

5.2 DANE Deployment - Domains

First, we take a look at the domains where at least one SMTP server offers DANE. From a security perspective, this is not recommended. An attacker could block the connections to the email servers with DANE support and force a connection to insecure email servers. Therefore, we also check how many domains offer DANE for *all* SMTP servers. We find that 25 (0.9%) domains provide DANE for *at least one* SMTP server and 25 (0.9%) domains offer DANE for *all* SMTP servers (see table 4). The fact that each of the domains supports DANE for all SMTP servers guarantees that the patient can take advantage of the security features.

Table 3. 2041 (72.3%) domains allow encrypted email sending with the deprecated TLS versions SSLv2 to TLS 1.1

Lowest possible TLS Version	Domains (at least one SMTP server)	Domains (all SMTP servers)
SSLv2	2 (0.1%)	1 (0%)
SSLv3	23 (0.8%)	19 (0.7%)
TLS 1	1845 (65.4%)	1823 (64.5%)
TLS 1.1	171 (6.1%)	172 (6.1%)
TLS 1.2	782 (27.7%)	812 (28.7%)
TLS 1.3	0 (0%)	0 (0%)

For DANE, according to RFC 7671, at least TLS 1.0 must be supported, and TLS 1.2 or higher should be supported [20]. In our dataset, all domains that support DANE for all SMTP servers use at least TLS 1.2.

Of the 25 domains that support DANE, 4 are from major German email providers ("web.de", "gmx.de", "gmx.net" and "posteo.de")[7]. The rest are, according to the name, own domains of the medical institutions. The email infrastructure of the latter domains is provided by 6 providers (see section 5.3 for details on identifying the email infrastructure provider). In 14 out of 21 (60%) cases, this is "one.com". The remaining 7 domains (40%) are distributed among 5 other providers.

The 25 domains are used by 81 medical institutions that offer at least one email address that supports DANE. 22 of the domains are used by 66 medical institutions, of which all email addresses support DANE. Of the latter, 47 facilities use email addresses from major providers ("gmx.de/net" and "web.de") and 19 have their own domain.

Table 4. Only a small percentage of domains (0.9%) supports DANE

Status	Domains (at least one SMTP server)	Domains (all SMTP servers)
DANE Ok	25 (0.9%)	25 (0.9%)
DANE Not Ok	2.811 (99.1%)	2.811 (99.1%)

5.3 Email Infrastructure Provider Analysis

If a medical facility decides to accept emails through its own domain, email server management can be outsourced to an external provider. These providers can provide the email infrastructure for many domains. This means that the security level of all managed domains depends on what the provider offers.

[7] web.de and GMX actually are part of the same company called "1 & 1 Mail & Media GmbH". Web.de and GMX are very popular freemailing services in Germany; there exist paid premium services as well, though.

We investigate which providers manage the most domains for our dataset and check which TLS versions are offered and whether DANE is supported.

Extracting Providers from MX Records In our dataset, the `MX` records are used to determine the providers. We found that the naming of the hosts of a provider differs only on the subdomain level. For example, a provider has the `MX` records *mail1.myprovider.de* and *mail2.myprovider.de*. For the analysis, the two `MX` records are combined into one provider named *myprovider.de*. All domains that have **.myprovider.de* (star as wildcard) as `MX` record are counted as managed domains of the provider *myprovider.de*. Note that not all providers may be captured correctly using this method, as the naming of `MX` records may differ.

It can be stated that only 6 providers (out of total 789) manage over 50% of domains[8]. A change in security measures at these providers could have a major impact on the security level of many domains.

TLS Deployment - Providers We only analyze the 6 top providers in more detail.

When looking at the SMTP servers per provider, we find that all servers offer the same maximum TLS version. For the minimum TLS version, this only applies to 5 of the 6 providers: kasserver.com has servers in our dataset that offer TLS 1.0 as well as ones that offer TLS 1.1 or TLS 1.2 as the minimum version.

All providers offer the latest TLS versions 1.2 and 1.3. However, at the same time 4 out of 6 providers still support the deprecated version TLS 1.0. The other 2 providers only allow encrypted receiving with at least the secure TLS version 1.2. Table 5 shows the detailed results.

