Androidサードパーティ SDKの プライバシーラベル開示ガイドラインの分析

秋山 由依[1,a)]　稲吉 弘樹[1]　齋藤 彰一[2]　門田 暁人[1]

概要：近年，スマートフォンアプリの開発者には，アプリが収集するユーザデータやその利用目的をアプリストア上でプライバシーラベルとして開示することが求められている．特に Google Play では，開発者が自己申告するデータセーフティが導入されている．多くのアプリにはサードパーティ製の SDK が数多く組み込まれており，開発者はこれら SDK によって収集されるデータの種類や利用目的についても把握し，正確に申告する必要がある．そこで，SDK 提供元は自社 SDK に関するプライバシーラベル開示ガイドラインを公開している．しかしこれらのガイドラインが，データセーフティ申告における一連の質問項目をどの程度カバーしているのか，また申告フォームで使用される公式な分類・用語に合っているのかについては，これまで十分な分析が行われていない．本稿では，広く利用される商用 SDK を対象に 57 個のガイドラインを収集し，データセーフティ申告の手順に基づいて設定した 18 の観点から分析を行った．その結果，申告手順の各段階に関するガイドライン中の記述において幅広く不備が存在し，正確な申告を困難とさせている実態が明らかとなった．

キーワード：Android，サードパーティ SDK，プライバシー漏洩，プライバシーラベル開示ガイドライン

Preliminary Analysis of Privacy Label Disclosure Guidelines for Android Third-party SDKs

Yui Akiyama[1,a)]　Hiroki Inayoshi[1]　Shoichi Saito[2]　Akito Monden[1]

Abstract: In recent years, mobile app developers have been required to disclose, via privacy labels on app stores, the types of user data collected and their purposes of use. On Google Play, this is implemented through the self-reported “Data Safety” section. Many apps integrate multiple third-party SDKs, requiring developers to understand and accurately declare data collection and usage by these SDKs. To assist developers, SDK providers have published privacy label disclosure guidelines for their products. However, there has been little systematic analysis of how comprehensively these guidelines address the full set of Data Safety disclosure questions or whether they align with the official classifications and terminology used in the reporting form. In this study, we collected 57 guidelines for SDKs listed in the Google Play SDK Index, a widely used commercial SDK repository, and analyzed them from 18 perspectives based on the Data Safety reporting procedure. This paper presents our findings.

Keywords: Android, third-party SDK, privacy leak, privacy label disclosure guideline

1. はじめに

近年のプライバシー保護の重要性向上に伴い，スマートフォンアプリの開発者は，アプリが収集するユーザデータやその目的をプライバシーラベルとしてアプリストア上で開示することが求められている．2022 年 7 月，Google Play ではデータセーフティ (DS) の申告が義務付けられ，アプリ開発者は，アプリによるユーザデータおよび端末データの収集・共有の有無，およびその利用目的について

[1] 岡山大学 Okayama University
[2] 名古屋工業大学 Nagoya Institute of Technology
[a)] p8wu2xb7@s.okayama-u.ac.jp

自己申告する必要が生じた．Google は DS 申告の全責任がアプリ開発者にあるとしている．

DS 導入以降，DS 申告結果の実態調査として大規模な測定と開発者インタビュー [1] やデータ削除リクエストに注目した研究 [2]，転送時の暗号化に注目した研究 [3] など幅広く実施されてきた．一方，SDK プロバイダが公開するプライバシー開示ガイドライン (これ以降単にガイドラインと呼ぶ) に関する研究は限られている．ガイドラインには SDK のプライバシー慣行が記載されており，アプリ開発者はこれを頼りに DS 申告を行う．従ってその記載の明瞭さが，ラベルの正確性や申告容易性を左右する．ガイドラインと SDK 本体の実際のプライバシー慣行を比較することでガイドラインの記載漏れの実態を調査した研究が Android 向け [4] や iOS 向け [5] に実施された．しかし，ガイドラインそのものを定量的に分析した研究は存在せず，実態の解明は進んでいない．

本稿ではガイドラインの多様な特徴を手動で分析する．まず，主要な商用 SDK が登録されている Google Play SDK Index (GPSDKI) [6] 内の SDK を対象にガイドラインを収集し，57 個のガイドラインを得た．次に，Android アプリの配布・管理・分析を行うための Google の公式プラットフォームである Google Play Console [7] で DS 申告の手順を確認し，18 の観点を設定した．これらの観点に基づいてガイドラインを分析した結果，質問項目をカバーしていない場合や公式な分類・用語に合っていない場合などの申告を困難にさせる不備を幅広く発見した．

以降，2 章でアプローチを説明する．3 章で結果を報告し，5 章で制限と今後の課題を議論する．6 章で関連研究について述べ，最後に 7 章でまとめを述べる．

2. アプローチ

2.1 ガイドラインの収集

2.1.1 収集対象と方法

Google が公開する GPSDKI には「広範に使用されている商用 SDK が 100 以上掲載」とあり，これを収集対象とする．GPSDKI の各 SDK のページは，ガイドラインへのリンクを明記している場合があり，そのリンク先ページのダウンロードを自動化した．リンクが得られない SDK については，Web 検索による各 SDK のホームページ (HP) からの手動収集を試みる．

