



Sony
Interactive
Entertainment

2025年
2月

PlayStation

セーフティレポート



はじめに

ソニー・インタラクティブエンタテインメント (SIE) は、PlayStationを提供する企業として、プレイヤーの皆さんに革新的な体験をお届けすることを目指しています。組織を横断して活動する当社のセーフティチームは、全世界で1億人以上の月間アクティブユーザーを持つ PlayStation Networkにおいて、あらゆる背景を持つプレイヤーが安心して楽しめる、温かく支え合える環境を提供するために尽力しています。

ソニーグループ行動規範のもと、私たちはゲームコミュニティの世界を「感動」で満たし、プレイヤーの皆さんの安全を守るために取り組んでいます。その一環として、正確でわかりやすい情報提供を目指しています。

本レポートでは、プレイヤーの安全を守るための基本理念を支える3つの柱である「コントロール」、「保護・防止」、「施行・推進」をご紹介します。

PlayStationは、すべてのプレイヤーを歓迎します。そして、誰もが尊重され、すべてのプレイヤーにとって安全なオンラインコミュニティを育み、維持するために、皆さんのご協力をお願い申し上げます。

Catherine Jensen (キャサリン・ジェンセン)

ソニー・インタラクティブエンタテインメント VP、
グローバルコンシューマーエクスペリエンス



安全を守るための基本理念は 3つの柱に支えられています

- ◆ **コントロール**: 私たちは、プレイヤーとその保護者がオンライン体験を自分に合ったものに調整できるよう、さまざまなツールを提供しています。これらのツールには、コミュニケーションやプレイ時間のカスタマイズ、プライバシー保護、アカウントのセキュリティ確保などが含まれます。また、ファミリーマネジメント機能により、保護者が家族のために適切な判断を下せるようサポートします。
- ◆ **保護・防止**: 私たちは、プレイヤーをさまざまなオンライン上での危害から積極的に保護するため、高度な技術を活用しています。こうした脅威には、不適切なURL、攻撃的なテキストや画像、搾取的な表現、さらには児童性的搾取・虐待（CSEA）のような深刻なものがあります。プロアクティブ行動により、プレイヤーの体験に影響を与える前に、有害なコンテンツへの接触を最小限に抑えます。
- ◆ **施行・推進**: プレイヤーが行動規範を遵守し、違反者に対して適切な措置が取れるよう、経験豊富な人間のモデレーターを配置しています。また、プレイヤーからの積極的な報告や自動ツールによるレポートを活用し、人間のモデレーターが確認することで、行動規範に違反するコンテンツを特定しています。当社のポリシーは、表現の自由の尊重とプレイヤーを有害なコンテンツから守る必要性とのバランスを慎重に考慮し、公平性、明確性、一貫性のある施行ができるように制作されています。

コントロール

プレイヤーや保護者の方が、自分や家族にとって最適な選択をできるように。

プレイヤーは、PlayStation Network上で、「ブロック」、「ミュート」、「報告」といった機能を活用し、不快なやり取りを制限することができます。また、私たちはプレイヤーの声に耳を傾け、さらなる保護を提供するために改善を続けています。たとえば、2024年初頭には「デフォルトブロック機能」を導入し、「報告」機能を利用した際にその相手を自動的にブロックすることで、同じ不快な体験が繰り返されないようにしました。

さらに、PlayStation Networkの安全性を高めるため、新しい技術やポリシーの導入を積極的に検討・試験しています。その一例として、一部の国でプレイヤーの新規アカウント登録時に行われる年齢確認プロセスを試験的に導入しています。

私たちは今後も、すべてのプレイヤーにとって安全でフレンドリーなオンラインコミュニティを推進するため、オンラインセーフティの革新と向上に取り組み続けます。

PlayStation Networkでのプレイヤーのブロック

PS5パーティーボイスチャット

PlayStation Networkでの不適切な行為の報告



保護・防止

プレイヤーを守るための機能と技術。

私たちは、プレイヤーがログインするたびに、喜びとコミュニティとのつながりを感じられる体験を提供したいと考えています。すべてのプレイヤーは、安全に、そして安心して自分らしくゲームを楽しむ権利があります。SIEでは、プレイヤーがポジティブで楽しい体験を得られるよう、さまざまな安全対策を導入しています。その中でも特に表立った対策の一つが「不快な言葉やヘイトスピーチ表現を防止するフィルタリング機能」です。このフィルターは多言語に対応しており、オンラインIDや「自己紹介」欄などのプレイヤー同士が目にする場所で、攻撃的または

