Digital Security

Метаданные файлов:

Невидимая информация, раскрывающая вашу личность

Содержание

Введение

В Сети хранится и передается невообразимое количество данных, за которыми стоят метаданные, или цифровые следы. Каждый раз, когда вы добавляете кого-то в друзья в Facebook, слушаете рекомендуемую музыку в Spotify, публикуете твит, создаете тематические коллекции связанных изображений в Pinterest, метаданные работают в фоновом режиме и информируют о действиях своего создателя, помогая составить карту информационного ландшафта.

Каждый день в мессенджерах и по почте пользователи отправляют миллиарды файлов. Только через один WhatsApp отправляется около миллиарда фотографий ежедневно. При этом, лишь немногие задумываются о том, содержат ли отправляемые файлы цифровые следы: дату и время создания, имя автора, версию и технические характеристики программы, в которой был создан документ, местоположение и т.п.

Зачастую при создании нового документа цифровые следы остаются незамеченными (для рядового пользователя). Так или иначе, метаданные используются не только из добрых побуждений и могут представлять угрозу конфиденциальности.

В данном исследовании аналитики Digital Security помогут разобраться в понятиях и расскажут о том, как и какие именно метаданные можно найти в создаваемых вами файлах при их размещении в Сеть, какие опасности они таят, и что необходимо предпринять, чтобы обезопасить себя от утечки вашей личной информации.

Дисклеймер:

Данная работа носит аналитический характер. Цель — провести аналитический обзор по проблеме использования метаданных и узнать, как различные интернет-ресурсы относятся к метаданным передаваемых или загружаемых пользователями файлов. Авторы пользовались информацией из открытых источников.

Понятие «метаданные» и где их используют

Метаданные автоматически приписываются каждому файлу, помогают систематизировать и искать данные. Простая аналогия: чтобы получить в библиотеке конкретное произведение, необходимо сообщить библиотекарю автора и название. Непосредственно текст произведения — это сами данные, а метаданные, т.е. «данные о данных» — это имя автора, название произведения, количество страниц в книге, издательство и т.п.

Аналогично, посты в блогах имеют свои метаданные: заголовок, автор, дата публикации, теги и любая другая информация о публикации. Метаданные помогают установить лицензионные ограничения на распространение информации, указывая на автора контента.

Встроенные в веб-сайты метаданные включают в себя описание сайта, ключевые слова, метатеги — все, что имеет значение для поиска. Они идентифицируют контент и дают специальную характеристику, которая отличает его от другого контента. Идентификация метаданных включает множество технических элементов, таких как уникальный идентификатор ресурса (или URI), справочный номер файла или имя файла, заголовок и автора.

Сайты интернет-магазинов используют метаданные, которые хранят информацию о загрузках, рейтинги пользователей, данные пересылки, условия поиска, данные о ссылках и т.д. Digital-маркетологи следят за каждым кликом пользователя, сохраняя такую информацию, как тип устройства, местоположение, дату и время и другие данные, которые разрешено собирать по закону. Оперируя этой информацией, маркетологи создают картину распорядка дня и взаимодействий, предпочтений, привычек пользователей и используют ее в своих целях.

Провайдеры интернет-услуг, правительственные организации и другие лица, имеющие доступ к большим массивам метаданных, могут использовать метаданные с веб-страниц, электронных писем, телефонных звонков для мониторинга активности в интернете. Поскольку метаданные представляют собой выжимку из более крупных данных, эту информацию можно просматривать и фильтровать, чтобы получить сведения о миллионах пользователей.

