AI Mobile Device Requirements Specification

Version 1.0

23 February 2021

This Industry Specification is a Non-binding Permanent Reference Document of the GSMA

Security Classification: Non-confidential

Copyright Notice

Disclaimer

Antitrust Notice

This GSMA Permanent Reference Document (PRD) is classified by GSMA as an Industry Specification, as such it has been developed and is maintained by GSMA in accordance with the provisions set out GSMA AA.35 - Procedures for Industry Specifications.

Table of Contents

1 Introduction

1.1 Purpose

This specification enables the mobile industry to design, develop, and test an Artificial Intelligence (AI) Mobile Device.

This specification defines the normative baseline for an AI Mobile Device covering use-cases, applications, requirements and technology, whilst also taking into account security and privacy aspects, to accelerate the deployment of AI technology across the industry for Mobile Network Operators, devices and component manufacturers.

This specification contains normative and informative sections. Unless otherwise specified, all sections are normative.

The explanation and background information for this specification is available in the GSMA AI Mobile Device Guidelines Study Report 2018 **Error! Reference source not found.**.

1.2 Consideration of Security and Privacy in AI Implementations

As an emerging and powerful domain of technology, AI can be used for incredibly beneficial purposes but has the potential to cause harm (whether intentionally or negligently). Principles are being established that the human should be in 'command' and 'control' of such functionality. Whilst an 'AI Mobile Device' represents one element of an overall implementation of AI, it is important that these factors are considered in any implementation. Implementers should adopt a 'Secure by Design' and 'Privacy by Design and by Default' approach. Primarily this means that functionality built into future devices is safe from the start from the User's perspective, based on the principle that the User can enable such functionality if they want to. Enabling the User to have a choice is a core principle of the security and privacy requirements within this document. Implementers of this specification are also invited to consider the broader ethical implications of how they integrate such functionality into User devices and the functionality itself, alongside the legislative and regulatory requirements in each country and jurisdiction such devices are sold into.

1.3 Scope

The scope of this specification is to define AI Mobile Device requirements. The AI Mobile Device in this version specifically refers to an AI mobile phone and tablet. Other types of mobile devices like IoT and wearable items may be considered in future releases.

1.4 Definition of Terms

Term	Description
AI Agent	A software entity which has the characteristic to learn autonomously to perform a task by trial and error, with minimal guidance from the User. Reference: http://ica2019.crowdscience.org/
AI Application	A core functionality of the device enhanced by AI computation or a third party application that applies Machine Learning (of which deep learning and reinforcement learning are specific examples).

Term	Description
AI Model	Mathematical algorithms that are trained using data and human expert input to replicate a decision an expert would make when provided that same information.
Biometric Data	Computer representations of human characteristics typically used for authentication and authorisation purposes. Examples include a face, fingerprint or voice. This data may be interpreted through an algorithm and stored in another form such as a template of the biometric.
Data Processor	The data processor processes personal data. Applies in this document to data processed through AI algorithms.
Deep Learning	Deep Learning is one type of Machine Learning.
Facial Photo Enhancement	An application that can do one or more of the following: remove spots, reduce wrinkles, reshape facial features (such as lips, nose, cheeks, ears etc.), remove dark circles, and alter skin tone when taking selfies.
Human in Command	As defined in the EU Commissions document 'Ethics Guidelines for Trustworthy AI' **Error! Reference source not found.**.
Inference	The process in which a trained machine learning model identifies patterns and makes predictions or decisions based on given inputs.
Machine Learning	Is the application of computer algorithms that automate improvement through experience. It is a type of artificial intelligence based on the concept that model built by algorithm can learn from data (known as training data), identify patterns and make decisions (known as Inferencing) with minimal human intervention.
Native API	APIs provided by the device manufacturer for access to AI hardware (e.g., NPU, CPU, GPU and DSP).
Native Application	An application that is pre-installed by the device manufacturer.
Operations	In the specific context of OPS, Operations only refers to multiply-accumulate (MAC) operations, not including input, output and other operations, and typically 1 MAC operation = 2 Deep Learning operations; the number of MACs needed to compute an inference on a single image is a common metric to measure the efficiency of the model. The widths of the integer matrix multiplication vary by architecture, dedicated hardware and supported topologies. Any claimed TOPS number depends on several assumptions such as frequency, number of MACs and various other hardware specifications.
Operations Per Second / Per Watt (OPS/w)	OPS per watt extend that measurement to describe performance efficiency.
Personal Data	Personal data are any information which are related to an identified or identifiable natural person. Encrypted Personal Data is considered as Personal Data. Anonymized Personal Data is not considered as Personal Data.
Privacy by Default	The practice during system design and deployment of ensuring that privacy is protected throughout the entire lifecycle of personal data. This includes that: personal data's collection, processing, period of storage and accessibility are minimised to that which is necessary for the identified purpose, and rights of data subjects are strongly respected.

