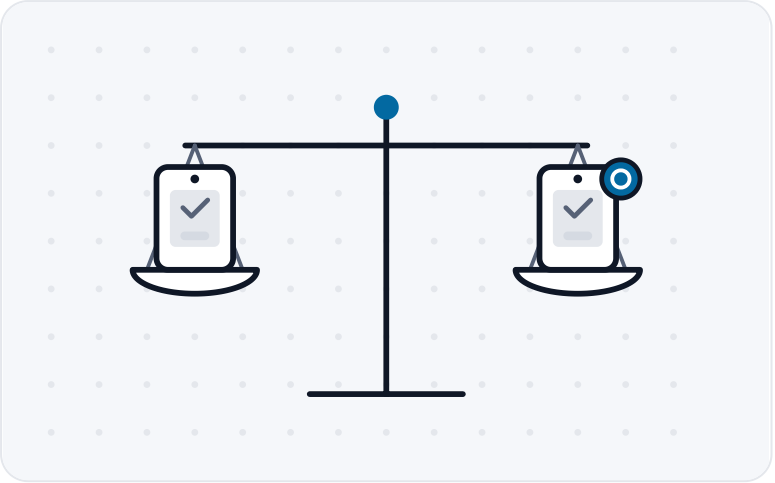
A-SIT Technology

E-Government Innovationszentrum

28. April 2023 Analyse · Mobile Sicherheit · Privatsphäre & Krypto · 23 Min. Lesezeit

Datenschutzvergleich zwischen kostenpflichtigen und kostenlosen Apps

Emina Ahmetovic



Autorinnen und Autoren	Emina Ahmetovic
Stand	2023-04-25
Quelle	https://technology2.a-sit.at/veroeffentlichungen/datenschutzvergleich-zwischen-kostenpflichtigen-und-kostenlosen-apps/
Hinweis	Maschinell übersetzt.

Kernaussagen

- Verglichen wurden 1620 Paare aus kostenloser und kostenpflichtiger Android-App; die Paare stammen vom selben Entwickler.
- Bei der Datenerhebung stimmen 76 % der Paare überein, bei der Datenweitergabe 78 % und bei Sicherheits- und Datenschutzpraktiken 79 %.
- Auch bei den Berechtigungen stimmen 76 % der Paare überein; kostenlose Apps deklarieren in den abweichenden Fällen mehr Berechtigungen.
- Wo die Praktiken abweichen, erheben und teilen kostenlose Apps bei einigen Datentypen mehr Daten, etwa Absturzprotokolle, Geräte-IDs und ungefähren Standort.

> - Nutzende könnten fälschlich annehmen, Bezahlen sorge für besseres Datenschutzverhalten; die Studie fand nur wenige Unterschiede.

Statistiken zeigen, dass 97 % der Android-Apps aus dem offiziellen Store kostenlos sind. Viele kostenlose Apps bieten auch Versionen zum Kauf an. Bei der Möglichkeit, zwischen kostenlosen oder kostenpflichtigen Versionen zu wählen, setzen User oft voraus, dass kostenpflichtige Apps einen besseren Datenschutz bieten.

In diesem Projekt wurden die Datenschutzfunktionen zwischen kostenlosen und kostenpflichtigen Apps evaluiert. Dazu wurde einen Datensatz der Top-Paid-Free App-Paare erstellt und Datenschutzmerkmale wie Berechtigungen und Datenschutzerklärung geprüft. Schließlich wurde auf Basis dieser Erkenntnisse analysiert, inwieweit die kostenpflichtigen Apps die angenommenen Datenschutzerwartungen erfüllen.

1. Einleitung

Kostenlose Android-Anwendungen dominieren den Play Store. Die Verteilung kostenpflichtiger und kostenloser Apps fiel 2022 mit 97 Prozent zugunsten der kostenlosen Apps aus[1]. Da nur 3 Prozent der Anwendungen im Store käuflich zu erwerben sind, wird sich dieser Trend in den kommenden Jahren voraussichtlich kaum ändern. Dennoch bieten mobile Märkte auch weitere Preismodelle, die die kostenlose und die kostenpflichtige Version ergänzen. Insbesondere Freemium- und Paidmium-Modelle gehören zu den beliebten App-Preismodellen, zwischen denen Endnutzende wählen können. Während kostenlose Apps für Nutzende gratis sind und keine In-App-Zahlungen voraussetzen, bezeichnet Freemium Apps, die kostenlos heruntergeladen werden können, aber dennoch In-App-Käufe anbieten. Ebenso verlangen kostenpflichtige Apps beim Herunterladen eine einmalige Zahlung und bieten keine In-App-Käufe an, während Paidmium kostenpflichtige Apps sind, die zusätzlich In-App-Käufe für weitere Einnahmen anbieten. Die Vielfalt der Preismodelle gibt Nutzenden die Möglichkeit, die Apps nach ihren Vorlieben und Bedürfnissen anzupassen. Sie erlaubt es Entwickelnden aber auch, ein Geschäftsmodell aufzubauen und Gewinn zu erzielen. Tatsächlich liegen die für 2022 prognostizierten Einnahmen im App-Markt bei 5,25 Mrd. US-Dollar für kostenpflichtige Apps, verglichen mit 204,90 Mrd. US-Dollar für Einnahmen aus In-App-Käufen[2].

1.1. Verständnis der Nutzenden von kostenpflichtigen Apps

Um Nutzende zu gewinnen, entscheiden sich viele Anwendungsentwickelnde für den Vertrieb kostenloser Versionen. Um einen Gewinn sicherzustellen, wählen Entwickelnde jedoch in der Regel ein Einnahmemodell oder eine Kombination mehrerer Modelle. Damit stellen sich Fragen zu den Datenschutzunterschieden zwischen kostenpflichtigen und kostenlosen Apps. Das Monetarisierungsmodell der Apps und die hohe Verbreitung kostenloser Apps werfen die Frage auf, ob kostenlose Apps auf Kosten der Privatsphäre gehen, ob also Privatsphäre gegen kostenlose Apps eingetauscht wird. Andererseits wird auch diskutiert, ob kostenpflichtige Apps die Privatsphäre besser schützen. Einige Studien zum Verständnis der Nutzenden von kostenlosen und kostenpflichtigen Apps und deren Datenschutzmerkmalen legen nahe, dass die verbreitete Überzeugung besteht, das Bezahlen einer App bedeute, dass die Daten der Nutzenden weder weitergegeben noch erhoben werden (Han C, 2020 Jan;2020(3)). Diese Überzeugung ist stark mit

kostenpflichtigen Apps verbunden, die mit Werbefreiheit, „keine Werbung“ oder ähnlichen Schlagwörtern werben, die Nutzende als Version mit besseren datenschutzfreundlichen Techniken interpretieren. Frühere Forschung, die auf statischer und dynamischer Analyse der Apps beruht, hat jedoch gezeigt, dass dies in der Praxis irreführend ist.