DANE Deployment - Providers To check whether a provider supports DANE, its `MX` records and the associated SMTP servers are considered. The difference to the domain DANE analysis is that it is not checked whether DNSSEC is also activated for querying the `MX` records. It is therefore possible that a provider enables DANE, but the domain owner has not enabled DNSSEC (e.g. missing `DS` entry in the parent zone).

None of the top providers supports DANE and, thus, leaving the email communication of their customers (in our scenario: medical institutions) with their users (here: patients) insecure.

Microsoft as a Service Provider We were surprised that 249 medical institutions use email services provided by Microsoft, because Microsoft has long been in the focus of criticism of German data protection authorities—justified or not. We first suspected that the medical institutions made use of a free outlook.com

[8] Since one provider can manage multiple domains (`MX` records of two providers on one domain) it is actually the sum of the managed domains of all providers

Table 5. Six providers handle the mail transport for over 50% of the domains

Provider	Domains	Min. TLS Version	Max. TLS Version	DANE
kundenserver.de	397 (13.6%)	TLS 1	TLS 1.3	No
rzone.de	373 (12.8%)	TLS 1.2	TLS 1.3	No
outlook.com	239 (8.2%)	TLS 1.2	TLS 1.2	No
kasserver.com	237 (8.1%)	TLS 1 (TLS 1.2)	TLS 1.3	No
ispgateway.de	163 (5.6%)	TLS 1	TLS 1.2	No
ionos.de	131 (4.5%)	TLS 1	TLS 1.3	No

account, which would have been definitely a problem in terms of privacy protection. One must know that Microsoft offers both private ("consumer") and business email services.

Microsoft differentiates between private and business customers both in terms of the *Service Agreements* and in terms of the *Privacy Policies*—and this differentiation is also relevant for the email service, and, thus, our scenario.

As an example, Microsoft states in its service agreements for private customers[9] that the provided services are only intended to be used for private and non-commercial usage. Medical institutions that use a (free) private email service from Microsoft might violate the service agreements[10]. Moreover, and this is the more critical part: Microsoft states in its privacy policy for private customers[11] that *personal data are used for*:

- Development and improvement of products
- Personalization of Microsoft products and provision of recommendations
- Targeted advertisement

To sum it up, Microsoft could use all the personal data (in our scenario: medical data of medical institutions' patients) sent and received via email through a medical institution's private customer email account as test data for product development, for feeding recommender systems and for advertisement purposes.

Medical institutions, as controllers according to the GDPR, shall only use processors that provide *sufficient guarantees* to implement appropriate technical and organisational measures according to Art. 28 GDPR. A private email service provided by Microsoft definitely does not provide sufficient guarantees and must not be used by medical institutions for communication with patients.[12]

[9] `https://www.microsoft.com/en-us/servicesagreement`

[10] And due to an absence of privacy guarantees for private accounts, such medical institutions might violate another term in the service agreement: The Code of Conducts requires: "Don't engage in activity that violates the privacy of others."

[11] `https://privacy.microsoft.com/de-de/privacystatement`

[12] It should be noted that this might not only be the case for Microsoft but for other freemail services (like web.de or GMX) as well; we did not check their service agreements and privacy protection policies in detail, though.

We thus needed to find a way to check whether a Microsoft email address is based on a private or business account. We can make a distinction between three different variants:

(1) As a private customer, an account can be created at outlook.com, whereupon an email address is created which ends with outlook.com. Depending on the country, other top level domains are also possible, for example outlook.de. The service outlook.com has been renamed several times over the years, which is why there are still a number of "legacy" domains[13] [8]. If such an email address is used, the private plan is used.
(2) Both private and business customers can use their own (paid) domain with the underlying email infrastructure provided by Microsoft [1,8]. This can be recognized by the fact that the `MX` record of the domain points to a Microsoft email server.
(3) Business customers are provided with an email address that ends with onmicrosoft.com [2]. In this case, a business plan is used.

Table 6 shows the breakdown of our analysis.

Table 6. Breakdown of the adoption of Microsoft's email services. A medical facility is included in a category if at least one email address from the facility meets the criterion.