Term	Description
Privacy by Design	The concept of taking privacy into account throughout the entire process of building a system or solution, by which privacy is protected throughout the entire lifecycle of personal data and the rights of data subjects are strongly respected. The user preferences and decisions are respected.
Secure by Default	The practice during system design and deployment of ensuring minimal exposure to security risk, by actions such as the minimisation of system and data exposure within software and hardware or across interfaces and ensuring that the system is preconfigured with secure settings.
Secure by Design	The concept of building software and hardware from the foundation with a secure approach. This includes the use of secure design patterns, the employment of secure design practices and approaches such as 'the principle of least privilege', system robustness testing and ensuring that no known vulnerabilities are present. It also covers other development practices that benefit the security and ongoing practices such as ensuring that vulnerabilities are appropriately addressed and that the system can maintain security during its lifetime.
Software Framework	A software framework is a universal, reusable software environment that provides particular functionality as part of a larger software platform to facilitate development of software applications, products and solutions. Software frameworks may include support programs, compilers, code libraries, tool sets, and application programming interfaces (APIs) that bring together all the different components to enable development of a project or system.
TensorFlow	TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications.
TensorFlow Lite	TensorFlow Lite is an open source deep learning framework for on-device inference. (https://tensorflow.org/)
Third-party Applications	An application installed by the User.
User	The user or owner of the Mobile device and/or user the Application for the Mobile Device.

1.5 Abbreviations

Term	Description
AI	Artificial Intelligence
API	Application Programming Interface
AR	Augmented Reality
ASR	Automatic Speech Recognition
Caffe	Convolutional Architecture for Fast Feature Embedding
Caffe2	Caffe2 is a deep learning framework that provides an easy and straightforward way for experimentation with deep learning by using community contributions of new models and algorithms. Users bring their creations to scale using the power of GPUs

Term	Description
	in the cloud or to the masses on mobile with Caffe2's cross-platform libraries. (https://caffe2.ai/docs/caffe-migration.html).
CPU	Central Processing Unit
DNN	Deep Neural Network
DSP	Digital Signal Processing
FAR	False Acceptance Rate
FPE	Facial Photo Enhancement
FRR	False Rejection Rate
GPU	Graphics Processing Unit
GSMA	Global System for Mobile Communications, originally Group Special Mobile Association
MAC	Multiply-accumulate
MACE	Mobile AI Compute Engine
MEC	Mobile Edge Computing
NPU	Neural Processing Unit
OPS	Operations Per Second
OPS/W	Operations Per Second / Per Watt
SDK	Software Development Kit
SDO	Standards Developing Organization
SE	Secure Element
TAR	True Acceptance Rate
TEE	Trusted Execution Environment
TOPS	Tera Operations Per Second
TTS	Text-To-Speech
VGG	Visual Geometry Group (Department of Engineering Science, University of Oxford)

1.6 References

Requirements SHALL be based on the exact versions as indicated below. However, if the manufacturers use a later release and/or another version this SHALL be indicated. The GSMA will take efforts to continually align with other SDOs for timely information about release plans.

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For updated references, that latest edition of the referenced document (including any amendments) applies.

Ref	Doc Number	Title
[1]	RFC 2119	"Key words for use in RFCs to Indicate Requirement Levels", S. Bradner, March 1997. Available at http://www.ietf.org/rfc/rfc2119.txt
[2]		Dark Patterns: https://www.darkpatterns.org/

[3]	ETSI GS MEC	Series standards, available at https://www.etsi.org/technologies/multi-access-edge-computing
[4]	GPD_SPE_009	TEE System Architecture Available at https://globalplatform.org/specs-library/
[5]		https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1#Human%20agency
[6]	RFC8174	Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words https://www.rfc-editor.org/info/rfc8174
[7]	TSG33_028	GSMA AI Mobile Device Guidelines Study Report 2018 https://infocentre2.gsma.com/gp/wg/TS/WorkingDocuments/TSG33_028%20TSG%20Study%20Report%20of%20AI%20Mobile%20Device%20Guidelines%20v2.0.pptx

1.7 Modal verbs terminology

The key words “MUST”, “MUST NOT”, “REQUIRED”, “SHALL”, “SHALL NOT”, “SHOULD”, “SHOULD NOT”, “RECOMMENDED”, “MAY”, and “OPTIONAL” in this document are to be interpreted as described in RFC 2119 **Error! Reference source not found.** (RFC8174) [6] when, and only when, they appear in all capitals, as shown here.

