Sustainability of Digital Formats: Planning for Library of Congress Collections

Android Package

>> Back

Table of Contents

Format Description Properties

- ID: fdd000592
- Short name: APK
- Content categories: aggregate
- Format Category: file-format
- Other facets: container-bundle, binary, structured, symbolic
- Last significant FDD update: 2024-04-09
- Draft status: Preliminary

Identification and description

Full name	Android Package
Description	An Android package (APK) is an archive file that contains the contents of an Android app required at runtime. Their primary usage is in Android-based devices that rely on Android package files to install applications. Android package files (APKs) comprise the contents of an Android app. APKs are based on Java's JAR format structure and packaged as a ZIP archive. The Android Developer's Guide states that Android applications are primarily written in Kotlin, the Java programming language, and C++ languages. The Android SDK tools compile code, data and resource files into an APK or an Android App Bundle (AAB) format. A single APK file contains all the necessary files that are required for its installation and execution. An APK file, when extracted as a ZIP, contains the following files and folders, according to the Android Developer's Guide: • assets/:A directory that contains applications assets. • lib/: A directory that contains compiled code prepared for specific hardware platforms.

	• META-INF/: A directory that contains the manifest file, signature, and a list of resources in the archive. Metadata about the application is stored here. • res/: A directory that contains any additional, non-compiled resources such as images. • AndroidManifest.xml: A file that contains the name, versioning information and contents of the package. • classes.dex: Compiled Java classes. The extension is short for "Dalvik executable format." This format is defined in the Android documentation. • resources.arsc: A file containing all compiled resource files. Within the META-INF/ folder are a few files that ensure package integrity. This includes MANIFEST.MF, the application certificate titled CERT.RSA, and a Signature File, CERT.SF, which contains a list of resources and a SHA-1 checksum for each line in the MANIFEST.MF file. Files, values and structures may be slightly different depending on their signature version. This structure is derived from the JAR specification. The AndroidManifest.xml file may be either regular XML or binary XML, but likely binary. This requires advanced tooling for people to access, parse, and read. Apktool, apkanalyzer, and Androguard are programs that make reading this file possible. The manifest file describes essential information about your app to the Android build tools, the Android operating system, and Google Play. Google provides a sample of this file. Researcher Johan van der Knijff provides a sample of extracted binary XML.
Production phase	A final-state format used for distribution of Android applications.
Relationship to other formats	
Extension of	JAR. Users can perform read and write operations on an APK file using Java APIs java.util.zip and java.util.jar". Not described separately at this time.
Extension of	ZIP, ZIP File Format (PKWARE)
Contains	XML, XML (Extensible Markup Language)
Contains	Dalvik executable format (DEX). Not described separately at this time.
Affinity to	Android App Bundle (AAB). The AAB format creates APK files. According to Google, An AAB "contains the contents of an Android app project, including some additional metadata that isn't required at runtime. An AAB is a publishing format and can't be installed on Android devices. It defers APK generation and signing to a later stage." Not described separately at this time.
Has extension	APKS. Files with the APKS extension contain a collection of APK files. See description in the Android Developer's Guide's bundletool documentation. Not described separately at this time.
Has extension	XAPK. Compressed Android Package. According to FileInfo.com, XAPK is similar to the APK format but may contain additional assets used by the app. Not described separately at this time.
Has extension	APKM. Android App Bundle Mirror. According to FileInfo.com, APKM is "an Android app bundle created for use with APKMirror Installer, an alternative Android app installer." Not described separately at this time.

Local use

LC experience or existing holdings	The Library of Congress has a small amount of APK files in its collections.
LC preference	The Library of Congress Recommended Formats Statement (RFS) lists

	APK as a preferred format for software and gaming distribution packages.