1.2. Beitrag

Diese Studie beleuchtet die Datenschutzunterschiede zwischen kostenpflichtigen und kostenlosen Apps. Zwar ist das Bezahlen von Apps durch die zusätzlichen Funktionen und Merkmale gerechtfertigt, die dadurch freigeschaltet werden, wir konzentrieren uns jedoch auf die verbreitete Überzeugung, dass sich von der kostenlosen und der kostenpflichtigen Version einer App die letztere durch besseres Datenschutzverhalten auszeichnet. Unser Ziel war es, Datenschutzmerkmale zu untersuchen und die folgende Forschungsfrage zu beantworten: Was sind die Datenschutzunterschiede zwischen der kostenlosen App und ihrem kostenpflichtigen Gegenstück? Damit wollen wir feststellen, ob das Bezahlen für Apps andere Datenschutzpraktiken bedeutet. Als Teil unseres Beitrags haben wir einen Datensatz mit datenschutzbezogenen Daten erstellt, die im Play Store für die 1620 Paare aus kostenlosen und kostenpflichtigen Apps verfügbar sind. Wir nutzen diese Daten, um datenschutzbezogene Merkmale wie Datensicherheit und Berechtigungen zu untersuchen und zu vergleichen. Abschließend ziehen wir auf Grundlage der Evaluierungsergebnisse unser Fazit. Die Ergebnisse unserer Analyse zeigen eine große Überschneidung zwischen kostenlosen und kostenpflichtigen Apps hinsichtlich der Datensicherheit. In den folgenden Abschnitten erläutern wir unsere Ergebnisse genauer.

1.3. Gliederung

Der Rest der Studie ist wie folgt aufgebaut. In Abschnitt 2 geben wir den Hintergrund zu den Datenschutzfunktionen in Android-Apps. In Abschnitt 3 stellen wir unsere Methodik vor. In Abschnitt 4 legen wir die Evaluierungsergebnisse dar. In Abschnitt 5 diskutieren wir verwandte Arbeiten. In Abschnitt 6 formulieren wir unser Fazit.

2. Hintergrund

In diesem Abschnitt geben wir einen kurzen Überblick über den Berechtigungsmechanismus in Android, den Abschnitt zur Datensicherheit, verschiedene Analysen von Anwendungen sowie die Monetarisierungsmodelle.

2.1. Berechtigungen unter Android

Android-Nutzenden wird eine große Anzahl mobiler Apps angeboten, die im Alltag nützliche Funktionen bereitstellen. Da die Apps jedoch in die Privatsphäre der Nutzenden eingreifen können, wurde das Berechtigungsmodell (Developers., 2023) als Schutzmechanismus eingeführt, der Nutzenden Kontrolle über ihre Daten gibt. Je nach Umfang der eingeschränkten Aktionen und dem Zugriff auf eingeschränkte Ressourcen definiert Android die Schutzstufen „normal“ und „gefährlich“.

Die normale Schutzstufe wird den Installationszeit-Berechtigungen (Developers, Install-time permissions, 2023) zugewiesen, wie etwa INTERNET, die sich durch begrenzten Zugriff auf eingeschränkte Ressourcen auszeichnen. Da sie Aktionen mit minimalen Auswirkungen auf andere Apps oder das System ausführen, werden sie vom System automatisch zum Installationszeitpunkt erteilt (dem Zeitpunkt, zu dem Nutzende

die App installieren). Die Installationszeit-Berechtigungen werden Nutzenden standardmäßig nicht angezeigt, sondern sind in den Details des Play Store zu finden. Sie können nicht widerrufen werden, und es gibt keine Prüfungen zur Laufzeit. Sie werden außerdem allen Nutzenden auf dem Gerät erteilt.

Dagegen werden Laufzeitberechtigungen (Developers, Runtime Permissions, 2023), wie SMS, CONTACTS, CALENDAR, LOCATION und ähnliche, der gefährlichen Schutzstufe zugeordnet, da sie Zugriff auf eingeschränkte Daten erlauben und der App die Ausführung eingeschränkter Aktionen ermöglichen. Sie werden zur Laufzeit angefordert (unmittelbar vor der eingeschränkten Aktion) und benötigen die ausdrückliche Zustimmung der Nutzenden. Nutzende sollten sich bewusst sein, dass Laufzeitberechtigungen auf private Daten zugreifen können, die möglicherweise sensible Informationen enthalten. Im Gegensatz zu den Installationszeit-Berechtigungen können Nutzende Laufzeitberechtigungen jederzeit widerrufen.

2.2. Analyse von Android-Anwendungen

Die frühere Forschung, die Datenschutzunterschiede zwischen kostenpflichtigen und kostenlosen Apps aufdecken will, hat Anwendungsanalysen durchgeführt, die in der Regel statisch, dynamisch oder eine Kombination aus beiden waren.

Die dynamische Anwendungsanalyse ist ein Verfahren zur Erkennung und Analyse bösartigen Verhaltens einer Anwendung, während diese läuft. Da die Überwachung und Verfolgung privater Daten in Echtzeit erfolgt, wird diese Methode nicht durch Obfuskation, Verschlüsselung oder ähnliche Faktoren beeinträchtigt. Dennoch kann die dynamische Analyse keine Datenschutzverletzungen erkennen, die zur Laufzeit nicht ausgelöst werden. Eine weitere Einschränkung ist der hohe Ressourcenverbrauch durch den Echtzeitbetrieb, der diese Verfahren auf manchen Mobilgeräten weniger praktikabel macht (Kang, 2021).

Im Gegensatz zur dynamischen Analyse erfolgt die statische Analyse, bevor eine App ausgeführt wird. Die statische Analyse ist ein Verfahren, bei dem anwendungsspezifische Merkmale wie Berechtigungen und API-Aufrufe durch Analyse des Quellcodes oder seiner Binärdarstellung extrahiert werden. Da die statische Analyse auf der Ebene des Quellcodes arbeitet, wird die Erkennung von Schadsoftware durch manche Umgehungstechniken beeinträchtigt, etwa durch Obfuskation des Codes oder dynamisches Nachladen von Code. Eine weitere Einschränkung der statischen Analyse ist, dass sie eine größere Zahl an Fehlalarmen erzeugen kann. Ein beliebtes Open-Source-Werkzeug für die statische Analyse unter Android ist FlowDroid (Arzt, 2014).