2 AI Mobile Device Definition

An AI Mobile Device refers to a mobile device that has all of the following characteristics:

1. On-device computational resources to enable AI deep learning and other AI algorithms based on either dedicated AI hardware or general hardware to support deep learning AI applications.
2. On-device software framework to support the updating of AI deep learning neural networks.
3. On-device AI software to perform inferencing using deep neural network models.

3 The Requirements of AI Mobile Device

3.1 Hardware requirements

AI Mobile Device hardware is required to support AI software applications efficiently.

Hardware performance measurements can be found in the Table below using the modified VGG 16. Alternatively, a better network may be used.

Requirement for the modified VGG 16 network	
TS47_3.1_REQ_001	An AI Mobile Device SHOULD have a minimum of (1) int8 TOPS.
TS47_3.1_REQ_002	An AI Mobile Device SHOULD have a minimum of (0.5) float16 TOPS.
TS47_3.1_REQ_003	An AI Mobile Device SHOULD have a minimum of (0.5) int8 TOPS/Watt.
TS47_3.1_REQ_004	An AI Mobile Device SHOULD have a minimum of (0.3) float16 TOPS/Watt.

3.2 Software requirements

AI Mobile Device software requirements:

TS47_3.2_REQ_001	An AI Mobile Device SHALL support on-device model updates of an existing deep learning network.
TS47_3.2_REQ_002	An AI Mobile Device SHALL support native APIs to expose the AI hardware functions.
TS47_3.2_REQ_003	An AI Mobile Device SHALL support application APIs (See Appendix A) for native and third-party applications to access Computer Vision (CV), Automatic Speech Recognition (ASR), Natural Language Understanding (NLU) models.
TS47_3.2_REQ_004	An AI Mobile Device SHOULD provide an SDK to convert DNN models from an existing format to the native format of the AI mobile device. Non-exhaustive examples of DNN model file format are: *.ckpt or *.pb, *.tflite, *.prototxt, *.pb or *.pth or *.pt, *.json and *.onnx.
TS47_3.2_REQ_005	An AI Mobile Device SHOULD provide an SDK to support definition of new customized Deep Learning operators.

For the existing SDKs and APIs refer to Annex A.1.

3.3 Performance

The device SHALL use a benchmark system (e.g. MLPERF.org, AI-benchmark.com, AIT China Telecom etc.) to generate an inferencing performance report.

3.4 AI Application Requirements

AI applications may include but are not limited to biometric functions, image processing, speech, augmented reality (AR) and system optimization categories. If any such functions are supported on the device then the following requirements apply.

3.4.1 Biometric Performance Requirements

TS47_3.4.1_REQ_001	An AI Mobile Device SHOULD support a 2D facial biometric system.
TS47_3.4.1_REQ_002	An AI Mobile Device SHOULD support a 3D facial biometric system.
TS47_3.4.1_REQ_003	An AI Mobile Device SHOULD support a fingerprint biometric system.
TS47_3.4.1_REQ_004	An AI Mobile Device supporting 2D facial biometric system SHALL support the biometric KPI requirement TS47_3.4.1_REQ_004.1 for each of the use cases: Device Unlock, Application Login and Payment Authorization.
TS47_3.4.1_REQ_004.1	2D Facial FAR <= (0.002)% and FRR <= (3)% simultaneously
TS47_3.4.1_REQ_005	An AI Mobile Device supporting 3D facial biometric system SHALL support the biometric KPI requirement TS47_3.4.1_REQ_005.1 for each of the use cases: Device Unlock, Application Login and Payment Authorization.
TS47_3.4.1_REQ_005.1	3D Facial FAR <= (0.001)% and FRR <= (3)% simultaneously.
TS47_3.4.1_REQ_006	An AI Mobile Device supporting fingerprint biometric system SHALL support the biometric KPI requirement TS47_3.4.1_REQ_006.1 for