2.1.2 収集結果

表 1 にガイドラインの収集結果を示す．先行研究 [4] で我々は，2023 年の 3 月 (自動：31 個)，7 月 (自動：3 個，手動：12 個)，10 月 (自動：1 個，手動：0 個) にガイドラインの収集を行い，合計 47 個のガイドラインを得た．SDK 掲載数はそれぞれ 153 個，156 個，175 個とやや増加傾向にある．より多くのガイドラインを調査するため，2024 年 9 月 (自動：1 個)，2025 年 1 月 (手動：10 個) に追加の収集を行い，新たに 11 個を得た．結果，合計は 58 ガイドラインであり，78 SDK についての記載がある．この内の 2 個は，2025 年 1 月収集時にページの一部のみしか保存できていなかったため，2025 年 6 月に再度サーバへアクセスし，ページ全体を保存した．また，3 個の SDK に関わる 1 個のガイドラインは iOS アプリ向けのものであると判明したため，除外した．従って，本稿で調査する SDK は 75 SDK の 57 ガイドライン (自動：36 個，手動：21 個) となった．

表 1 ガイドライン収集の日付，概要と収集数
Table 1 Guideline collection dates, summary, and numbers.

収集日	ソース	方式	SDK 掲載	取得 (SDK)
2023 年 3 月 4 日	GP SDK Index	自動	153	31
2023 年 7 月 4 日	GP SDK Index	自動	156	3
	各 SDK の HP	手動	-	12
2023 年 10 月 4 日	GP SDK Index	自動	175	1
	各 SDK の HP	手動	-	0
2024 年 9 月 4 日	GP SDK Index	自動	214	1
2025 年 1 月 2 日	各 SDK の HP	手動	-	10
累積	-	-	214	58 (78)
採用	-	-	-	57 (75)

表 2 分析対象の SDK カテゴリおよびガイドライン数
Table 2 SDK categories and the number of guidelines.

SDK カテゴリ	ガイドライン数 (57)
Advertising and monetization	23
Analytics	25
Data management	1
Location	2
Marketing and engagement	15
Payments	4
Social	2
User authentication	1
User support	4

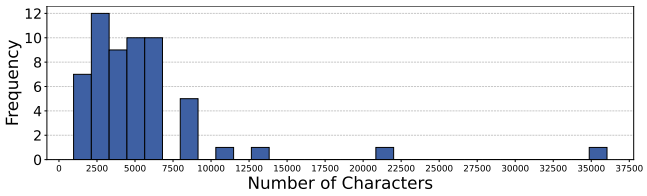

図 1 分析対象ガイドラインの文字数の分布
Fig. 1 Distribution of character counts in the guidelines.

表 2 に調査対象 57 個の SDK カテゴリごとの数を示す．カテゴリは執筆時点の GPSDKI の定義に基づく 9 種類であり，本調査では各カテゴリで最低 1 個のガイドラインの収集に成功した．ただし，SDK によっては複数のカテゴリに属するものもある．**図 1** は文字数の分布を示している．文字数はページ全体，またはガイドラインに関わる部分のみを数えている．約 2,500 文字のガイドラインが 12 個と最も多く，またその前後にほとんどのガイドラインが分布している．一方で 10,000 文字を越えるガイドラインも 4 個存在し，10,000〜35,000 文字以上と幅がある．

表 3 申告フォームのフェーズ，質問事項と，対応する観点数と節
Table 3 Data safety form questionnaire and the number of corresponding features we defined.

フェーズ	質問事項	観点数
データの収集とセキュリティ	収集・共有の有無	-
	転送時の暗号化	1
	データの削除リクエスト	2
データの種類	データの種類の選択	3
データの使用と処理	収集・共有の選択	3
	一時的処理	1
	オプションのデータ収集	3
	収集・共有の目的	5
合計	-	18

2.2 ガイドラインの分析

DS 申告のフェーズ，質問事項に対して，合計で 18 の観点を作成した (**表 3**)．DS 申告は 3 つのフェーズからなり，申告者は大きく 8 種類の質問事項に回答する．本節は各質問事項と観点について説明する．

2.2.1 転送時の暗号化

DS では，収集または共有するユーザデータがあると答えた場合，はじめにデータ転送時の暗号化の有無を問われる．ガイドラインに暗号化に関する記載があれば開発者はその通りに回答できるが，なければ回答できない．ガイドライン内のデータ暗号化有無の記載の実態を明らかにする．

2.2.2 データの削除リクエスト

続いて，ユーザにデータの削除リクエストの方法を提供しているかを問われる．そこで，この質問に回答するための記載に関して以下の 2 つの観点を定めた．

ユーザデータ削除の記載： これに関する記載がガイドラインになければ開発者は申告できない．従って 2.2.1 項の暗号化の観点と同様に，データ削除記載の有無を調査する．

90 日以内のユーザデータ削除の記載： データ削除リクエストの質問では，リクエスト可能・不可能に加えて，「方法は未提供だが 90 日以内に自動でデータ削除を行う」という選択肢がある．そこで 90 日以内の削除に言及しているガイドラインを調査するが，本観点では一概にガイドラインの良し悪しを言えず，特徴の 1 つとして分析を行う．