Taint Analysis ist eine Technik der Informationsflussanalyse, mit der der Fluss sensibler Informationen von definierten Quellen zu definierten Senken verfolgt wird. Quellen sind Ressourcen, die Nutzende schützen wollen, etwa der Standort oder die Telefonnummer, während Senken die Stellen sind, an denen die Ressourcen das Gerät verlassen könnten, etwa Methoden, die mit der Übertragung der Daten zusammenhängen. Die Taint Analysis zielt darauf ab, Datenlecks zu erkennen, indem sie den Informationsfluss zwischen Quellen und Senken beobachtet und prüft, ob Daten aus Quellen eine Senke erreichen (Zhang, 2021). Sie kann statisch oder dynamisch erfolgen und wird meist zur Erkennung von Datenschutzverletzungen eingesetzt. Ein weit verbreitetes Werkzeug für statische Taint Analysis in Android-Apps ist TaintDroid (Enck, 2014).

2.3. Datensicherheit unter Android

Um die Transparenz zu erhöhen und Nutzenden zu helfen zu verstehen, wie Apps mit ihren Daten umgehen, nutzen Entwickelnde den Abschnitt zur Datensicherheit, um die erforderlichen Informationen auszufüllen und Nutzenden anzuzeigen. Der Zweck der Datensicherheit unter Android ist es offenzulegen, welche Daten erhoben und weitergegeben werden, und mehr Informationen zu Sicherheits- und Datenschutzpraktiken in der App bereitzustellen. So erhalten Nutzende genügend Informationen, um eine fundierte Entscheidung über die von ihnen verwendeten Apps zu treffen. Im Abschnitt zur Datensicherheit werden die folgenden Daten offengelegt:

- Datenerhebung, also die Übertragung von Daten von einem Gerät durch die App, sei es durch Bibliotheken oder SDKs, durch das Erheben von Daten aus einer innerhalb der App geöffneten Webansicht oder durch übertragene Daten, die ephemer verarbeitet wurden, sowie Nutzerdaten, die pseudonym erhoben wurden. Daten, die nur lokal verarbeitet und nicht vom Gerät gesendet werden, zählen ebenso wenig zur Datenerhebung wie die Ende-zu-Ende-Verschlüsselung.

- Datenweitergabe, also die Weitergabe erhobener Daten aus der App an Dritte. Die Datenweitergabe umfasst die folgenden Fälle: außerhalb des Geräts, auf dem Gerät, aus den App-Bibliotheken und SDKs sowie aus der von der App geöffneten Webansicht. Die folgenden Fälle gelten nicht als Weitergabe: die Übertragung vollständig anonymisierter Daten, die Übertragung von Daten an Dienstleister, die Übertragung von Daten für rechtliche Zwecke oder die Übertragung von Daten mit Einwilligung der Nutzenden.

- Datenverarbeitung, bei der Entwickelnde offenlegen können, welche erhobenen Datentypen optional und welche erforderlich sind. Im optionalen Fall haben Nutzende Kontrolle über die Datenerhebung und können entscheiden, im Gegensatz zum erforderlichen Fall, in dem die erhobenen Daten für die Funktion der App benötigt werden.

- Datenschutz- und Sicherheitspraktiken, bei denen Entwickelnde die Möglichkeit haben offenzulegen, ob ihre App bei der Datenübertragung Verschlüsselung verwendet oder ob sie einen Löschmechanismus bietet, mit dem Nutzende die Löschung ihrer Daten verlangen können.

- Familienrichtlinie. Anwendungen, die sich an Kinder als Zielgruppe richten, müssen die Anforderungen der Families-Richtlinie von Google Play erfüllen[3]. Die Information, ob die App der Familienrichtlinie entspricht, wird im Abschnitt zur Datensicherheit angezeigt.

- Unabhängige Sicherheitsprüfung. Wurde die Anwendung gegen einen globalen Sicherheitsstandard validiert, können Entwickelnde diese Information ergänzend zum Abschnitt zur Datensicherheit hinzufügen.

Es gibt verschiedene Typen und Kategorien von Daten, die von der App erhoben oder weitergegeben werden. Außerdem müssen Entwickelnde den Zweck[4] der Datenverarbeitung angeben, etwa App-Funktionalität, Analyse, Kommunikation mit Entwickelnden, Werbung oder Marketing, Betrugsprävention, Sicherheit und Compliance, Personalisierung sowie Kontoverwaltung. Datentypen und ihre Kategorien sind in Tabelle 1 aufgeführt.

Kategorie	Datentyp
Standort	Ungefährer Standort
Standort	Genauer Standort
Personenbezogene Daten	Name
Personenbezogene Daten	E-Mail-Adresse
Personenbezogene Daten	Nutzer-IDs
Personenbezogene Daten	Adresse
Personenbezogene Daten	Telefonnummer
Personenbezogene Daten	Ethnische Herkunft
	Politische oder religiöse Überzeugungen
	Sexuelle Orientierung
	Sonstige Informationen
Finanzinformationen	Zahlungsinformationen der Nutzenden
Finanzinformationen	Kaufverlauf
Finanzinformationen	Bonität
Finanzinformationen	Sonstige Finanzinformationen
Gesundheit und Fitness	Gesundheitsinformationen
Gesundheit und Fitness	Fitnessinformationen
Nachrichten	E-Mails
Nachrichten	SMS oder MMS
Nachrichten	Sonstige In-App-Nachrichten
Fotos und Videos	Fotos
Fotos und Videos	Videos
Audiodateien	Sprach- oder Tonaufnahmen
Audiodateien	Musikdateien
Audiodateien	Sonstige Audiodateien
Dateien und Dokumente	Dateien und Dokumente
Kalender	Kalendereregnisse

Kategorie	Datentyp
Kontakte	Kontakte
App-Aktivitäten	App-Interaktionen
App-Aktivitäten	Suchverlauf in der App
App-Aktivitäten	Installierte Apps
App-Aktivitäten	Sonstige nutzergenerierte Inhalte
App-Aktivitäten	Sonstige Aktionen
Webbrowsing	Webbrowserverlauf
App-Informationen und Leistung	Absturzprotokolle
App-Informationen und Leistung	Diagnose
App-Informationen und Leistung	Sonstige Daten zur App-Leistung
Geräte- oder andere IDs	Geräte- oder andere IDs

Tabelle 1. Die Tabelle zeigt Datentypen und ihre Kategorien. Details zu den Datentypen und ihre Beschreibungen sind auf der offiziellen Website[5] aufgeführt.