	each of the use cases: Device Unlock, Application Login and Payment Authorization.
TS47_3.4.1_REQ_006.1	Fingerprint FAR <= (0.002)% and FRR <= (3)% simultaneously.
TS47_3.4.1_REQ_007	The biometric key performance indicators (KPIs) for the supported biometric system SHOULD be certified by one or more of the following programs: Fast IDentity Online (FIDO) Alliance Biometric Component Certification Program. Internet Finance Authentication Alliance (IFAA) biometric Certification Program.

3.4.2 On-Device Image Processing Requirements

This section defines the requirements for on device computer vision capabilities and Device Image Processing Application

TS47_3.4.2_REQ_001	An AI Mobile Device SHOULD have optical character recognition (OCR) capability on the device.
TS47_3.4.2_REQ_002	An AI Mobile Device SHOULD have image detection, image classification and image segmentation capabilities on the device.
TS47_3.4.2_REQ_003	An AI Mobile Device SHOULD have face detection and face clustering capabilities within a group of photos on the device.
TS47_3.4.2_REQ_004	An AI Mobile Device SHOULD have video super-resolution capabilities on the device.
TS47_3.4.2_REQ_005	An AI Mobile Device SHOULD have video classification capabilities on the device.

3.4.2.1 On-Device Image Processing Applications

TS47_3.4.2.1_REQ_001	The AI Mobile Device SHOULD support photo scene detection and recognition where the User has the ability to consent to their use.
TS47_3.4.2.1_REQ_001.1	If REQ_001 is supported then the AI Mobile Device SHALL support Identification of one or more objects in different scenes such as portraits, landscapes, foods, night scenes and texts, etc.
TS47_3.4.2.1_REQ_001.2	If REQ_001 is supported then the AI Mobile Device SHALL support Scene detection capabilities to optimize camera settings for image capture based on scene content.
TS47_3.4.2.1_REQ_002	The AI Mobile Device SHOULD support text detection and recognition of installed language packages, where the User has the ability to consent to the text detection and recognition use.
TS47_3.4.2.1_REQ_003	The AI Mobile Device SHOULD support automatic language detection.
TS47_3.4.2.1_REQ_004	The AI Mobile Device SHOULD provide personalized FPE for Users based on gender, age, and skin tone.
TS47_3.4.2.1_REQ_005	The AI Mobile Device SHOULD support FPE of multiple people in a single photo.

TS47_3.4.2.1_REQ_006	The FPE functionality SHOULD be switched off by default and the AI Mobile Device SHOULD support User adjustment of the FPE level from no enhancement to the max FPE.
TS47_3.4.2.1_REQ_007	The AI Mobile Device SHOULD support automatic classification of photos in an album by different categories.

Note: FPE functionality is recommended to be automatically off by default in order to give the User the choice of whether to turn this feature on. This is in recognition of mental health and ethical concerns.

3.4.3 Speech

Requirements for speech ability include such functions as voice recognition, text to speech, voice activation etc.

TS47_3.4.3_REQ_001	The AI Mobile Device SHOULD have speech ability.
TS47_3.4.3_REQ_002	The AI Mobile Device SHOULD support Automatic speech recognition (ASR) capabilities where the User has the ability to consent to ASR.
TS47_3.4.3_REQ_003	The AI Mobile Device SHOULD support Natural Language Understanding (NLU) capabilities where the User has the ability to consent to NLU.
TS47_3.4.3_REQ_004	The AI Mobile Device SHOULD support Synthesized Voice (Text-To-Speech (TTS) capabilities where the User has the ability to consent to TTS.
TS47_3.4.3_REQ_005	If the AI Mobile Device supports Voice Assistant then the requirements in section 3.4.3.1 SHALL apply.