Sustainability factors

Disclosure	Limited. Ownership of the format belongs, de facto, to the Android Open Source Project, with influence from Google. The only known published documentation is via Google's Android Developer's Guide. Comments welcome.
Documentation	Some guidance is available on Google's Android Developer's Guide under "Application fundamentals".
Adoption	APK is used to distribute application software on Android operating systems.
Licensing and patents	The Android operating system is developed by the Open Handset Alliance consortium, with its primary and most widely used version developed by Google. The core operating system, the Android Open Source Project is licensed primarily under the free and open-source Apache License 2.0. The Android name and logo are trademarked by Google. There is no specific license or patent for the APK format. Comments welcome.
Transparency	Limited. Depends upon algorithms and tools to read; will require sophistication to build tools.
Self-documentation	APK's AndroidManifest.xml file holds metadata and self-documentation about technical dependencies, including versions, hardware or software requirements, permissions, and any additional arbitrary metadata. The meta-data element can hold any additional, arbitrary metadata. The uses-sdk element contains minSdkVersion and targetSdkVersion, which define the minimum and target levels of the application. The uses-feature element is used to declare hardware or software features used by the application. The uses-library element details any shared library dependencies. The Android Developer's Guide outlines in further detail and includes a sample manifest file. The amount of documentation depends on what is provided by the developer. Additional details are explained in Johan van der Knijff's blog post "Towards a preservation workflow for mobile apps". **Accessibility Features** Specific support for accessibility is unknown. Although the focus is on Windows applications, Microsoft's Accessibility (Windows Installer) page has some helpful information about installer packages in general, stating "the user interface of an installer package should facilitate accessibility of the application or product to all users." These include tips about "tooltip text is contained in the Help column of the Control table" to be shown by screen readers, "Controls in dialog boxes should be linked using the Control_Next field of the Control table. The controls need to be authored such that they can all be reached by using the TAB key." and "Shortcut keys should be provided for gaining access to controls directly." Andriod Developers Build accessible apps has information for application packages but not specifically APK installers. See also ZIP. Comments welcome.
External dependencies	The British Library asserts that through their research, they were not able to identify tools that could unpackage and migrate APKs for use in non-Android environments. Only application developers or publishers are able to update applications for compatibility with new versions of the operating system. They note that the Android SDK allows for development and testing on a PC, allowing developers to test on a variety of device configurations and versions without requiring the physical device.

	The Google Play store requires that a compressed APK must be less than 100MB.
Technical protection considerations	APKs have the concept of Application Signing: "Application signing allows developers to identify the author of the application and to update their application without creating complicated interfaces and permissions. Every application that is run on the Android platform must be signed by the developer. Applications that attempt to install without being signed will be rejected by either Google Play or the package installer on the Android device." APKs require Authentication: "Android uses the concept of user-authentication-gated cryptographic keys that requires the following components: 1. Cryptographic key storage and service provider. 2 User authenticators." Android 5.0 up to Android 9 support full-disk encryption. Android 7.0 and later supports file-based encryption. This is expanded in Android 9, which introduced support for metadata encryption. Other security issues described on the Android Security Features page from the Android Open Source Project.

Quality and functionality factors

Aggregate	
Compression	See ZIP.
Support for Error Dectection	See ZIP.
Beyond normal functionality	None.

File type signifiers and format identifiers

Tag	Value	Note
Filename extension	apk	See Google's Android for Developers guide. Other extensions may be used, such as those listed on Wikipedia and Wikidata. Comments welcome.
Internet Media Type	application/vnd.android.package-archive	Not listed in IANA. See Apache Tika and Wikidata.
Magic numbers	See related format.	See ZIP.
Pronom PUID	See note.	PRONOM has no corresponding entry as of March 2024.
Wikidata Title ID	Q596391	See https://www.wikidata.org/wiki/Q596391.

Notes

General	According to reporting from the Verge, after August 2021, Google required that new Play apps will have to be published using the Android App Bundle (AAB) format. Devices that support Android packages (APKs) will still be able to download and install this format. The Verge reporting also notes that there are potential preservation issues around difficulties "playing back" APKs on a phone due to "sideloading" (installing an application without using the specified hardware device) and installing

	from "unknown sources" (sources not directly vetted by the Google Play store). Comments welcome. APKs may be "split." Google's Android Developer's Guide explains "Split" APKs. These files are similar to standard APKs, but the Android platform is able to treat multiple installed split APKs as a single application. Android 4.4 and lower do not support split APKs. There have been at least 17 major versions of the Android operating system since its first release in 2008. Early versions were limited to mobile phone devices. The British Library notes the "rapid rate of development poses a challenge for apps with specific operating system dependencies, as these apps may not function as intended once a device's underlying operating system is updated."
History	

Format specifications

- See Disclosure under Sustainability factors.