2.4. Datenschutzerklärung unter Android

Um besser zu beurteilen, welche Apps sie herunterladen, können Nutzende in der Datenschutzerklärung der App weitere Informationen zum Datenschutz erhalten. Dieses Feld enthält den Link zur Erklärung, die von Nutzenden heruntergeladen werden kann. Diese optionale Information befindet sich im Detailabschnitt von Google Play und wird von Entwickelnden verfasst, die ihre Datenschutzbestimmungen mit den Nutzenden teilen möchten[6].

2.5. Überblick über die Monetarisierungsmodelle

Ein Monetarisierungsmodell ist ein Geschäftsmodell, das zur Erzielung von Einnahmen aus einem Produkt, einer Dienstleistung, einem Werk, geistigem Eigentum oder Ähnlichem verwendet wird[7]. Es bezeichnet den Prozess, in dem ein Gegenstand ohne Einnahmen Einkommen erzeugen kann[8].

Es gibt viele verschiedene Monetarisierungsmodelle[9], mit denen kostenlose Apps Gewinn erzielen können. Diese Geschäftsmöglichkeiten umfassen häufig:

- Werbung bezeichnet das Modell, bei dem Entwickelnde in ihrer Anwendung Werbung Dritter einblenden, um Einnahmen zu erzielen. Das ist auch eines der beliebtesten Modelle, da es vielseitig ist und in die meisten Apps passt, die Apps aber gleichzeitig für Endnutzende kostenlos bleiben können[10].

- Affiliate-Marketing bezeichnet ein der Werbung ähnliches Modell; hier entstehen die Einnahmen jedoch in Form einer Provision jedes Mal, wenn Nutzende die App herunterladen oder definierte Vorgänge in der App ausführen.
- In-App-Käufe bezeichnen eines der beliebtesten Monetarisierungsmodelle, bei dem Nutzende zusätzliche Funktionen der App oder Waren kaufen können. Das ist zum Beispiel oft bei Spiele-Apps der Fall, in denen Nutzende für das nächste Level zahlen können, und allgemein dort, wo Käufe zusätzliche Funktionen der App freischalten.
- Kostenpflichtige Abonnements sind ein beliebtes Modell bei Software-as-a-Service-Apps, bei denen zu Beginn eine kostenlose Testphase angeboten wird, auf die das kostenpflichtige Abonnement folgt, das einen höheren Umfang an Inhalten bietet. So nutzen zum Beispiel viele bekannte Zeitungen das Abonnementmodell.
- Freemium ist ein Modell, das Premium- und kostenlose Apps verbindet; die für Nutzende verfügbare kostenlose Version ist üblicherweise mit Basisfunktionen verbunden. Für zusätzliche Funktionen, Inhalte oder Möglichkeiten müssen Nutzende jedoch zahlen. Für dieses Modell ist es wichtig, dass die Kernfunktionalität der App kostenlos bleibt.
- Sponsoring ist der Werbung ähnlich, beruht jedoch auf einem einzelnen Werbetreibenden, der in irgendeiner Weise mit dem Zweck der App verbunden ist.
- Crowdfunding ist ein bei Start-ups beliebtes Modell, bei dem Finanzierungsplattformen die App durch Geldspenden unterstützen können. Zu den beliebten Crowdfunding-Plattformen gehören Kickstarter[11], Indiegogo[12], CrowdFunder[13], AppFunder[14].
- Transaktionserlöse und Ähnliches. Wenn zu den Funktionen der App Geldtransaktionen gehören, kann dies genutzt werden, um eine Gebühr auf die Transaktionen zu erheben. Das ist bei Hotel- und Flugbuchungs-Apps verbreitet.

3. Methodik

Einer der ersten Schritte war die Erstellung eines Datensatzes. Um die Datenschutzmerkmale der kostenlosen und der kostenpflichtigen Apps zu vergleichen, erstellen wir einen Datensatz, der aus datenschutzbezogenen Metadaten der Apps besteht. Wir haben einen Ansatz gewählt, bei dem wir kostenlose Apps mit den entsprechenden kostenpflichtigen Paaren vergleichen. Das bedeutet, dass wir für jede kostenlose App versuchen, das kostenpflichtige Gegenstück desselben Entwicklers zu finden. Da wir keine Apps kaufen, verwendet unsere Analyse nur Daten, die im Play Store verfügbar sind. Für diesen Zweck haben wir die beliebtesten appIDs aus dem Google Play Store ausgelesen, die den Rang TOP FREE haben. Anschließend haben wir die Suche auf jene eingegrenzt, die nur auf Englisch verfügbar sind, und erhielten rund 50 000 appIDs. Um zu identifizieren, ob und welche Apps kostenpflichtige Gegenstücke haben, suchen wir nach Informationen zu den App-Entwickelnden. Nachdem wir die developerID jeder App gefunden hatten, suchten wir kostenpflichtige Apps dieses Entwickelnden. Nicht alle kostenlosen Apps haben kostenpflichtige Versionen, und da kostenlose und kostenpflichtige Versionen derselben Entwickelnden im Play Store nicht wirklich miteinander verknüpft sind, haben wir uns für Algorithmen in Python entschieden, die die Ähnlichkeit zwischen zwei Zeichenketten messen. Um herauszufinden, welche kostenpflichtige appID das Gegenstück zur kostenlosen appID wäre, haben wir einige Ansätze verglichen

und uns für FuzzyWuzzy[15] entschieden, eine Python-Bibliothek, die die Levenshtein-Distanz[16] verwendet, um die Unterschiede zwischen Sequenzen und Mustern mit einem Ähnlichkeitswert von größer oder gleich 80 % zu berechnen. Das folgende Code-Snippet zeigt einen Screenshot der Paare mit dem jeweiligen Wert.


```
[{
"paidApp": "lysesoft.andftppro", "freeApp": "lysesoft.andftp", "score": 91
"paidApp": "air.LecturaMusicalPracticaPRO", "freeApp":
"air.LecturaMusicalPractica", "score": 95
"paidApp": "com.atypicalgames.radiationisland", "freeApp":
"com.atypicalgames.radiationislandfree", "score": 94
"paidApp": "com.jumobile.manager.systemapp.pro", "freeApp":
"com.jumobile.manager.systemapp", "score": 94
"paidApp": "pro.cryptotab.android", "freeApp": "max.cryptotab.android", "score": 86
"paidApp": "com.shenyaocn.android.usbdualcamerapro", "freeApp":
"com.shenyaocn.android.usbdualcamera", "score": 96 }]
```