2.2.3 データの種類の選択

DS では 14 カテゴリ 38 データタイプについて収集・共有するデータタイプを申告する．ガイドラインがすべてのデータタイプをカバーしていれば，開発者はその通りに申告できる．しかし，未記載のデータタイプがある場合，それを申告しなくてもよいのかどうか不明確である．そこで，3 つの観点を設定して分析する．

データカテゴリの記載： ガイドラインをカテゴリ (C) の一致度に応じて 4 個のグループに分類する．公式に準拠した 14 カテゴリを過不足なく記載している (C-完全)，記載されているカテゴリはすべて公式に準拠しているが一部未記載のカテゴリがある (C-不完全)，独自のカテゴリを含む (C-非公式追加)，カテゴリの記載なし (C-なし) である．

データタイプの記載： ガイドラインに記載されたデータタイプ (T) の網羅性の有無を分析する．データカテゴリの記載が C-完全の場合，公式のデータタイプ 38 種類をすべて記載していれば網羅性ありとして T-完全，そうでない場合は T-不完全に分類する．また C-不完全と C-非公式追加の場合は，記載されたデータカテゴリの範囲で公式のデータタイプがすべて記載されていれば T-完全とする．一方で C-なしのガイドラインは公式データカテゴリが存在しないため，データタイプの分析は実施しない．また，5 つのデータカテゴリはそれぞれ 1 つのデータタイプのみを含んでおり，データタイプを省略しても申告上の問題はないため，データタイプ記載ありとみなして分析を行う．

データの種類の記載方法： データカテゴリやタイプがガイドラインにどのような形式で掲載されているかを調査する．表形式とその他に分類する．形式だけで良し悪しは決まらないが，ガイドラインの特徴の 1 つとして調査する．

2.2.4 収集・共有の選択

DS では 14 カテゴリ 38 データタイプのそれぞれについて，収集 (collect) と共有 (share) の有無を申告する．これに関して以下の 3 つの観点を作成した．

収集・共有動作の記載の種類： ガイドラインの記載で “collect” や “share” といった単語が 2 種類あれば収集・共有が明確に区別でき，その通りに申告できる．一方その他の単語が 1 種類だけの場合，収集・共有のどちらとして申告するかは開発者の判断に委ねられるため，申告が困難になる．そこで，動作の記載が 2 種類のもの，“collect” 1 種類のもの，その他 1 種類のものの 3 グループに分類する．

共有曖昧ガイドライン (G_{SA})： ガイドラインの記載上は収集 (collect) であっても，実際にはサードパーティによる収集，すなわち共有 (share) として申告すべき場合がある．このとき，共有として申告するかどうかはアプリ開発者の判断に依存し，誤申告のリスクがある．そこで，以下のいずれかの条件を満たすガイドラインを共有曖昧ガイドライン (G_{SA}) と定義し，その数を調査する：

(1) 収集が “yes” または “optional” であり，共有が “no” である．この時，収集の記載が実際には共有動作を含意する可能性があり，単純に共有申告を無視できない．
(2) 記載動作が「共有」以外の 1 種類のみであり，そのラベルが “yes” または “optional” である．この時，動作の意味解釈が開発者に依存し，申告漏れの恐れがある．

これを論理式で表すと次のようになる：

$$
\begin{aligned}
G_{\mathrm{SA}} \iff & (A_{\mathrm{c}} \in \{\text{“yes”}, \text{“optional”}\} \ \wedge \ A_{\mathrm{s}} = \text{“no”}) \\
& \vee \ (|O| = 1 \ \wedge \ O \neq \{\mathrm{share}\} \\
& \quad \wedge \ A_o \in \{\text{“yes”}, \text{“optional”}\})
\end{aligned}
$$

ここで，データ収集・共有に関する記載内容をそれぞれ

A_c，A_s と表す．O はガイドラインに記載されている動作の集合であり，A_o は O の要素に関する記載内容を表す．

収集曖昧ガイドライン (G_{CA})： G_{SA} とは異なり，収集の有無が不明瞭なガイドラインに注目する．デバイス上のアプリ内で動作する SDK がデータをサードパーティへ送信する動作は「共有」であるが，データがデバイス外へ出る場合は総じて「収集」として申告が必要であるため，SDK による「共有」を申告する場合は同時に「収集」の申告も必要な可能性が高い．以下のいずれかの条件を満たすガイドラインを収集曖昧ガイドライン (G_{CA}) と定義する：

(1) 収集についての記載がなく，かつ共有が行われる場合．
(2) 初期設定を意味するデフォルト収集のラベルが “no” であり，かつ共有が選択式 (optional) である場合．

これを論理式で表すと次のようになる：

$$G_{\mathrm{CA}} \iff (A_c = \varnothing \ \wedge \ A_s \in \{\text{“yes”}, \text{“optional”}\}) \vee (A_{\mathrm{dc}} = \text{“no”} \ \wedge \ A_s = \text{“optional”})$$