3.4.3.1 Voice assistant

TS47_3.4.3.1_REQ_001	AI Mobile Device SHALL support the following functions. Automatic speech recognition (ASR) capabilities. Natural Language Understanding (NLU) capabilities. Synthesized Voice (Text-To-Speech (TTS)) capabilities.
TS47_3.4.3.1_REQ_002	The AI Mobile Device SHALL support voice trigger, and its specific requirements are listed in the following sub requirements: TS47_3.4.3.1_REQ_002.1, 002.2 and_002.3
TS47_3.4.3.1_REQ_002.1	The AI Mobile Device SHOULD support voiceprint recognition for preventing people other than the device's owner from triggering voice assistant.
TS47_3.4.3.1_REQ_002.2	In a quiet environment, the following SHALL be required: The true acceptance rate (TAR) >= (90)%, and the false acceptance rate (FAR) of voiceprint recognition <= (20)%.
TS47_3.4.3.1_REQ_002.3	In a noisy environment, the following SHALL be required: TAR >=(80)%, and FAR of voiceprint recognition <= (20)%.
TS47_3.4.3.1_REQ_003	The AI Mobile Device SHALL have on-device speech recognition library (i.e. with no access to the Internet) for changing the system

	setting (e.g. Turn Bluetooth on/off via voice assistant) and invoking the native applications (e.g. send SMS via voice assistant).
TS47_3.4.3.1_REQ_004	The AI Mobile Device SHOULD have access to different categories of applications and invoke these applications' services and functions via voice assistant.
TS47_3.4.3.1_REQ_005	The AI Mobile Device SHALL support information search by on-device voice assistant.
TS47_3.4.3.1_REQ_006	The AI Mobile Device SHOULD support interaction with smart devices (e.g. home appliances) via voice assistant.

3.4.4 Augmented Reality (AR)

TS47_3.4.4_REQ_001	The AI Mobile Device SHOULD provide the following AI capabilities for AR native and third-party applications: 1. Hand gesture recognition. 2. Hand skeleton tracking. 3. Human body pose recognition. 4. Human body skeleton tracking.
TS47_3.4.4_REQ_002	The AI Mobile Device SHOULD support the following applications: 1. AR Emoji a. Creating customized AR-based Emoji. b. Tracking User's facial movement and expression and render these on the AR-based Emoji. 2. AR video a. Compositing real objects with virtual objects and/or virtual background. b. Minimum (30) fps frame rate. c. AR shadow effect and occlusion handling. d. AR enhanced information text labels should not deviate or disappear from the actual target scene when the AI Mobile Device moves.

3.4.5 System Optimization

TS47_3.4.5_REQ_001	Only with the explicit permission of the User in order to respect the User's right to privacy around their habits: the AI Mobile Device SHOULD support dynamic system resource allocation and optimization based on feedback provided by on-device sensors measuring environmental conditions combined with continuous learning of User habits and behaviours or device or network usage or performance indicators: 1. Dynamic application management (e.g. pre-loading, closing, put to sleep, control network access) based on User's habits (e.g. usage duration, frequency). 2. Dynamic application management based on abnormal behaviour detection (e.g. increased memory usage, abnormal power consumption, self-starting in the background).

	3. Dynamic system resource management based on continuous learning of system performance (e.g. memory and storage defragmentation, off-line storage during off-peak periods). 4. Dynamic system resource allocation for high performance applications (e.g., gaming and video).

4 Privacy and Security Requirements

The following section defines the privacy and security requirements. Many privacy, data protection, and information security laws, rules, and regulations (including those found in GDPR) calibrate their obligations based on the degree of risk posed to individuals as a result of the processing.[1] GSMA, and these privacy and security requirements, recognize that differing AI Functions may pose differing degrees of risk. These privacy and security requirements, therefore, should be interpreted and applied in relation to the degree of risk to the privacy, data protection, and information security rights and interests of Users.

4.1 Privacy Requirements

Applicable law(s) and regulations as related to privacy and data protection must be complied with in connection with AI on mobile device. For avoidance of doubt, where laws are not in place in certain jurisdictions, manufacturers should respect the User and not leave AI functionality 'on' by default. It should be 'Private by Design and by Default'. Any choice to turn off functionality by the User must be fully respected and techniques, such as 'Dark Patterns' **Error! Reference source not found.**, that seek to manipulate a User's free choice should be avoided.

TS47_4.1_REQ_001	AI on mobile device SHOULD comply with the privacy laws in the country where the device is commercially retailed.
TS47_4.1_REQ_002	Appropriate technical and organisational safeguards SHOULD be implemented to ensure that, by default, only the personal data reasonably necessary for a specific purpose are processed.
TS47_4.1_REQ_003	AI Applications that process Personal Data SHALL be off by default unless processing exclusively takes place locally on the device.
TS47_4.1_REQ_003.1	The User SHOULD be allowed to control whether individual AI applications are switched on.
TS47_4.1_REQ_003.2	The User SHOULD be allowed to control whether individual AI applications are switched off.
TS47_4.1_REQ_004	The AI Application on the AI Mobile Device SHALL be designed in such a way that a Data Processor will have the responsibility to: 1) Be transparent with the User on the nature of the input data used in the AI processing (e.g. personal files, biometrics, …).