Nach der manuellen Prüfung des erstellten Datensatzes stellten wir fest, dass die Paare in den meisten Fällen korrekt identifiziert worden waren. In einigen Fällen jedoch, etwa bei der Zuordnung von “paidApp”: “pro.cryptotab.android”, als Gegenstück zu “freeApp”: “max.cryptotab.android”, stellten wir fest, dass das Verfahren weiter verfeinert werden sollte. Um unsere Zuordnungstechniken zu optimieren, haben wir eine zusätzliche Bedingung eingeführt: Die Unterschiede zwischen zwei Zeichenketten dürfen nur die folgenden Schlüsselwörter enthalten, etwa ‘full’, ‘pro’, ’lite’, ‘free’, ‘paid’, ‘ads’, ’light’ und ähnliche.

Am Ende unseres Vorgehens hatten wir 1 640 Paare gesammelt, mit den für unsere Datenschutzanalyse relevanten Daten, zu denen Datensicherheit, Berechtigungen, Datenschutzerklärung, Kategorie und Ähnliches gehören. Ein Beispiel für ein solches Paar ist im folgenden Code-Snippet zu sehen:

```
[{
"paidApp": { "genreID": "TOOLS", "securityPractices": "[]", "score": "3.9",
"sharedData": "[]", "ratings": "1043", "permissions": ["Phone", "Device ID & call
information"], "appID": "lysesoft.andftppro", "collectedData": "[]", "description":
" 'AndFTP Pro unlocks advanced features for AndFTP application. AndFTP is a FTP,
FTPS, SCP and SFTP client. It provides commands to rename, delete, set permissions
on remote files and folders. It can upload or download files and folders
recursively. It supports RSA and DSA keys for SSH. You need AndFTP free installed.
Features in Pro version are SCP support, folder synchronization, custom commands
and import settings from file.&lt;br&gt;&lt;br&gt;Pro version acts as an unlock
key, it does not have any icon and you cannot open it. Once installed it unlocks
all features of the free application. You can check it by running the free
application, then Menu-&gt;Options-&gt;Advanced and you should see &quot;License:
Pro&quot;.',\n ", "title": "AndFTPPro", "privacy policy url":
"http://www.lysesoft.com/about/privacy.html" },
"freeApp": { "genreID": "TOOLS", "securityPractices": "[]", "score": "3.6",
"sharedData": "[]", "ratings": "31196", "permissions": ["Photos/Media/Files",
"Storage"], "appID": "lysesoft.andftp", "collectedData": "[]", "description": "
'AndFTP is a FTP, FTPS, SCP, SFTP client. It can manage several FTP configurations.
It comes with both device and FTP file browser. It provides download, upload,
synchronization and share features with resume support. It can open (local/remote),
rename, delete, update permissions (chmod), run custom commands and more. SSH
RSA/DSA keys support. Share from gallery is available. Intents are available for
third party applications. Folder synchronization are available in Pro version
only.',\n ", "title": "AndFTP (your FTP client)", "privacy policy url":
"http://www.lysesoft.com/about/privacy.html" }]
```


Die erfassten Apps gehören zu 45 Kategorien, und die Verteilung der Apps nach Kategorien ist in der folgenden Tabelle dargestellt:

SHOPPING	5	AUTO_AND_VEHICLES	33
ENTERTAINMENT	0	NEWS_AND_MAGAZINES	4
SPORTS	20	MEDICAL	23
VIDEO_PLAYERS	49	HOUSE_AND_HOME	3
SOCIAL	7	COMICS	2
PHOTOGRAPHY	55	GAME_PUZZLE	40
ART_AND_DESIGN	12	GAME_BOARD	11
FINANCE	28	GAME_EDUCATIONAL	38
COMMUNICATION	31	PARENTING	2
BOOKS_AND_REFERENCE	41	GAME_ARCADE	42
MAPS_AND_NAVIGATION	24	GAME_CASUAL	25
HEALTH_AND_FITNESS	54	GAME_SIMULATION	55
TOOLS	246	LIBRARIES_AND_DEMO	4
WEATHER	28	GAME_WORD	11
GAME_CASINO	5	GAME_TRIVIA	5
BUSINESS	25	GAME_RACING	12
EVENTS	1	GAME_ADVENTURE	44
MUSIC_AND_AUDIO	115	GAME_SPORTS	6
LIFESTYLE	32	GAME_ACTION	41
TRAVEL_AND_LOCAL	17	GAME_MUSIC	6
GAME_STRATEGY	52	PERSONALIZATION	46
EDUCATION	114	PRODUCTIVITY	118
GAME_ROLE_PLAYING	34	GAME_CARD	27

Tabelle 2. Diese Tabelle zeigt die Verteilung der Apps nach ihrer Kategorie.

4. Evaluierung

Wir haben die Daten aus unserem Datensatz genutzt, um einen Datenschutzvergleich zwischen Paaren aus kostenpflichtigen und kostenlosen Apps durchzuführen. Statt nur die Datenschutzpraktiken aller kostenlosen und kostenpflichtigen Apps zu analysieren, haben wir einen Ansatz gewählt, bei dem wir Free-

Paid-Paare von Apps gefunden haben und unsere Analyse auf diesen Paaren aufbauen. Dieser Ansatz liefert zuverlässigere Ergebnisse; das Finden von Paaren hat jedoch den Nachteil, dass die Zahl der Apps deutlich sinkt, weil Paare nicht korrekt identifiziert werden konnten oder weil nicht jede kostenlose App ein Gegenstück hat.

Wir gliedern unsere Evaluierung in vier Teile: erhobene Daten, weitergegebene Daten, Sicherheits- und Datenschutzpraktiken sowie Berechtigungen. Die Gesamtevaluierung zeigt für die genannten Kategorien erhebliche Überschneidungen zwischen den Praktiken kostenpflichtiger und kostenloser Apps, wie im folgenden Diagramm zu sehen ist. In den einzelnen Abschnitten erläutern wir unsere Ergebnisse genauer.