不適切な言葉をブロックすることで、有害なコンテンツの拡散を防ぎます。

また、あまり目立たないものの同様に重要な保護対策として、高度なスパムブロック機能を導入しています。機械学習ツールを活用し、技術を悪用してプレイヤーに向けて大量の有害なメッセージを送信しようとする行為などを検出します。さらに、コミュニティをより安全に保つために、独自のボットを運用して、プレイヤーの妨害を目的とする悪性ボットを検出し、削除しています。

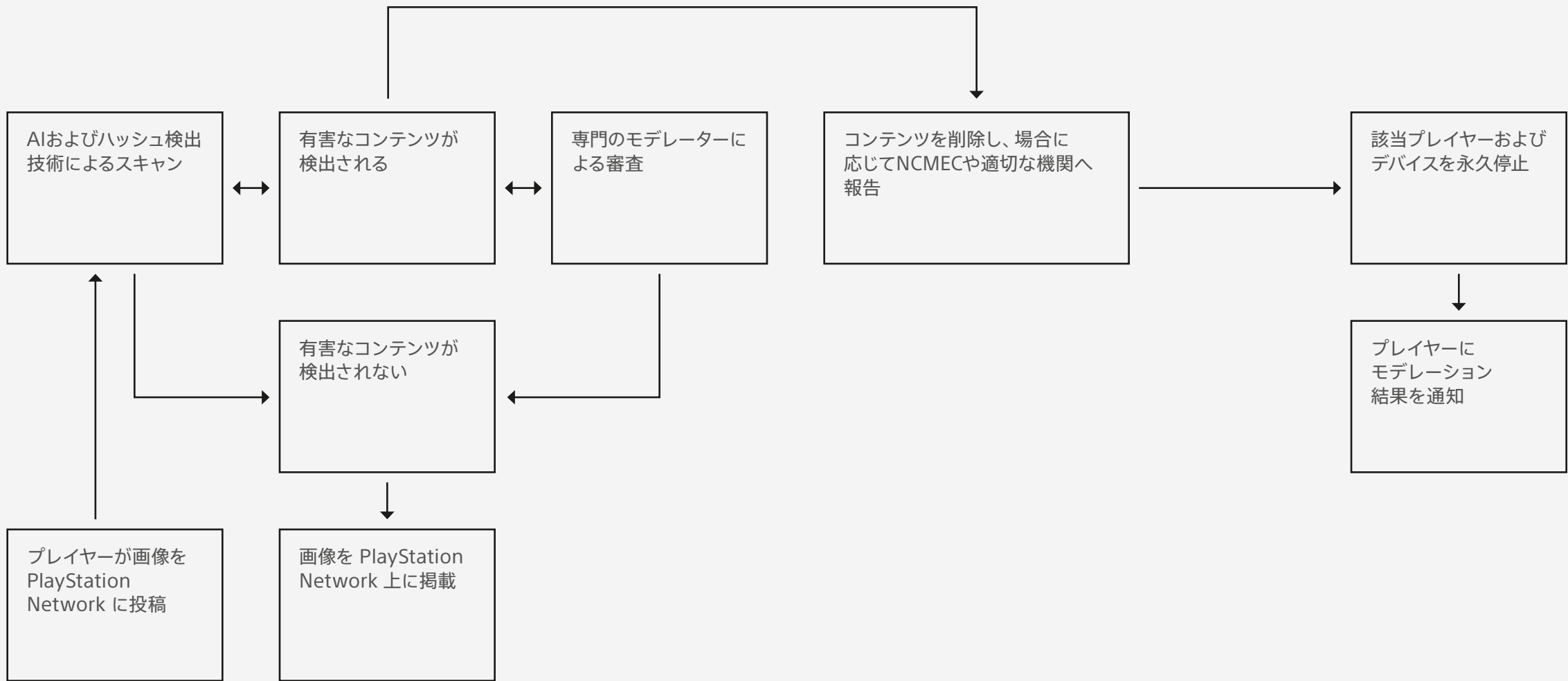
PlayStation Networkでは、オープンなコミュニケーションとゲーム関連コンテンツの共有を促進するためにリンクの共有を許可していますが、警戒も怠りません。行動規範に違反するURLは迅速に特定・制限され、プラットフォーム全体の安全性と敬意ある環境を維持しています。行動規範に対する違反が確認されたURLはブロックリストに追加され、それ以降PlayStation Network上では表示されなくなります。ヘイトコンテンツ、ポルノ、極端な暴力を含むサイトは、PlayStation Networkの行動規範に違反するため、厳しく禁止されています。