[1] See, e.g., GDPR Article 35(1) requiring a Data Protection Impact Assessment where the processing, "is likely to result in a high risk to the rights and freedoms of natural persons", and GDPR Article 34(1) requiring a notice of personal data breach where it, "is likely to result in a high risk to the rights and freedoms of natural persons."

	2) Forbid transferring personal data processing off the device except if the User has explicitly agreed or other legal basis has been satisfied in accordance with the law. 3) Forbid transferring results of on-device AI processing containing personal data off the device except if the User has explicitly agreed or other legal basis has been satisfied in accordance with the law.

4.2 Security Requirements

Applicable law(s) and regulations as related to security and data protection must be complied with in connection with AI on mobile device. For avoidance of doubt, where laws are not in place in certain jurisdictions, manufacturers should respect the User and not leave AI functionality 'on' by default. From a security perspective this also follows the 'principle of least privilege', ensuring that systems have no more access than is necessary, as a default starting point. The AI Mobile Device needs to operate as 'Secure by Default'. Any choice to turn off functionality by the User must be fully respected and techniques, such as 'Dark Patterns' **Error! Reference source not found.** that seek to manipulate a User's free choice should be avoided. This assists in retaining User trust and helps prevent subversion by malicious actors.

TS47_4.2_REQ_001	The AI Mobile Device SHALL use reasonable safeguards appropriate to the sensitivity, confidentiality and integrity of the information.
TS47_4.2_REQ_002	Except as required or permitted by applicable law, the User SHALL always remain in control of the collection of their personal data and its usage, in order to minimise the risk of malicious usage or data leakage.
TS47_4.2_REQ_003	Off 'toggle' switches SHALL turn off the functionality, except as permitted or required by applicable law.
TS47_4.2_REQ_004	Techniques, such as 'Dark Patterns', that manipulate the User's choice SHALL NOT be used.

4.2.1 Security for AI Applications

TS47_4.2.1_REQ_001	The AI models used by an AI Mobile Device SHOULD be secure and robust, and be protected with appropriate safeguards to prevent and to mitigate attacks.
TS47_4.2.1_REQ_002	Defence techniques SHOULD be employed to protect the training data for protecting models. For example, in evasion attacks, data can be manipulated to mislead AI models.
TS47_4.2.1_REQ_003	Autonomous AI Mobile Device operations SHALL be controlled, and/or authorized by the authenticated User.
TS47_4.2.1_REQ_004	AI Mobile Device operations SHOULD be performed in the Secured Environment **Error! Reference source not found.**, e.g. a secure boot and upgrade is enforced, and the system integrity is protected.
TS47_4.2.1_REQ_005	Data and metadata for AI Mobile Device SHALL be stored with encryption with keys that are stored securely in a Secured Environment, e.g. Trusted Execution Environment (TEE) **Error! Reference source not found.**.

TS47_4.2.1_REQ_006	Biometric Data, which are processed by an AI Application (e.g. templates) used for authentication within the AI Mobile Device, SHALL NOT be transferred off the device.

AI applications for high security requirements should use the following defence techniques on AI models:

- Network distillation, adversarial training, adversarial sample detection, etc. are recommended to prevent AI models from evasion attacks.
- Training data filtering, regression analysis, ensemble analysis, etc. are recommended to be used to prevent AI models from poisoning attacks.
- Encryption algorithm or better, input pre-processing, model pruning, etc. are recommended to prevent AI models from backdoor attacks.

4.2.1.1 Biometric Authentication

TS47_4.2.1_REQ_007	Users' Biometric Data (such as facial data, fingerprint data, etc.) SHALL be encrypted when at rest on the device. Encryption/decryption of this data SHALL be performed in a Secured Environment **Error! Reference source not found.**.
TS47_3.2.1_REQ_007.1	Biometric Data SHALL also be stored in the Secured Environment.
TS47_4.2.1_REQ_008	Biometric algorithms (such as face recognition algorithms, fingerprint algorithms, etc.) SHOULD run in a private and Secure Environment such as a Trusted Execution Environment (TEE) **Error! Reference source not found.**.
TS47_4.2.1_REQ_009	If Users' Biometric Data is replaced, the previous Biometric Data before the replacement SHALL be deleted completely and permanently and not be recoverable by data rollback.
TS47_4.2.1_REQ_010	The Biometric Data SHALL be wiped and made unrecoverable by a device factory reset.