ここで $\varnothing$ は記載なしを，A_{dc} はデフォルトのデータ収集に関する記載内容を表す．

2.2.5 一時的処理

DS では収集ありと申告したデータタイプについて，それぞれ「一時的処理」に該当するかどうかを回答する．ガイドラインに一時的処理について記載がなければ申告が困難になる．この記載について調査し実態を明らかにする．

2.2.6 オプションのデータ収集

DS では収集ありと申告したデータタイプについて，それがエンドユーザによって設定変更可能かどうかを申告する．これに関する分析のため，観点を 3 つ設定した．

オプションのデータ収集の記載： 記載がない場合，申告が困難になる．記載の有無を調査し実態を明らかにする．

オプションのデータ収集の目線の記載： ガイドラインによってはエンドユーザ目線ではなく SDK 利用者 (アプリ開発者) 目線で記載しており，誤申告の恐れがある．目線を調査するために，設定方法や動作の主体に注目する．まず，アプリ開発者が行う設定 (例．プログラムの変更) の場合は開発者目線と判定する．主語が “you” や “app developers” など開発者を指している場合も，開発者目線と判定する．次に，ユーザが行う設定 (例．デバイス設定の変更) や主語が “end user” の場合はエンドユーザ目線と判定する．また，目線が複数種類記載されている場合は「両方」として数え，目線の判定が困難な場合は「不明」とする．

共有オプショナルガイドライン (G_{SO})： DS 申告では収集と共有の 2 種類の動作があるが，「オプション」は収集専用であり「オプションの共有」はあり得ない．これに反して，データの収集と共有の 2 種類の動作が記載され，収集のラベルが「no」，共有のラベルが「optional」となっているものを，共有オプショナルガイドライン (G_{SO}) と呼ぶ：

$$G_{\mathrm{SO}} \iff A_c = \text{“no”} \ \wedge \ A_s = \text{“optional”}$$

2.2.7 収集・共有の目的

DS では申告した各データタイプについて，その収集・共有それぞれの目的を回答する．ガイドライン中で，申告するすべてのデータタイプに関して，そしてそのすべての収集・共有動作に関して目的の記載があれば，開発者はその通りに申告できる．しかし，目的の記載漏れがあると申告できず，また動作やデータタイプと目的の対応付けの曖昧さも問題である．分析のために 5 つの観点を設定した．

目的の記載： 目的の記載自体がない場合，申告は困難である．そこで，目的の記載の有無を調査する．

目的の記載方法： ガイドライン中に記載される「目的」の形式を 4 グループに分類する．データカテゴリやタイプの表に目的を組み込んでいるもの，文章形式のもの，目的専用の表を掲載するもの，その他である．本観点は良し悪しの評価ではなく，特徴の 1 つとして調査する．

目的に対応する収集・共有の区別： 動作と目的の対応付けの曖昧さを調査する．「区別可能・自明」と「区別不可能」の 2 つに分類する．「区別可能・自明」は例えば，収集の目的と共有の目的を分けて記載する場合や，目的が 1 種類しかなく両動作同じ目的とみなせる場合である．「区別不可能」は例えば，目的が複数あるが動作ごとに分けられていない場合や，一方の動作の目的が未記載の場合である．

目的に対応するデータ： ガイドラインに記載されたデータタイプと目的の対応付けの曖昧さを分析する．目的に関する記載において，データタイプまたは目的のどちらか一方，またはその両方が 1 つのみであれば対応は明確であり「自明」と分類する．一方で，複数のデータタイプに対して複数の目的が記載されている場合は「多対多」と分類する．

公式に従った目的の記載： 目的の申告では 7 種類 (App functionality; Analytics; Developer communications; Advertising or marketing; Fraud prevention, security, and compliance; Personalization; Account management) から目的を選択する．ガイドラインの記載がこれら 7 種類に従っていない場合，アプリ開発者が独自に判断して申告する必要がある．そこで，記載された目的がすべて公式に従っているもの，一部従っているもの，完全に従っていないものの 3 グループにガイドラインを分類する．

3. 結果

3.1 転送時の暗号化

分析の結果，記載ありが 40 個 (70%) と多数を占める一方で，記載なしが 17 個 (30%) と一定数存在する実態が明らかになった．記載ありの 40 個はすべて「暗号化を行う」と示している．記載なしとした 1 個のガイドラインは「シリアライズ化を行う」旨の記述があるが，これは暗号化とは異なると判断し，記載なしに分類した．また単に暗号化

表 4 ユーザデータ削除の記載を含む・含まないガイドライン
Table 4 User data deletion feature analysis result.

記載あり	プライバシーポリシー参照を指示	未記載
35	1	21

表 5 90 日以内データ削除の記載を含む・含まないガイドライン
Table 5 Feature analysis result of data deletion within 90 days.