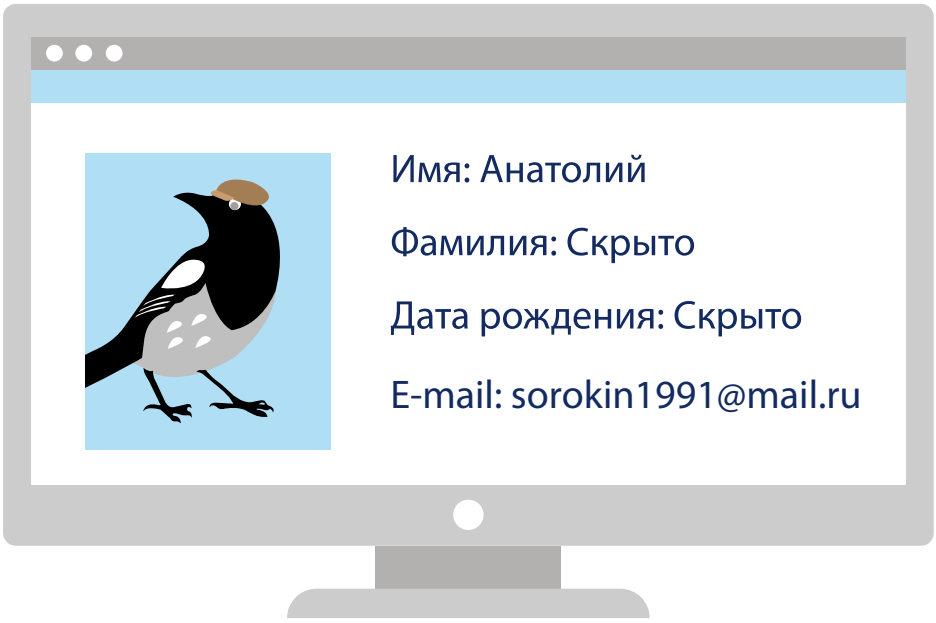
Злоумышленники могут использовать метаданные для осуществления вредоносных действий.

> ***Как правило, метаданные скрыты для пользователей. Каждое электронное письмо, которое вы отправляете или получаете, имеет ряд метаданных, часть из которых не видна в почтовом клиенте. Эти метаданные включают: тему, отправителя и адресата, дату и время отправки, имена отправляющего и принимающего серверов, формат сообщения — обычный текст или HTML, сведения о программном обеспечении для защиты от спама.***

Метаданные — угроза, о которой вы не знали?

Примечательно, что многие помнят о важности скрытия IP-адреса, знают способы его замены, но при этом даже не слышали о метаданных, которые могут предоставить больше личной информации, чем тот же IP-адрес.

Нет смысла передавать документы, фото или видео с включенным прокси, VPN и через TOR-браузер, если в метаданных в поле «Автор» указан «Анатолий Сорокин» и есть геометка с его точным адресом проживания. В таком случае ни о какой анонимности речи и быть не может.



Известно об их существовании или нет, метаданные способны повлиять на жизнь пользователей. Доказательством может служить небольшое исследование репортера ABC Уилла Окендена. Законы Австралии обязывают телефонные и интернет-компании хранить метаданные о звонках, сообщениях и случаях выхода в интернет в течение двух лет. Репортер запросил весь объем хранимых о нем метаданных и провел их анализ.

> *В России есть требование хранить не только метаданные, но и записи разговоров, текстовые сообщения, изображения, аудио- и видеозаписи, а также иные электронные сообщения пользователей в течение полугода.*

Объем метаданных за два года раскрывает адрес проживания, рабочий адрес, основные маршруты передвижения по городу и за его пределами, факты отсутствия по месту проживания или пребывания, а также основные контакты.

Ход исследования

Авторы решили изучить, как интернет-ресурсы, почтовые сервисы, мессенджеры и фотостоки ведут себя с метаданными различных файлов.

Для исследования понадобились фотографии, видео и текстовые документы формата DOCX и выборка из 16 различных сервисов, включая мессенджеры, облачные хранилища, фотостоки и почтовые сервисы. С помощью специального ресурса Jeffrey's Image Metadata Viewer исследователи проверили, какие метаданные есть в выбранных файлах.

Шаг 1

Загрузили файлы в Jeffrey's Image Metadata Viewer

Шаг 2

Получили метаданные файлов

Шаг 3

Передали файлы через почтовые сервисы, облака и мессенджеры, загрузили на фотостоки

Шаг 4

Скачали загруженные ранее файлы

Шаг 5

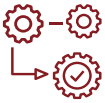
Проверили файлы в Jeffrey's Image Metadata Viewer

Шаг 6

Получили метаданные загруженных файлов

Шаг 7

Сравнили результаты, полученные на шагах 2 и 6



Шаг 8

Сделали выводы о взаимодействии сервисов с загружаемыми материалами

Далее приводим результаты исследования.