4.2.1.2 Speech

TS47_4.2.1_REQ_011	Voiceprint Data SHOULD be stored on the device with encryption.
TS47_4.2.1_REQ_012	The temporary Voiceprint Data SHALL NOT remain in the memory after processing.
TS47_4.2.1_REQ_013	When the Voiceprint Data is permanently and completely deleted, it SHALL NOT be recoverable by data rollback.
TS47_4.2.1_REQ_014	The Voiceprint Data SHALL be wiped and made unrecoverable by a device factory reset.
TS47_4.2.1_REQ_015	The device SHOULD be resistant to voice replay attacks.

4.2.1.3 Augmented Reality

TS47_4.2.1_REQ_016	Appropriate safeguards SHOULD be used to protect AR applications from malicious application attacks, such as spoofing a User with information about the real and/or virtual world, sensory overload attacks, hijacking the User's clicks, etc.

5 AI Agent (informative)

This section and its subsections are informative

5.1 General

An achievement of deep learning is its extension to the domain of reinforcement learning. In the context of reinforcement learning, an autonomous agent learns to perform a task by trial and error, with minimal guidance from the User.

Examples of AI agent capabilities but not limited to:

1. If granted permission from the User or owner of the data, the agent is responsible for the decision-making of AI computation offloading, and may implement a MEC-first strategy, i.e. abstract the computation offloading decision function from specific application and make it become a functional entity on AI mobile device.

On-device deep reinforcement learning will enable a device to perceive the environment and react autonomously. Supporting more and more autonomous applications will be the trend, which will make an AI Mobile Device significantly different from the smartphone of today.

AI agents are software entities which can carry out some actions on behalf of clients with some degree of autonomy.

In general, agents possess five common properties which are autonomy (some level of self-control), adaptiveness (the ability to learn and improve performance with experience), reactivity (the ability to perceive the environment and to respond in a timely fashion to changes that occur), proactivity (the ability not only to act simply in response to their environment but also to exhibit goal-directed behaviour by taking the initiative) and sociability (the ability to interact, communicate and work with other agents).

Incorporating an AI agent will dramatically change the landscape of mobile devices. It can act as the “brain” of the mobile device, to control the behaviour and system performance of the device. It can act as the new “entrance of services”, recommend services (applications) to the end User based on context.

In the future, the AI agent will become an important feature for defining an AI mobile device.

5.2 Privacy and Security Requirements for AI Agent

The User and/or management entity needs to be provided with notice about how the AI agent may affect them.

At all times, the principle of ‘Human in Command’ needs to be adhered to when an AI agent makes decisions to transfer data off the device.

The User will be able to provide express permission, or other appropriate legal basis, for specific data to be transferred away from the device and be able to learn the categories of personal data being processed (e.g. images, categories of information and so on).

Permission rules will comply with TS47_4.1_REQ_001 of this document, a device respects the decision of the User.

Example: If a User states a period of time that expires and subsequently then chooses to say ‘once’, the device needs to not store the data in the intervening period between permissions on-device and then upload / offload this data.

Personal data from 3rd parties who are within the proximity of the device should not be transferred off the device by the AI agent. The expectation is that best endeavours would be made to use AI processes to filter out background information and not to inadvertently capture third party information (e.g. voices, faces etc.).

The decisions and recommendations made by the AI agent need to be understandable by a User.

An AI agent needs to be protected from external threats.

6 Network Requirements to Support AI Mobile Devices (informative)

Computation on AI mobile devices may be improved by offloading to MEC or Cloud to reduce latency and mobile power consumption if permitted by regulation and law. The ubiquitous AI Mobile Device will make AI computation a very important task for the network to bear, which will ultimately drive the network to change.

1. Cloud computing centres may have the ability to provide AI as a service.
2. MEC may have the ability to provide AI as a service, which is equivalent to location service, bandwidth management service and radio network information service, and provide unified open APIs **Error! Reference source not found.**.
3. Networks may gradually evolve from a communication platform to a platform that supports both communication and computation, in order to better support edge learning.