記載あり	未記載	対象外
0	36	21

の申告を適切に行うよう促すだけの記述や，暗号化ありの記載が別ページに見られる場合もあった．前者は当該 SDK の暗号化状況が未記載であり，後者はガイドラインページ自体に未記載であるため，両者を未記載と分類した．

3.2 データの削除リクエスト

3.2.1 ユーザデータ削除の記載

記載あり 35 個 (61%) に対して，完全に未記載のものは 21 個 (37%) と相当数発見した (**表 4**)．また，1 個のガイドラインは直接的な回答ではなくプライバシーポリシーの参照を促しており，当該プライバシーポリシーにおいて「削除申請可能」の記載を発見した．このようなガイドラインでは，申請自体は可能だがプライバシーポリシーの走査に手間がかかり，申請に時間を要する可能性がある．

3.2.2 90 日以内のユーザデータ削除の記載

表 4 で未記載に分類したガイドライン 21 個を除き，36 個を対象に調査した．結果，記載ありのガイドラインは 0 個であった (**表 5**)．未記載の 36 個中，特に「削除方法を提供しない」とあるものは 2 個であった．これらは 90 日以内のユーザデータ削除の有無を明記することで，本質問項目への回答を明確化すべきである．

いくつかのガイドラインに掲載された DS 申告フォームのスクリーンキャプチャには「90 日以内のユーザデータ削除」が存在していない点を発見した．本質問項目は 2023 年 4 月に新たに導入されたものであり [8]，その後，これらのガイドラインが更新されていない可能性が考えられる．

3.3 データの種類の選択

3.3.1 データカテゴリの記載

表 6 に結果を示す．C-完全は 27 個 (47%)，C-不完全は 10 個 (18%)，C-非公式追加は 6 個 (10%)，C-なしは 14 個 (25%) であり，公式 14 を過不足なく記載しているガイドラインは半数にも満たない実態が明らかになった．

C-非公式追加 6 個のうち 1 個は，公式のデータタイプ名をカテゴリ名として一部使用しており，作成者の不注意が疑われる．記載されたデータタイプ名は公式に準拠しているため，申告は可能である．また，2 個は公式に準拠したカテゴリ 14 個すべてに加えて独自カテゴリ “other data” と “data usage” を記載している．これらは公式 14 カテゴリ外であり，申告の必要性やどの公式カテゴリとして申告すべきかの判断がアプリ開発者に依存するため適切ではない．同様に残りの 3 個も独自カテゴリを含むが，収集・共有ありのカテゴリのみ掲載しており，公式 14 カテゴリの一部が未記載であった．よって，独自カテゴリが公式カテゴリに対応付く可能性も考えられたが，著者らの判断により対応付けは困難とした．独自カテゴリは，“ad interactions” と “usage data” である．本調査における公式表記と異なる記載の読み替えについては， 5 章で詳述する．

表 6 ガイドライン記載上のカテゴリ名と公式カテゴリ名の比較結果
Table 6 Comparison of guideline and official category names.

		データカテゴリ			
		C-完全	C-不完全	C-非公式追加	C-なし
データタイプ	T-完全	18	4	4	-
	T-不完全	9	6	2	-
合計	-	27	10	6	(14)

3.3.2 データタイプの記載

表 6 はデータカテゴリ記載の観点で分類した結果ごとに本観点の結果を示している．C-完全において T-完全は 18 個 (67%) であり，過半数はデータタイプを完全に記載している．一方，C-不完全では T-完全が 4 個 (40%) と低く，カテゴリが不完全なガイドラインはデータタイプも不足しやすい傾向が見られる．なお，4 個中 2 個は DS 申告フォームのスクリーンキャプチャを用いており，読み換え不要や情報欠落抑止の点からスクリーンキャプチャの有用性が見られた．最後に，C-非公式追加では，T-完全は 4 個 (67%) 存在した．割合は C-完全と同様に高い．これらのガイドラインは公式データタイプをすべて網羅した上で独自のデータカテゴリを追加している傾向がある．

C-完全かつ T-完全の 18 個はすべての公式データカテゴリとタイプに関する申告内容を網羅しており理想的である．一方，C-完全かつ T-不完全 9 個のうち 3 個は，どのデータタイプも収集・共有しないカテゴリについてデータタイプの列挙を省略しカテゴリ全体で否定している．T-完全より簡潔であり，より望ましいと言える．残りの 6 個は，あるカテゴリに対して一部データタイプが抜けており，一貫性がない．よって，問題なしと考えられるものは，C-完全の T-完全 18 個と T-不完全 3 個の合計 21 個である．

また，C-不完全と C-非公式追加の T-不完全のガイドライン，それぞれ 6 個と 2 個においても，カテゴリを一括で収集・共有を拒否，もしくはあるカテゴリについて部分的にデータタイプが抜けている場合のどちらかであった．

3.3.3 データの種類の記載方法

表形式は 51 個 (89%)，それ以外の 6 個 (11%) は文章や箇条書きであり，多くは表形式を用いていることが明らかになった．6 個中 4 個は C-なし，2 個は C-不完全と，両者データカテゴリの抜けが存在し，表形式以外の形式を用い

表 7 ガイドラインに記載された収集や共有動作を表す用語の種類数
Table 7 Number of terms repersenting data collection and sharing in the guidelines.

2 種類 (collect&share)	1 種類 (collect)	1 種類 (その他)
32	20	5

表 8 共有曖昧ガイドラインの分析結果
Table 8 Number of ambiguous data-sharing guidelines.

条件 1	条件 2	該当なし
19	25	13

表 9 収集曖昧ガイドラインの分析結果
Table 9 Number of ambiguous data collection guidelines.