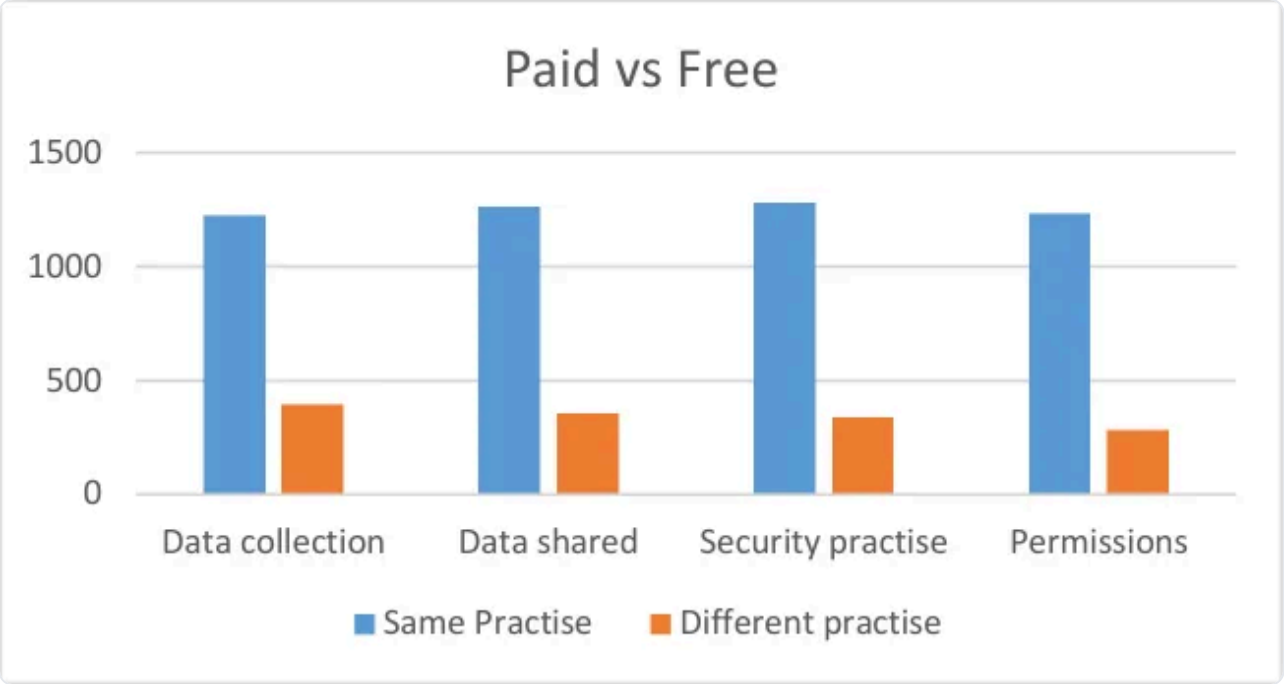


Abbildung 1. Das Diagramm zeigt die Unterschiede zwischen kostenpflichtigen und kostenlosen Apps in vier Kategorien: Datenerhebung, Datenweitergabe, Sicherheitspraktiken und Berechtigungen.

4.1. Erhobene Daten

Wir vergleichen den Abschnitt zur Datenerhebung der kostenlosen App und ihres kostenpflichtigen Gegenstücks. Insgesamt haben wir 1620 Paare analysiert. Unsere Studie zeigt, dass 76 % (1225) bei der kostenlosen und der kostenpflichtigen Version dieselbe Datenerhebungspraxis haben. Wir definieren als dieselbe Erhebungspraxis jene, bei der kostenlose und kostenpflichtige Apps denselben Datentyp derselben Kategorie für denselben Zweck erheben. Die verbleibenden 24 % (395) haben bei der Datenerhebung unterschiedliche Praktiken. Wir untersuchen weiter, worin diese Unterschiede bestehen. Die Ergebnisse sind in der folgenden Tabelle 3 zusammengefasst. Die erste Spalte, Datenerhebungstyp, ist der Typ der erhobenen Daten. Die zweite Spalte, Kostenpflichtige App, ist die Anzahl der Vorkommen dieses Datentyps in kostenpflichtigen Apps, und die dritte Spalte, Kostenlose App, gibt die Anzahl der Datenerhebungen in kostenlosen Apps an. Die Ergebnisse zeigen deutlich, dass kostenlose Apps bei einigen Typen im Allgemeinen mehr Daten erheben als kostenpflichtige Apps, etwa Absturzprotokolle, Geräte- oder andere IDs, Diagnose, App-Interaktionen, ungefährer Standort, E-Mail-Adresse, sonstige Daten zur App-Leistung und Nutzer-IDs. Bei den übrigen Typen sind die Ergebnisse recht ähnlich.

Datenerhebungstyp	Kostenpflichtige App	Kostenlose App
Absturzprotokolle	150	250
Geräte- oder andere IDs	143	312
Diagnose	121	216
App-Interaktionen	119	220
Ungefährer Standort	63	136
E-Mail-Adresse	57	61
Sonstige Daten zur App-Leistung	56	97
Nutzer-IDs	42	57
Sonstige Aktionen	37	63
Genauer Standort	32	34
Name	29	26
Sonstige nutzergenerierte Inhalte	15	15
Fotos	13	14
Kaufverlauf	11	25
Sonstige Informationen	7	5
Sonstige In-App-Nachrichten	7	6
Installierte Apps	6	24
Suchverlauf in der App	5	5
Sprach- oder Tonaufnahmen	3	4
Gesundheitsinformationen	3	3
Dateien und Dokumente	3	5
Zahlungsinformationen der Nutzenden	2	1
Sonstige Finanzinformationen	2	2
Videos	2	3
Telefonnummer	2	4
Fitnessinformationen	2	7
Ethnische Herkunft	1	1

Datenerhebungstyp	Kostenpflichtige App	Kostenlose App
Sexuelle Orientierung	1	1
Kalenderereignisse	1	1
Kontakte	1	2
SMS oder MMS	1	1
Adresse	1	0
Musikdateien	0	1
Sonstige Audiodateien	0	1
Webbrowserverlauf	0	1

Tabelle 3. Die Tabelle zeigt, welche Datentypen sowohl von kostenpflichtigen als auch von kostenlosen Apps erhoben werden und wie oft die erhobenen Daten jeweils in kostenpflichtigen und kostenlosen Apps vorkommen.