Variant	Medical institutions
(1) private outlook.com address	9 (0.3%)
(2) Microsoft email service with own domain	240 (8.6%)
(3) Microsoft business email account	0 (0%)

Especially variant 2 was of special interest for us, as most medical institutions use this variant and we could not directly derive from the data whether those email addresses are based on a private or business account. We then analyzed the `MX` records in more detail and noticed that they ended either with "mail.protection.outlook.com", "msv1.mail.outlook.com" or " mail.eo.outlook.com". In the next step, we needed to find out whether private or business accounts are provided with such records. We therefore created different Microsoft email accounts: private and business accounts, which we all linked to our own paid domain. We found that private email accounts get an `MX` record ending with "pamx1.hotmail.com" and business accounts get an `MX` record ending with "mail.-protection.outlook.com". Help pages on the internet indicate that the other `MX` records were previously used for the business offer. Thus, we are quite sure that the 240 medical institutions that use a Microsoft email service with their own domain use a business account with proper guarantees in terms of security and

[13] We used outlook.com, outlook.de, live.com, live.de, msn.de, msn.com, hotmail.com, hotmail.de for the analysis

privacy protection and these cases are not as problematic as they first seemed to be.[14]

6 Conclusion

In this paper we investigated the state of email security of German medical institutions. For this purpose, we built a dataset of email addresses from a medical portal and analyzed the corresponding email servers. We can conclude that (1) the support of STARTTLS is high (99.5% of domains). Furthermore, we can state that (2) 99.5% of the domains support the current TLS versions 1.2 and 1.3. The availability of DANE (3) is very low with 0.9% of the domains, however. In addition, we derived from the `MX` records which provider manages the email infrastructure of a domain. We can see (4) that only few providers (0.6%) manage more than 50% of the domains. The top providers all enable current TLS versions, but DANE is not supported by any of them. If the providers would implement state-of-the-art security properly, all of their customers' email communication would be secured at once.[15]

Only 25 of 2836 analyzed domains of medical institutions (used by 66 different medical institutions) are fully protected with DANE, thus meeting the requirements of the German data protection authorities for high risk email communication.[16] All the other medical institutions must not exchange medical data via email, unless they employ further technical measures like E2EE.

Medical institutions do not need to operate their email servers themselves—they can make use of processors. However, what many forget: the medical institutions still stay the responsible controllers (in terms of the GDPR). If email providers are chosen that do not provide state-of-the-art security, medical institutions risk GDPR fines from data protection authorities.

We will regularly repeat our investigation in the future and provide the results at `https://www.mail-sicherheit.jetzt`

References

1. Alle Microsoft 365-Pläne vergleichen | Microsoft. `https://www.microsoft.com/de-de/microsoft-365/business/compare-all-microsoft-365-business-products?market=de`, date of last access: 14th June 2022

[14] Rather, those email accounts will soon be protected with DANE, as Microsoft announced a DANE roll-out for 2022: `https://techcommunity.microsoft.com/t5/exchange-team-blog/releasing-outbound-smtp-dane-with-dnssec/ba-p/3100920`

[15] It is a good sign that Microsoft started DANE roll-out for Exchange Online in early 2022: `https://techcommunity.microsoft.com/t5/exchange-team-blog/releasing-outbound-smtp-dane-with-dnssec/ba-p/3100920`

[16] However, 47 of these medical institutions employ a gmx or web.de address and it is not clear whether it is a freemail or a business account.