条件 1	条件 2	該当なし	対象外
0	2	30	25

るガイドラインは公式 14 カテゴリを網羅できていない傾向が見られた．よって公式 14 カテゴリの網羅性の観点からは，表形式を用いることが望ましいと示唆される．

3.4 収集・共有の選択

3.4.1 収集・共有動作の記載の種類

2 種類のものが 32 個 (56%)，収集 1 種類のみのものが 20 個 (35%)，その他 1 種類のものが 5 個 (9%) であり，2 種類の動作が明記されたものは全体の約半分であった (**表 7**)．収集・共有の識別においては “collect” や “share” という直接的な表現がない場合があり，“off-device” や “with third parties” などの語句に基づき収集・共有の定義と照らし合わせて読み換えた．「その他」はそれぞれ “capture”，“track”，“transmit”，“receive”，“collect and transmit” の 5 種類の動作であり，収集・共有判定が困難であった．

3.4.2 共有曖昧ガイドライン (G_{SA})

表 8 に結果を示す．G_{SA} に該当したのは条件 1：19 個，条件 2：25 個の合計 44 個 (77%)，該当なしは 13 個 (23%) であり，多くのガイドラインで共有動作の有無の明瞭性に問題が見つかった．さらに，該当ありのうち 2 個は “Data sharing” と記載があるが，当該 SDK から他のサードパーティへのデータ共有についての説明であり，DS 申告の「共有」とは意味が異なっており，混乱を招く恐れがある．また該当なしのうち 2 個は，共有を否定しており判定基準に合致したが，当該 SDK がサービスプロバイダである旨の記載があった．DS 公式ではサービスプロバイダについては「共有」は該当しないため，「該当なし」と判定した．

3.4.3 収集曖昧ガイドライン (G_{CA})

記載動作が 1 種類かつその動作が「共有」と確かに判定可能なガイドラインはない (表 7) ため，記載動作 1 種類の 25 個は対象外とし，よって条件 1 の該当はない (**表 9**)．2 個が条件 2 に該当し，G_{CA} は少数ではあるが存在する．

表 10 一時的処理の記載を含む・含まないガイドライン
Table 10 Ephemeral processing feature analysis result.

記載あり	未記載	一部のデータについて記載
17	38	2

表 11 オプションのデータ収集とその目線の記載の分析結果
Table 11 Optional collection and perspective analysis result.

オプションの記載	開発者	ユーザ	両方	不明	対象外	合計
あり	19	9	18	3	4	53
なし	-	-	-	-	-	4

表 12 共有オプショナルガイドラインの分析結果
Table 12 Analysis result of guidelines with optional sharing.

該当あり	該当なし	対象外
2	30	25

3.5 一時的処理

記載ありが 17 個 (30%)，未記載が 38 個 (67%)，一部のデータについて記載が 2 個 (4%) であった (**表 10**)．記載ありの 17 個中 10 個と過半数で一時的処理を否定する記載があった．半数以上のガイドラインが未記載である実態が明らかとなったが，「一時的処理あり」のガイドラインが少数である点を踏まえると，未記載の場合は一時的処理の否定を示唆している可能性もある．また，一部のデータのみ記載のものは，収集が必ず行われるデータについてのみ一時的処理に関する記載があった．

3.6 オプションのデータ収集

3.6.1 オプションのデータ収集の記載

結果を**表 11** の最終列に示す．記載ありは 53 個 (93%)，記載なしは 4 個 (7%) であり，ほとんどのガイドラインはオプションのデータ収集について記載している．

3.6.2 オプションのデータ収集の目線の記載

3.6.1 項で「記載なし」に分類した 4 個は対象外とした．またオプション収集を否定する記載がある 4 個も対象外とした．結果を表 11 に示す．開発者目線が 19 個 (39%) と多く発見された．これらはエンドユーザ目線の情報が欠如しており，申告を困難にさせる．次に，エンドユーザ目線のものは，4 分類の中で唯一問題がないグループであるが，9 個 (18%) と少なかった．「両方」は 18 個 (37%) あり，そのほとんどは複数のデータタイプについて開発者目線のみの記載であるため，その部分について申告が困難となる．ただし，すべてのデータタイプについて両方の目線があり問題ない場合もあった．最後に「不明」は 3 個 (6%) あり，これらは記載が最小限で判定の手がかりとなる語句が得られなかった．これらの目線が開発者であると仮定した場合，誤った申告につながるため注意が必要である．

3.6.3 共有オプショナルガイドライン (G_{SO})

記載動作が 1 種類かつその動作が「共有」と確かに判定

表 13 目的記載形式と，目的に対応する収集・共有の区別，目的に対応するデータの分析結果
Table 13 Analysis results of purpose description formats, distinctions between corresponding collection and sharing, and data associated with purposes.