4.2. Weitergegebene Daten

Wir vergleichen den Abschnitt zu weitergegebenen Daten der kostenlosen App und ihres kostenpflichtigen Gegenstücks. Insgesamt haben wir 1620 Paare analysiert. Unsere Studie zeigt, dass 78 % (1263) bei der kostenlosen und der kostenpflichtigen Version dieselbe Datenweitergabepraxis haben. Wir definieren als dieselbe Weitergabepraxis jene, bei der kostenlose und kostenpflichtige Apps denselben Datentyp derselben Kategorie für denselben Zweck weitergeben. Die verbleibenden 22 % (357) haben bei der Datenweitergabe unterschiedliche Praktiken. Wir untersuchen weiter, worin diese Unterschiede bestehen. Die Ergebnisse sind in der folgenden Tabelle 4 zusammengefasst. Die erste Spalte, Datenweitergabetyp, ist der Typ der weitergegebenen Daten. Die zweite Spalte, Kostenpflichtige App, ist die Anzahl der Vorkommen dieses Datentyps in kostenpflichtigen Apps, und die dritte Spalte, Kostenlose App, gibt die Anzahl der Datenweitergaben in kostenlosen Apps an.

Datenweitergabetyp	Kostenpflichtige App	Kostenlose App
Absturzprotokolle	92	300
Geräte- oder andere IDs	87	177
Diagnose	79	174
App-Interaktionen	75	169
Ungefährer Standort	44	156
E-Mail-Adresse	7	5
Sonstige Daten zur App-Leistung	23	55
Nutzer-IDs	12	25
Sonstige Aktionen	23	39
Genauer Standort	15	42
Name	2	0
Sonstige nutzergenerierte Inhalte	4	2
Fotos	5	3
Kaufverlauf	6	16
Sonstige Informationen	0	5
Sonstige In-App-Nachrichten	1	1
Installierte Apps	3	8
Suchverlauf in der App	5	7
Sprach- oder Tonaufnahmen	3	3
Gesundheitsinformationen	0	0
Dateien und Dokumente	3	2
Zahlungsinformationen der Nutzenden	3	1
Sonstige Finanzinformationen	1	1
Videos	1	1
Telefonnummer	2	0
Fitnessinformationen	0	2
Ethnische Herkunft	1	0

Datenweitergabetyp	Kostenpflichtige App	Kostenlose App
Adresse	1	0
Webbrowserverlauf	0	4

Tabelle 4. Die Tabelle zeigt, welche Datentypen sowohl von kostenpflichtigen als auch von kostenlosen Apps erhoben werden und wie oft die erhobenen Daten jeweils in kostenpflichtigen und kostenlosen Apps vorkommen.

Tabelle 4 zeigt die Unterschiede zwischen den Datenweitergabepraktiken kostenloser und kostenpflichtiger Apps. Ähnlich wie bei der Datenerhebung unterscheiden sie sich vor allem durch eine deutlich größere Zahl an Datenweitergaben in kostenlosen Apps bei einigen Typen, etwa Absturzprotokolle, Geräte- oder andere IDs, Diagnose, App-Interaktionen und ungefährer Standort. Andererseits sind die Unterschiede zwischen kostenlosen und kostenpflichtigen Apps bei weniger verbreiteten Datentypen vernachlässigbar.

4.3. Sicherheitspraktiken

Wir vergleichen die Sicherheits- und Datenschutzpraktiken der kostenlosen App und ihres kostenpflichtigen Gegenstücks. Insgesamt haben wir 1620 Paare analysiert. Unsere Studie zeigt, dass 79 % (1279) bei der kostenlosen und der kostenpflichtigen Version dieselbe Sicherheits- und Datenschutzpraxis haben. Die verbleibenden 21 % (341) haben bei der Datenweitergabe unterschiedliche Praktiken. Wir untersuchen weiter, worin diese Unterschiede bestehen, und erhalten die folgenden Ergebnisse:

Sicherheits- und Datenschutzpraktiken	Kostenpflichtige Apps	Kostenlose Apps
Daten werden bei der Übertragung verschlüsselt	52	256
Daten können nicht gelöscht werden	43	184
Die Löschung der Daten kann beantragt werden	29	102
Daten werden nicht verschlüsselt	20	30
Verpflichtung zur Einhaltung der Play- Familienrichtlinie	29	18
Unabhängige Sicherheitsprüfung	0	1

Tabelle 5. Die Tabelle zeigt, welche Sicherheits- und Datenschutzpraktiken sowohl kostenpflichtige als auch kostenlose Apps haben und wie oft sie jeweils in kostenpflichtigen und kostenlosen Apps vorkommen.

Ähnliche Ergebnisse zeigen sich auch bei den Sicherheits- und Datenschutzpraktiken. Einige Merkmale, die in kostenpflichtigen Apps häufiger vorkommen als in kostenlosen, sind jedoch die Verpflichtung zur Einhaltung der Play-Familienrichtlinie.

4.4. Berechtigungen

Wir vergleichen die Berechtigungen der kostenlosen App und ihres kostenpflichtigen Gegenstücks. Insgesamt haben wir 1620 Paare analysiert. Unsere Studie zeigt, dass 76 % (1235) bei der kostenlosen und der kostenpflichtigen Version dieselben Berechtigungen haben. Die verbleibenden 24 % (385) haben bei der Datenweitergabe unterschiedliche Praktiken. Wir untersuchen weiter, worin diese Unterschiede bestehen, und die erhaltenen Ergebnisse zeigen, ähnlich wie im vorigen Abschnitt, dass kostenlose Apps mehr Berechtigungen deklarieren im Vergleich zu kostenpflichtigen Apps.

Berechtigungen	Kostenpflichtige Apps	Kostenlose Apps
WLAN-Verbindungsinformationen	30	138
Speicher	68	96
Fotos/Medien/Dateien	68	96
Telefon	22	57
Geräte-ID und Anrufinformationen	19	55
Standort	17	46
Kontakte	19	33
Identität	17	27
Geräte- und App-Verlauf	6	19
Mikrofon	5	10
Kamera	19	14
Kalender	5	7
Wearable-Sensoren/Aktivitätsdaten	1	3

Tabelle 6. Die Tabelle zeigt, welche Arten von Berechtigungen sowohl kostenpflichtigen als auch kostenlosen Apps zugeordnet sind und wie oft diese Berechtigungen jeweils in kostenpflichtigen und kostenlosen Apps vorkommen.