Фото

Метаданные фотоматериалов включают в себя следующую информацию:

- модель устройства, с помощью которого сделано фото, и его технические характеристики;
- размер фотографии;
- дата и время съемки;
- геолокация, если у камеры есть доступ к ней;
- автор фотографии;
- расстояние от камеры в момент съемки;
- теги и ключевые слова, описывающие содержание фото (создаются автором).

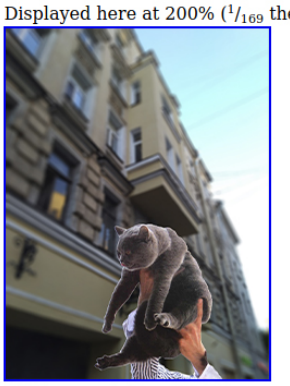
На скриншоте, сделанном в онлайн-сервисе, вы видите таблицу с данными о загруженной фотографии.

Basic Image Information

Target file: IMG_20200914_184539 (2).jpg

Caption:	Sweet cute fat cat Artist
Artist:	**Anatolii Sorokin**
Camera:	asus ZB602KL
Lens:	3.5 mm
Exposure:	Auto exposure, Not Defined, 1/50 sec, f/2.2, ISO 108
Flash:	Off, Did not fire
Keywords:	House, street, evening, old town.
Date:	**September 14, 2020** 6:45:39PM (timezone not specified) (6 days, 23 hours, 1 minute, 53 seconds ago, assuming image timezone of 3 hours ahead of GMT)
Location:	Latitude/longitude: **59°** 57' 13.4" North, **30°** 17' 34.1" East (59.953732, 30.292815) Map via embedded coordinates at: Google, Yahoo, WikiMapia, OpenStreetMap, Bing (also see the Google Maps pane below) Altitude: 32 meters (104 feet) Camera Pointing: South-southeast Timezone guess from earthtools.org: 3 hours ahead of GMT
File:	**3,120 × 4,160** JPEG (**13.0** megapixels) 6,645,100 bytes (6.3 megabytes)
Color Encoding:	Embedded color profile: "sRGB"

Extracted **120 × 160** 6.8-kilobyte "Photoshop:PhotoshopThumbnail" JPG
Displayed here at 200% (1/169 the area of the original)



Click image to isolate; click this text to show histogram

Раскрытие геолокации настораживает в первую очередь. Публикация фотографии в Сети с геометкой и комментарием "Едем в отпуск!" может подсказать злоумышленнику, в какой момент вас не будет дома, а также, что самое важное, где ваш дом находится.

Видео

Цифровые следы, которые остаются на отснятом материале:

- размер;
- формат;
- модель устройства, с помощью которого сделано видео;
- дата и время съемки;
- геолокация, если у камеры есть доступ к ней;
- автор видео;
- теги, описывающие содержание видеоматериала (создаются автором).

File — basic information derived from the file.

File Size	11 MB
File Type	MP4
File Type Extension	mp4
MIME Type	video/mp4

Composite
This block of data is computed based upon other items. Some of it may be wildly incorrect, especially if the image has been resized.

Avg Bitrate	16.3 Mbps
GPS Latitude	59.865600 degrees N
GPS Longitude	30.385000 degrees E
Megapixels	2.1
Rotation	90
GPS Position	59.865600 degrees N, 30.385000 degrees E

На сайте отобразилась геометка того места, где был отснят материал, как и в случае с фотографиями.

Текстовый документ в формате DOCX

Файлы документов MS Office (docx, xlsx и т.п) используют формат Office Open XML, который по факту является ZIP-архивом, содержащим в себе текстовые данные, медиа и прочие компоненты, снабженные разметкой в формате XML. Для текстовых файлов онлайн-ресурсы, как правило, раскрывают следующие метаданные:

- автор документа (если пользователь его указал);
- формат файла;
- название файла;
- даты создания и редактирования файла;
- размер файла.