目的記載	観点	分類	データタイプ表 (26)	文章 (12)	目的専用表 (3)	その他 (4)	合計 (45+12)
あり	収集・共有の区別	可能・自明	8	1	0	1	10
		不可能	8	1	1	0	10
		対象外	10	10	2	3	25
	データとの対応	自明	26	5	1	2	34
		多対多	0	7	2	2	11
なし	-	-	-	-	-	-	(12)

可能なガイドラインはない (表 7) ため，記載動作 1 種類の 25 個を除外して調査した．結果，2 個が該当し，記載の修正が望まれる (**表 12**)．このうち 1 個は，オプションのデータ共有が開発者目線で記載されており，申告者がその点を理解できれば，DS 申告における「ユーザ目線のオプションのデータ収集」とは区別されるべきものであることが分かるであろう．残る 1 つはオプションのデータ共有がユーザ目線で記載されているため，「オプションのデータ収集」として申告すべき可能性が浮上するが，同時に G_{CA} に該当しており，収集の有無が不明瞭であるため，収集の申告そのものをすべきかどうかに疑問が残る．

3.7 収集・共有の目的

3.7.1 目的の記載

結果を**表 13** の最終列に示す．記載あり 45 個 (79%) に対して，完全に未記載は 12 個 (21%) であり，約 2 割のガイドラインには目的の記載がなく申告不可能であることが明らかになった．一部のデータタイプのみに目的を記載する場合も「記載あり」とし，以降の観点の調査対象とした．

3.7.2 目的の記載方法

目的の記載のあった 45 個を記載方法ごとに分類した結果を表 13 の表頭に示す．データタイプ表に組み込まれている場合が 26 個 (58%) と半数以上を占め，その後，文章形式 12 個 (27%)，目的専用の表 3 個 (7%) と続いた．目的専用の表は，最左列に目的名を列挙し，それぞれについて申告必要・不必要や，その詳細として当該目的を選択すべき場合の説明がある．「その他」は，データタイプごとの目的の記載に加えて目的専用の表を用いたもの，申請時のスクリーンキャプチャを掲載しているもの，全データタイプ一括で「共通の目的」と明記しているものがあった．

3.7.3 目的に対応する収集・共有の区別

結果を表 13 に示す．なお，記載動作が 1 種類のもの，または 2 種類かつ一方の動作を完全に否定するものは動作を区別する必要がなく，該当する 25 個を対象外とした．分析の結果，「可能・自明」と「不可能」はいずれも 10 個 (50%) であり，両者は同数であった．

目的の記載方法では，データタイプ表で対象外を除く 16 個中「可能・自明」は 8 個 (50%)，文章形式で対象外を除く 2 個中「可能・自明」は 1 個 (50%) であり，どちらの形式も半数のガイドラインが目的に対応する収集・共有の区別に問題がある．

表 14 ガイドラインに記載された目的名と公式目的名の比較結果
Table 14 Comparison of guideline and official purpose names.

分類	数	例
準拠	35	Crash logs: Analytics, App Functionality
一部不一致	5	IP address: Analytics and Reporting
完全不一致	5	IP address: To improve services
対象外	12	-

3.7.4 目的に対応するデータ

分析の結果，34 個 (76%) は「自明」であるのに対し，11 個 (24%) は「多対多」に分類され，その割合は無視できない水準であった (表 13)．目的の記載方法がデータタイプ表の 26 個はすべて「自明」となった．また，文章形式かつ多対多と分類した中の 1 個は，特定のデータタイプに関する目的の記述が散在していた．これは対応付けの曖昧さの問題に加え，開発者の見逃しによる申告漏れの恐れもある．また，文章形式や目的専用表を用いたものはいずれも過半数が多対多に分類され，本観点においてもデータタイプ表を用いる方が明確な情報を提供できる傾向が見られた．

3.7.5 公式に従った目的の記載

表 14 は分類結果と例を示している．分析対象は 45 個であり，目的の記載のない 12 個は対象外としている．結果，「準拠」が 35 個 (78%) と多く見つかった．例えばデータタイプ “Crash logs” は，公式に沿った目的 2 種類 (“Analytics” と “App Functionality”) を記載している．「一部不一致」は 5 個 (11%) であり，例えば “IP address” において公式の “Analytics” と公式には存在しない “Reporting” の 2 つを記載していた．「完全不一致」も 5 個 (11%) であり，例えば “IP address” において公式のいずれにも該当しない記載であった．この内，2 個のガイドラインでは例えば “ad performance” という独自記述に対して，公式には “Advertising or marketing” という比較的類似した目的が存在していた．しかし類似性の判断は各人に依存するため，本分析では読み替えせず不一致として扱った．

4. 議論

本研究の知見は次のような示唆を与える．まず，プラットフォーマーである Google は，プライバシーラベル開示ガイドラインの表記形式について統一的な規則を設けることが望ましい．DS 申告は質問項目が多く複雑であるため，ガイドラインの記載の曖昧さは開発者の理解や対応において解釈のばらつきを助長し得る．次に，SDK 提供者においては，現状ガイドラインの表記に関する統一的な規定は存在しないが，可能な範囲で公式の表記に準拠するなどの配慮が望ましい．最後にアプリ開発者は，ガイドラインにのみ依拠するのではなく，他の公式ドキュメントや補助的情報源とも照合しつつ，適切な申告を行うことが望ましい．