5. Verwandte Arbeiten

In einer Studie von Han et al. (Han C, 2020 Jan;2020(3)) wurde die Erwartung der Nutzenden an kostenlose und kostenpflichtige Apps untersucht. Die Autoren befragten 998 Teilnehmende, und die Ergebnisse legen stark nahe, dass die Teilnehmenden erwarten, dass kostenpflichtige Apps bessere Datenschutz- und Sicherheitspraktiken bieten als ihre kostenlosen Gegenstücke. Die Studie hat gezeigt, dass die meisten Nutzenden Werbung als Hinweis darauf sehen, dass die App in gewissem Maß Daten verfolgt und erhebt, dass sie aber auch annehmen, dass das Fehlen von Werbung Schutz für ihre persönlichen Daten bedeutet. In ihrer Studie wendeten sie statische und dynamische Analyse an. Sie stellten fest, dass 45 % der

kostenpflichtigen Versionen dieselben Bibliotheken von Drittanbietern verwendeten wie die kostenlosen Versionen und dass 74 % dieselben gefährlichen Berechtigungen hatten wie die kostenlose App. Außerdem zeigten sie, dass 32 % der kostenpflichtigen Apps dasselbe Datenerhebungs- und Übertragungsverhalten aufweisen wie ihre kostenlosen Gegenstücke. Da wir die kostenpflichtigen Apps in unserer Studie nicht kaufen und deshalb statische oder dynamische Analysen ausschließen, bauen wir unsere Studie auf den Metadaten auf, die mit der App bereitgestellt werden. Da Android-Anwendungen zentral vertrieben werden, steht Nutzenden auch eine Vielzahl an Metadaten zur Verfügung, etwa Datensicherheit der App, Datenschutzerklärung, Beschreibungen, Bewertungen, Berechtigungen und Ähnliches. Im Gegensatz zu früheren Arbeiten konzentrieren wir unsere Studie hauptsächlich auf den Unterschied bei der Datensicherheit der Apps.

6. Fazit

Aus den Ergebnissen des Vergleichs der im Play Store verfügbaren Datenschutzmerkmale lässt sich Folgendes zusammenfassen:

- 76 % der Datenerhebungspraktiken sind bei kostenlosen und kostenpflichtigen Apps gleich
- 78 % der Datenweitergabepraktiken sind bei der kostenlosen App und ihrem kostenpflichtigen Gegenstück gleich
- 79 % der Sicherheits- und Datenschutzpraktiken sind bei der kostenlosen App und ihrem kostenpflichtigen Gegenstück gleich
- 76 % der Berechtigungen sind bei der kostenlosen App und ihrem kostenpflichtigen Gegenstück gleich

Angesichts dieser Zahlen lässt sich schließen, dass Nutzende fälschlicherweise annehmen könnten, das Bezahlen für Apps sorge für besseres Datenschutzverhalten. Wie auch andere frühere Arbeiten argumentieren wir, dass unsere Studie zwischen kostenlosen und kostenpflichtigen Apps nur sehr wenige Unterschiede bei den genannten Datenschutzbereichen gezeigt hat.

Diese Studie hat jedoch ihre Grenzen. Erstens berücksichtigt sie nur kostenlose und kostenpflichtige Modelle; andere Preismodelle liegen außerhalb des Umfangs dieser Studie. Zweitens verlassen wir uns auf den Zuordnungsprozess, um Paare zu finden, im Gegensatz zu anderen Studien, in denen diese Merkmale manuell geprüft wurden, indem diese Aufgabe ausgelagert wurde. Außerdem konnten wir keine statische oder dynamische Analyse durchführen, da wir keine kostenpflichtigen Apps gekauft haben. Diese Studie zeigt das Potenzial, die Unterschiede zwischen kostenpflichtigen und kostenlosen Apps weiter zu untersuchen, da sie einige interessante Ergebnisse geliefert hat, die den Überzeugungen oder Erwartungen der Nutzenden widersprechen könnten. Ein Ansatz wäre eine klassische Analyse, die auch eine Codeanalyse umfassen könnte, um mehr Informationen über die Unterschiede im Verhalten zu erhalten.

Literaturverzeichnis

Arzt, S. a. (2014). Flowdroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps. *Acm Sigplan Notices*, 49(6):259–269.

Developers, G. (2023). Install-time permissions. pp. https://developer.android.com/guide/topics/permissions/overview#install-time.

Developers, G. (2023). Runtime Permissions. p. https://developer.android.com/guide/topics/permissions/overview#runtime.

Developers., G. (2023). Permissions on Android. p. https://developer.android.com/guide/topics/permissions/overview.

Enck, W. a.-G. (2014). Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems (TOCS)*, 32(2):1–29.

Han C, R. I.-R. (2020 Jan;2020(3)). The price is (not) right: Comparing privacy in free and paid apps. *Proceedings on Privacy Enhancing Technologies.*

Kang, H. a. (2021). {A modified flowdroid based on chi-square test of permissions. *Entropy*, 23(2):174.

Zhang, J. a. (2021). Analyzing android taint analysis tools: FlowDroid, Amandroid, and DroidSafe. *IEEE Transactions on Software Engineering* .

---

1. https://www.statista.com/statistics/266211/distribution-of-free-and-paid-android-apps/ ↩
2. https://www.statista.com/outlook/dmo/app/worldwide ↩
3. https://support.google.com/googleplay/android-developer/answer/9893335?hl=en ↩
4. https://support.google.com/googleplay/android-developer/answer/10787469?hl=en#zippy=%2Cdatacollection%2Cpurposes ↩
5. https://support.google.com/googleplay/androiddeveloper/answer/10787469?hl=en#zippy=%2Cpurposes%2Cdata-types ↩
6. https://support.google.com/googleplay/answer/2666094?hl=en ↩
7. https://www.mightynetworks.com/encyclopedia/monetization-model ↩
8. https://www.investopedia.com/terms/m/monetize.asp ↩
9. https://theappsolutions.com/blog/marketing/free-app-money/ ↩
10. https://www.bornfight.com/blog/how-to-make-money-with-your-mobile-app-top-10-app-monetizationmodels/ ↩
11. http://www.kickstarter.com/ ↩
12. http://www.indiegogo.com/ ↩
13. https://www.crowdfunder.co.uk/ ↩
14. https://appsfunder.guru/ ↩
15. https://github.com/seatgeek/fuzzywuzzy/blob/master/fuzzywuzzy/fuzz.py#L236 ↩
16. https://people.cs.pitt.edu/~kirk/cs1501/Pruhs/Spring2006/assignments/editdistance/Levenshtein%20Distance.htm ↩