ZIP

Ziped component #0	
ZipBitFlag	0x0006
ZipCRC	0x7102bc4a
ZipCompressedSize	365
ZipCompression	Deflated
ZipFileName	[Content_Types].xml
ZipModifyDate	**1980:01:01** 00:00:00 40 years, 8 months, 13 days, 16 hours, 30 minutes, 8 seconds ago
ZipRequiredVersion	20
ZipUncompressedSize	1,576

Вернуть отправителю

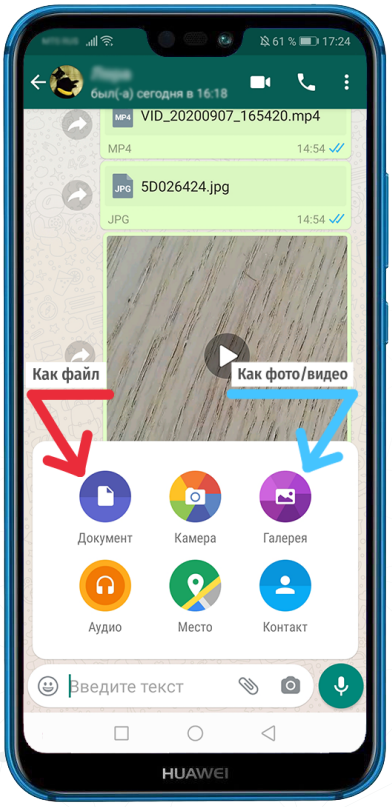
Над файлами было произведено множество действий: мы отправляли их в мессенджерах, соцсетях и почтовых сервисах, загружали на фотостоки и сервисы для просмотра видео и т.д. После этого снова проверили метаданные файлов через онлайн-ресурс и сравнили результаты.

Почтовые сервисы

Мы изучили различные почтовые сервисы на предмет сохранения в них метаданных и выяснили, что Gmail, Яндекс.Почта, Mail.ru, Protonmail передают файлы в целости и сохранности.

Почтовый сервис	Фото	Видео	Текстовый документ
Gmai	неизменно	неизменно	неизменно
Яндекс.Почта	неизменно	неизменно	неизменно
Mail.ru	неизменно	неизменно	неизменно
Protonmail	неизменно	неизменно	неизменно

Почтовые сервисы оставляют метаданные на месте и предоставляют полную информацию о документах.



Мессенджеры

Столь популярные сегодня мессенджеры нередко оказываются под прицелом как исследователей безопасности, так и злоумышленников. Вот одно из недавних сообщений о том, что они раскрывают персональные данные пользователей. В ходе данного исследования мы проверили, как те же распространенные мессенджеры обращаются с метаданными.

У каждого из пяти рассмотренных мессенджеров (Telegram, WhatsApp, Viber, Signal, Messenger) есть функция передачи вложения в виде файла со всеми сопутствующими данными.

Передача фото и видео через функцию “Камера” или с помощью галереи нередко убирает все метаданные. Впрочем, не без потери качества изображения.

Сводная таблица о сохранении мессенджерами метаданных при передаче вложения.

Мессенджер	Фото	Фото как файл	Видео	Видео как файл	Текстовый документ
Telegram	убираются полностью	неизменно	частично убираются	неизменно	неизменно
WhatsApp	убираются полностью	неизменно	убрана геометка	неизменно	неизменно
Viber	убираются полностью	неизменно	неизменно	неизменно	неизменно
Signal	убираются полностью	убираются полностью	убрана геометка	убрана геометка	неизменно
Messenger	остается только информация об авторе	остается только информация об авторе	неизменно	неизменно	неизменно

Как видно из таблицы, Telegram и Viber удаляют метаданные только с отправляемых изображений, но оставляют их неизменными при передаче фотографий и видео в форме файла. WhatsApp и Signal удаляют отметку о местоположении на отправленном видео. При отправке фотографий в Facebook Messenger приложение оставляет только информацию об авторе и удаляет остальные данные об изображении. Текстовые же документы сохраняют свои метаданные во всех рассмотренных мессенджерах.

Авторы также изучили метаданные фотографий, которые можно получить непосредственно в мессенджере через опцию “Камера”. При включенном доступе к геолокации как мессенджера, так и камеры метаданных обнаружить не удалось.

Фотостоки и облачные хранилища

Множество фотографов со всего мира загружают свои работы на фотостоки, чтобы поделиться ими бесплатно или же продать. Мы изучили те сайты, которые дают возможность скачать понравившуюся фотографию бесплатно: Pinterest, Pixabay, Unsplash, Pexels.

При скачивании изображений обнаружить метаданные не удалось ни на одном из попавших в данное исследование ресурсов.