2. Change your Microsoft 365 email address to use your custom domain. https://docs.microsoft.com/en-us/microsoft-365/admin/email/change-email-address?view=o365-worldwide, date of last access: 14th June 2022
3. Cryptographic Message Syntax (CMS). https://datatracker.ietf.org/doc/html/rfc3369, first CMS RFC that mentions S/MIME, Date of last access: 15th Jun 2022
4. dig(1) - OpenBSD manual pages. https://man.openbsd.org/dig.1, date of last access: 18th June 2022
5. Google Online Security Blog: Google Public DNS Now Supports DNSSEC Validation. https://security.googleblog.com/2013/03/google-public-dns-now-supports-dnssec.html, date of last access: 19th June 2022
6. Google Public DNS. https://developers.google.com/speed/public-dns, date of last access: 15th Apr 2022
7. Mail Transfer Protocol. https://datatracker.ietf.org/doc/html/rfc772, original Mail Transfer Protocol RFC, Date of last access: 15th Jun 2022
8. Microsoft 365 Single kaufen – Premium-Office-Paket | Microsoft. https://www.microsoft.com/de-de/microsoft-365/p/microsoft-365-single/cfq7ttc0k5bf?rtc=1&activetab=pivot:overviewtab, date of last access: 14th June 2022
9. OpenINTEL. https://www.openintel.nl/, date of last access: 24th Mai 2022
10. PGP Message Exchange Formats. https://datatracker.ietf.org/doc/html/rfc1991, first PGP RFC, Date of last access: 15th Jun 2022
11. shuque/gotls: Diagnostic tool to perform DANE & PKIX authentication of a TLS server. https://github.com/shuque/gotls, date of last access: 18th June 2022
12. Simple Mail Transfer Protocol. https://datatracker.ietf.org/doc/html/rfc5321, date of last access: 18th Apr 2022
13. SMTP Service Extension for Authentication. https://datatracker.ietf.org/doc/html/rfc4954, sMTP Auth Extension RFC, Date of last access: 15th Jun 2022
14. SMTP Service Extension for Secure SMTP over TLS. https://datatracker.ietf.org/doc/html/rfc2487, original SMTP STARTTLS Extension RFC, Date of last access: 15th Jun 2022
15. SMTP Service Extension for Secure SMTP over Transport Layer Security. https://datatracker.ietf.org/doc/html/rfc3207, updated SMTP STARTTLS Extension RFC, Date of last access: 15th Jun 2022
16. SMTP Service Extensions for Transmission of Large and Binary MIME Messages. https://datatracker.ietf.org/doc/html/rfc1830, sMTP Chunking Extension RFC, Date of last access: 15th Jun 2022
17. Testing TLS/SSL encryption. https://testssl.sh/, date of last access: 15th Apr 2022
18. Braun, S., Oostveen, A.M.: Encryption for the masses? An analysis of PGP key usage. Mediatization Studies (2) (2018)
19. Dukhovni, V., Hardaker, W.: SMTP security via opportunistic DNS-based authentication of named entities (DANE) transport layer security (TLS). RFC 7672, IETF (2015)
20. Dukhovni, V., Hardaker, W.: The DNS-based authentication of named entities (DANE) protocol: Updates and operational guidance. Tech. rep. (2015)

21. Durumeric, Z., Adrian, D., Mirian, A., Kasten, J., Bursztein, E., Lidzborski, N., Thomas, K., Eranti, V., Bailey, M., Halderman, J.A.: Neither snow nor rain nor MITM... an empirical analysis of email delivery security. In: Proceedings of the 2015 Internet Measurement Conference. pp. 27–39 (2015)
22. Foster, I.D., Larson, J., Masich, M., Snoeren, A.C., Savage, S., Levchenko, K.: Security by any other name: On the effectiveness of provider based email security. In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security. pp. 450–464 (2015)
23. Kambourakis, G., Gil, G.D., Sanchez, I.: What email servers can tell to Johnny: an empirical study of provider-to-provider email security. IEEE Access **8**, 130066–130081 (2020)
24. Lee, H., Gireesh, A., van Rijswijk-Deij, R., Chung, T., et al.: A Longitudinal and Comprehensive Study of the DANE Ecosystem in Email. In: 29th USENIX Security Symposium (USENIX Security 20) (2020)
25. Petrlic, R.: The General Data Protection Regulation: From a Data Protection Authority's (Technical) Perspective. IEEE Security Privacy **17**(6), 31–36 (2019). https://doi.org/10.1109/MSEC.2019.2935701
26. Zhu, L., Wessels, D., Mankin, A., Heidemann, J.: Measuring DANE TLSA deployment. In: International Workshop on Traffic Monitoring and Analysis. pp. 219–232. Springer (2015)