Annex A Informative

A.1 SDK & API

Currently, each chipset vendor has its own set of APIs, which leads to a fragmented ecosystem. Standardising and unifying application APIs is very necessary and highly recommended.

A.1.1 The Android Neural Networks API (NNAPI)

The Android Neural Networks API (NNAPI) is an Android C API designed for running computationally intensive operations for machine learning on mobile devices. NNAPI is designed to provide a base layer of functionality for higher-level machine learning frameworks (such as TensorFlow Lite, Caffe2, or others) that build and train neural networks.

< Official website URL, https://developer.android.com/ndk/downloads>

A.1.2 The Snapdragon Neural Processing Engine (SNPE)

The Snapdragon Neural Processing Engine (SNPE) is a Qualcomm Snapdragon software accelerated runtime for the execution of deep neural networks. The Qualcomm Neural Processing SDK for artificial intelligence (AI) is designed to help developers run one or more neural network models trained in Caffe/Caffe2, ONNX, or TensorFlow on Snapdragon mobile platforms, whether that is the CPU, GPU or DSP.

Official website URL, https://developer.qualcomm.com/software/qualcomm-neural-processing-sdk

A.1.3 HiAI

HiAI is a mobile terminal–oriented artificial intelligence (AI) computing platform that constructs three layers of ecology: service capability openness, application capability openness, and chip capability openness. The three-layer open platform that integrates terminals, chips, and the cloud brings more extraordinary experiences for Users and developers.

Official website URL, https://developer.huawei.com/consumer/en/devservice/doc/2020301

A.1.4 NeuroPilot

NeuroPilot is MediaTek's AI ecosystem. It embraces the advantages of 'Edge AI', which means the AI processing is done on-device rather than relying on a fast internet connection and Cloud service. However, NeuroPilot doesn't have to use a dedicated AI processor. Its software can intelligently detect what compute resources are available, between CPU, GPU and APU, and automatically choose the best one.

A.1.5 Core ML

Core ML is an Apple framework that allows developers to easily integrate machine learning (ML) models into apps. Core ML is available on iOS, watchOS, macOS, and tvOS. Core ML introduces a public file format (.mlmodel) for a broad set of ML methods including deep neural networks (convolutional and recurrent), tree ensembles (boosted trees, random forest, decision trees), and generalized linear models.

Official website URL, https://developer.apple.com/documentation/coreml

A.1.6 MACE

Mobile AI Compute Engine (MACE) is a deep learning inference framework optimized for mobile heterogeneous computing on Android, iOS, Linux and Windows devices. The design focuses on the following targets:

[1] Performance: Runtime is optimized with NEON, OpenCL and Hexagon, and Winograd algorithm is introduced to speed up convolution operations. The initialization is also optimized to be faster.

[2] Power consumption: Chip dependent power options like big.LITTLE scheduling, Adreno GPU hints are included as advanced APIs.

[3] Responsiveness: UI responsiveness guarantee is sometimes obligatory when running a model. Mechanism like automatically breaking OpenCL kernel into small units is introduced to allow better pre-emption for the UI rendering task.

[4] Memory usage and library footprint: Graph level memory allocation optimization and buffer reuse are supported. The core library tries to keep minimum external dependencies to keep the library footprint small.

[5] Model protection: Model protection has been the highest priority since the beginning of the design. Various techniques are introduced like converting models to C++ code and literal obfuscations.

[6] Platform coverage: Good coverage of recent Qualcomm, MediaTek, Pinecone and other ARM based chips. CPU runtime supports Android, iOS and Linux.

[7] Rich model formats support: TensorFlow, Caffe and ONNX model formats are supported.

Official website URL, https://github.com/XiaoMi/mace

Annex B Document Management

B.1 Document History

Version	Date	Brief Description of Change	Approval Authority	Editor / Company
1.0	23/02/211	New PRD	TSG42e (Call) TG - email	Kay Fritz / Vodafone

B.2 Other Information

Type	Description
Document Owner	Terminal Steering Group (TSG)
Editor / Company	Kay Fritz / Vodafone

It is our intention to provide a quality product for your use. If you find any errors or omissions, please contact us with your comments. You may notify us at prd@gsma.com

Your comments or suggestions & questions are always welcome.