Фотосток	Фото
Pinterest	убираются полностью
Pixabay	убираются полностью
Unsplash	убираются полностью
Pexels	убираются полностью

Далее мы проверили популярные облачные хранилища: Яндекс.Диск, Google Drive и Mega.nz.

Облачное хранилище	Фото
Яндекс.Диск	неизменно
Google Drive	неизменно
Mega.nz	неизменно

Как видно из таблицы, все исследованные облачные хранилища оставляют метаданные в целости и сохранности.

Выводы

Мы рассмотрели отношение различных ресурсов к метаданным файлов распространенных форматов JPG, MP4, DOCX. Резюмируем, что нам удалось выяснить.

При отправке изображений через почтовые сервисы или загрузке файлов в облачные хранилища метаданные остаются неизменными, как и качество отправляемых файлов. Рассмотренные нами бесплатные фотостоки убирают с фотографий метаданные, включая имя фотографа.

Популярные мессенджеры, выбранные для данного исследования, влияют на метаданные по-разному. Например, в Viber, Telegram, WhatsApp и Facebook Messenger сами файлы можно отправить без потери качества, при этом их метаданные не изменятся и не удалятся. В таком случае адресат или перехвативший сообщения злоумышленник с легкостью сможет узнать автора файла, дату создания и, если были отправлены фото- и видеоматериалы, посмотреть модель устройства, его технические характеристики и даже место, где была произведена съемка.

Следовательно, если данные о файлах могут являться конфиденциальной информацией, эксперты Digital Security рекомендуют удалять их перед загрузкой, так как с точки зрения злоумышленника, метаданные могут быть куда более полезны, чем сам файл.

Самый простой способ найти метаданные файлов — это посмотреть их свойства. Так, нажав правой кнопкой мышки на иконку файла, вы обнаружите раздел «Свойства», а в открывшемся окне отобразятся нужные данные. Также посмотреть метаданные можно в графических редакторах или с помощью специальных утилит. Другой способ — специализированные сайты, где можно загрузить файл или вставить ссылку на него и получить максимально подробное описание.

Изменить или удалить метаданные возможно через тот же раздел «Свойства». Для этого необходимо кликнуть на вкладку «Подробно» и отредактировать или удалить метаданные, нажав на ссылку «Удаление свойств и личной информации» и выбрав нужные пункты.

Некоторые сервисы имеют ограничения, например, на размер и вес загружаемых фотографий. Отредактировав или удалив метаданные, вы не только скроете личную информацию, но также уменьшите вес фотографий и, следовательно, ускорите загрузку страницы.

Заключение

Современные пользователи постоянно загружают и передают файлы через различные сервисы. При этом немногие обращают внимание на автоматически присвоенные каждому файлу метаданные, воспринимая их как нечто вторичное.

Метаданные, конечно, делают жизнь пользователя проще, например, они удобны для поиска информации. Там, где речь идет об авторском праве, метаданные тоже полезны. Фотографам важно наличие метаданных на сделанных ими фотографиях. На официальных музыкальных сервисах также — от наличия или отсутствия метаданных может зависеть, получат ли причастные к созданию композиции свой гонорар.

Однако те же метаданные могут стать и причиной неприятностей для пользователя. Возьмем отметку местоположения. С одной стороны, это крайне полезная функция современных девайсов: с ее помощью можно классифицировать фотографии, узнать погоду, поделиться местоположением, найти кафе или любое нужное место поблизости. С другой — доступ к вашему месту нахождения могут получить самые разные люди и организации. В лучшем случае ваша геолокация будет использована для настройки и показа таргетированной рекламы, в худшем — злоумышленники узнают больше о том, где вы живете и работаете.

В заключение повторим основную мысль: если вам не хочется сообщать о себе и своих файлах лишнюю информацию, удаляйте метаданные. Или хотя бы наиболее критичные из них — местоположение, для этого достаточно отключить геолокацию в настройках камеры.

Помните о том, что метаданные есть у любого файла, и они могут содержать конфиденциальную информацию. Следите за открытой о себе информацией и расскажите об этом тем, кто не знает. Будьте в безопасности!

Digital Security

Свяжитесь с нами

inbox@dsec.ru

7 (495) 223-07-86