5. 制限と今後の課題

データカテゴリ，タイプの記載調査における「公式の表記と異なるが意味は同じ表記」について，その許容範囲は著者らが独自に決定した．許容される例として “Crash logs” と “Error reports” がある．ページの都合上，掲載はできないが，本調査を通して公式 14 カテゴリ中 12 に関して独自表現を 22 個，公式 38 データタイプ中 31 に関して独自表現を 71 個発見した．これらの妥当性の保証については今後の課題である．また，本研究におけるガイドラインの分析はすべて手作業で行っているため，誤りを含む可能性がある．これへの対処と再現性確保の観点から，使用したガイドラインデータの公開を検討している．

6. 関連研究

2009 年にプライバシーラベルという新たなコンセプトが提唱 [9] されて以来，十数年を経て，スマートフォンアプリの主要プラットフォームである Google Play や Apple は，それぞれ DS (2022 年 7 月義務化) と Privacy Nutrition Label (2020 年 12 月義務化) を導入した．本稿で扱う Google Play の DS に関しては，その申告結果が大規模に測定され不正確な情報表示の蔓延が指摘 [1] された．アプリ開発者に対する調査では，サードパーティ SDK のプライバシー慣行に関する情報不足が課題として明らかとなった [1]．DS 申告の特定の質問事項に注目した実態調査として，転送時の暗号化状況についての記載と実動作の不一致調査 [3] や，ユーザアカウントデータの削除リクエストに注目したコンプライアンス調査が実施された [2]．Google Play で公開されアクセス可能なアカウント削除リンクを持つ 494 アプリの内，Google Play のアカウント削除要件を満たすものは 8.5%とわずかである実態が明らかにされた．

本稿が対象とするプライバシーラベル開示ガイドラインに関して，記載内容と実際のプライバシー慣行を比較し記載漏れの実態を調査した研究 [4], [5] が実施された．ガイドラインの改善に向けて，より包括的かつ詳細なプライバシー開示メカニズムを実現するために，ソフトウェア部品表から着想を得たプライバシー部品表の概念が提案された [10]．また，本稿の観点の 1 つである「目的に対応するデータ」に類似した分析として，対応付けを 3 パターンの構造とその組み合わせに分類し，利用者数の多いプライバシーポリシーを分析した実態調査 [11] が行われた．

7. まとめ

本稿は Android サードパーティ SDK に付随して公開されるプライバシーラベル開示ガイドライン 57 個を分析した．DS 申告の手順に基づいて設定した 18 の観点から分析を行った結果，申告を困難にする複数の不備を発見した．今後の展望には，不備の自動検出・修正，ガイドライン公開元へのフィードバック及びその応答の分析が挙げられる．

参考文献

[1] Khandelwal, R., Nayak, A., Chung, P. and Fawaz, K.: Unpacking Privacy Labels: A Measurement and Developer Perspective on Google’s Data Safety Section, *33rd USENIX Secur. Symp.*, pp. 2831–2848 (2024).

[2] Yan, J., Liao, S., Ma, J., Aldeen, M., Kumar, S. and Cheng, L.: No Way to Sign Out? Unpacking Non-Compliance with Google Play’s App Account Deletion Requirements, *34th USENIX Secur. Symp.*, pp. 3277–3296 (2025).

[3] Sakuraba, Y., Inayoshi, H., Saito, S. and Monden, A.: Plaintext in the Wild: Investigating Secure Connection Label Accuracy for Android Apps, *Proc. 25th IEEE Int. Conf. Source Code Anal. Manipulation* (2025).

[4] Inayoshi, H., Kakei, S. and Saito, S.: Detection of Inconsistencies between Guidance Pages and Actual Data Collection of Third-party SDKs in Android Apps, *Int. Conf. Mobile Softw. Eng. Syst.*, pp. 43–53 (2024).

[5] Xiao, Y., Zhang, C., Qin, Y., Alharbi, F. F. S., Xing, L. and Liao, X.: Measuring Compliance Implications of Third-party Libraries’ Privacy Label Disclosure Guidelines, *Proc. ACM SIGSAC Conf. Computer Communications Secur.*, pp. 1641–1655 (2024).

[6] Google: Google Play SDK Index, `https://play.google.com/sdks`. (Accessed 2025-08-01).

[7] Google: Google Play Console, `https://play.google.com/console`. (Accessed 2025-08-06).

[8] Staffbase: Data Types for the Google Data Safety Section, `https://support.staffbase.com/hc/en-us/articles/4677057837074-Data-Types-for-the-Google-Data-Safety-Section`. (Accessed 2025-07-25).

[9] Kelley, P. G., Bresee, J., Cranor, L. F. and Reeder, R. W.: A “Nutrition Label” for Privacy, *Proc. Symp. Usable Privacy Secur.* (2009).

[10] Xiao, Y., Nadkarni, A. and Liao, X.: Enhancing Transparency and Accountability of TPLs with PBOM: A Privacy Bill of Materials, *Proc. Workshop Softw. Supply Chain Offensive Research Ecosystem Defenses*, (online), available from ⟨https://doi.org/10.1145/3689944.3696159⟩ (2024).

[11] 原　亨，長谷川彩子，Jamieson, J.，秋山満昭：プライバシーポリシーにおける収集データと目的の対応関係の実態調査，研究報告セキュリティ心理学とトラスト（SPT），Vol. 2024-SPT-54, No. 18 (2024).