Useful references

URLs

- "Application fundamentals". Google. Last updated October 10, 2023. (https://developer.android.com/guide/components/fundamentals).
- "The Android App Bundle format". Google. Last updated December 6, 2021. (https://developer.android.com/guide/app-bundle/app-bundle-format).
- "Dalvik executable format". Android Open Source Project. Last updated February 6, 2024. (https://source.android.com/docs/core/runtime/dex-format).
- "Application Signing". Android Open Source Project. Last updated February 7, 2024. (https://source.android.com/docs/security/features/apksigning).
- "RSA (cryptosystem)". Wikipedia. (https://en.wikipedia.org/wiki/RSA_(cryptosystem)).
- "Android APK files & signature types (part 2)". Vlad Iftimie. March 3, 2020. (https://medium.com/@vlad.iftimie88/some-points-on-android-apk-files-part-2-8ad522ad002c).
- "JAR File Specification". Oracle. (https://docs.oracle.com/javase/7/docs/technotes/guides/jar/jar.html).
- "APK". ArchiveTeam. (http://fileformats.archiveteam.org/wiki/APK).
- "Apktool". (https://apktool.org/).
- "apkanalyzer". Google. Last updated April 12, 2023. (https://developer.android.com/tools/apkanalyzer).
- "Androguard". (https://github.com/androguard/androguard).
- "bundletool". Google. (https://developer.android.com/tools/bundletool).
- "XAPK". FileInfo.com. (https://fileinfo.com/extension/xapk).
- "APKM". FileInfo.com. (https://fileinfo.com/extension/apkm).
- "App manifest overview". Google. Last updated February 8, 2024. (https://developer.android.com/guide/topics/manifest/manifest-intro).
- "arize-androidManifest.xml". Johan van der Knijff. February 17, 2021. (https://github.com/KBNLresearch/mobile-apps/blob/main/sample-files/arize-androidManifest.xml). Sample AndroidManifest.xml file.
- "Android". Open Handset Alliance. (https://www.openhandsetalliance.com/android_overview.html).
- "AOSP frequently asked questions (FAQ)". Android Open Source Project. Last updated February 1, 2024. (https://source.android.com/docs/setup/about/faqs).
- "Brand guidelines". Android Open Source Project. Last updated February 8, 2024. (https://source.android.com/docs/setup/about/brands).
- "<meta-data>". Android Open Source Project. Last updated March 29, 2023. (https://developer.android.com/guide/topics/manifest/meta-data-element).
- "<uses-sdk>". Android Open Source Project. Last updated February 16, 2024. (https://developer.android.com/guide/topics/manifest/uses-sdk-element).
- "<uses-feature>". Android Open Source Project. Last updated December 15, 2023. (https://developer.android.com/guide/topics/manifest/uses-feature-element).

- "<uses-library>". Android Open Source Project. Last updated March 29, 2023. (https://developer.android.com/guide/topics/manifest/uses-library-element).
- "Considerations on the Acquisition and Preservation of Mobile eBook Apps". 16th International Conference on Digital Preservation (iPRES). British Library. M. Pennock, et al. September 25, 2019. (https://doi.org/10.5281/zenodo.3460450).
- "APK Expansion Files". Google. Last updated January 3, 2024. (https://developer.android.com/google/play/expansion-files).
- "Application Signing". Android Open Source Project. Last updated February 7, 2024. (https://source.android.com/docs/security/features/apksigning).
- "Authentication". Android Open Source Project. Last updated February 7, 2024. (https://source.android.com/docs/security/features/authentication).
- "Encryption". Android Open Source Project. Last updated February 7, 2024. (https://source.android.com/docs/security/features/encryption).
- "Android Security Features". Android Open Source Project. Last updated February 7, 2024. (https://source.android.com/docs/security/features).
- "APK (file format)". Wikipedia. (https://en.wikipedia.org/wiki/Apk_(file_format)).
- "tika-mimetypes.xml". Apache Tika. (https://github.com/apache/tika/blob/main/tika-core/src/main/resources/org/apache/tika/mime/tika-mimetypes.xml#L350).
- "Google is moving away from APKs on the Play Store". Jay Peters. June 30, 2021. (https://www.theverge.com/2021/6/30/22557390/google-apk-app-bundles-package-format-play-store).
- "Windows 11 will be able to sideload Android apps". Mitchell Clark. June 25, 2021. (https://www.theverge.com/2021/6/25/22550689/windows-11-android-app-sideload-amazon-app-store).
- "Sideloading and Unknown Sources on Android: How to do it and fix errors". Jerry Hildenbrand. April 16, 2020. (https://www.androidcentral.com/unknown-sources).
- "Analyze your build with the APK Analyzer". Google. Last updated January 3, 2024. (https://developer.android.com/studio/debug/apk-analyzer).
- "The Structure of Android Package (APK) Files". Google. 2008. (https://web.archive.org/web/20111224184753/http://sites.google.com/site/io/inside-the-android-application-framework). Available via Internet Archive.
- "Four Android emulators, two apps". Johan van der Knijff. February 9, 2021. (https://www.bitsgalore.org/2021/02/09/four-android-emulators-two-apps).
- "Running archived Android apps on a PC: first impressions". Johan van der Knijff. 23 October 23, 2014. (https://openpreservation.org/blogs/running-archived-android-apps-pc-first-impressions/).
- "Towards a preservation workflow for mobile apps". Johan van der Knijff. February 24, 2021. (https://www.bitsgalore.org/2021/02/24/towards-a-preservation-workflow-for-mobile-apps).
- "APK - What is an APK file?" FileFormat.com. (https://docs.fileformat.com/compression/apk/).
- Wikidata entry for Q596391 (https://www.wikidata.org/wiki/Q596391). Information in Wikidata about APK. Wikidata Title ID: Q596391.

Last Updated: 05/17/2024

Digital Preservation Home | Digital Formats Home