オンラインゲームの醍醐味のひとつは、他のプレイヤーと協力したり、時には競い合ったりすることです。ゲーム内で難しい課題を乗り越え、スキルを磨くことで得られる強い達成感は、プレイヤーにとって大きな喜びとなります。プレイヤーは健全な競争を歓迎しますが、不正なソフトウェアを使って不当に有利となるような相手とは対戦したくないはずです。そこで、当社のシステムではチートツールの使用を積極的に検出し、不正行為が確認されたアカウントには一時停止などの措置を講じ、公正でバランスの取れたゲーム環境を維持しています。

PlayStationは、一人でも、友達とでも、誰もが楽しく遊べる場です。プレイヤーの安全は最優先事項であり、PlayStation Networkでは子どもたちを守るためにさらなる対策を講じています。

私たちは、Family Online Safety Institute (FOSI)、全米行方不明・被搾取児童センター (NCMEC)、Tech Coalitionをはじめとする国際的に評価の高い団体と協力し、プレイヤーの安全を推進しています。この協力には、FOSIと

画像検出の流れ



連携した若年層向けのオンラインセーフティに関する教育やデジタル・シチズンシップ(良識あるインターネットの利用)の推進活動から、NCMECやTech Coalitionとの共同による、CSEA(児童性的搾取・虐待)素材を含む有害コンテンツの検出・削除ツールの開発まで、さまざまな取り組みが含まれます。

コミュニティの安全を確保するため、PlayStation Networkにアップロードされるすべての画像は、ハッシュマッチング技術を使用してスキャンされ、潜在的な有害画像との一致が検出されるようになっていきます。ハッシュマッチングツールは、コンテンツに「ハッシュ」と呼ばれる固有のデジタル署名を付与し、既知の署名のデータベースと照合します。一致が検出

された場合、そのコンテンツは専門のモデレーションチームにエスカレーションされ、審査が行われます。そのコンテンツが児童性的虐待素材(CSAM)であると確認された場合、NCMECへ直接報告されます。さらに、ハッシュマッチングに加え、未報告のCSAMを特定できる機械学習技術も導入しています。検出手法は異なりますが、モデレーションのプロセスは統一されており、CSAMの検出は非常に重要であるため、フラグされたコンテンツはすべて専門家による徹底的な審査が行われます。違反が確認された場合、該当アカウントおよび関連するPlayStationコンソールは、ネットワークで永久に利用停止となります。



施行・推進

熟練したモデレーション、プレイヤーとのコミュニケーション、および異議申し立てプロセス。

モデレーションでは、PlayStation Network上でプレイヤーから報告されたコンテンツが行動規範と照らし合わせて審査し、違反が確認された場合に適切な措置を適用します。モデレーションチームは、コンテンツ評価の専門家で構成されており、年中無休で24時間体制で運用されています。報告を受けたメッセージ、ボイスチャット、プロフィールの内容、Game Baseのグループ名、オンラインID、ゲーム内のコミュニケーションなどについて、迅速に審査を行います。

モデレーションは30以上の言語に対応しており、文化的背景や歴史的な配慮事項、各地域の法規制を考慮しながら、公平で敬意ある施行を徹底しています。すべての報告は慎重に取り

扱われ、報告されたアカウントに対する措置は人間のモデレーターによってコンテンツが確認され、その後を実施されます。

私たちは、施行措置を明確にプレイヤーに伝えるだけでなく、PlayStation Network上での不適切な行為がもたらす影響や、一時的または永久的なアカウント停止などの処分の可能性についてプレイヤーに積極的に啓発しています。また、PlayStation Networkが提供されている40カ国と地域において、永久停止処分に対する異議申し立ての機会をプレイヤーに提供しています。今後は一時的な停止処分にも異議申し立ての適用範囲を拡大するとともに、さらなる地域に展開する予定です。

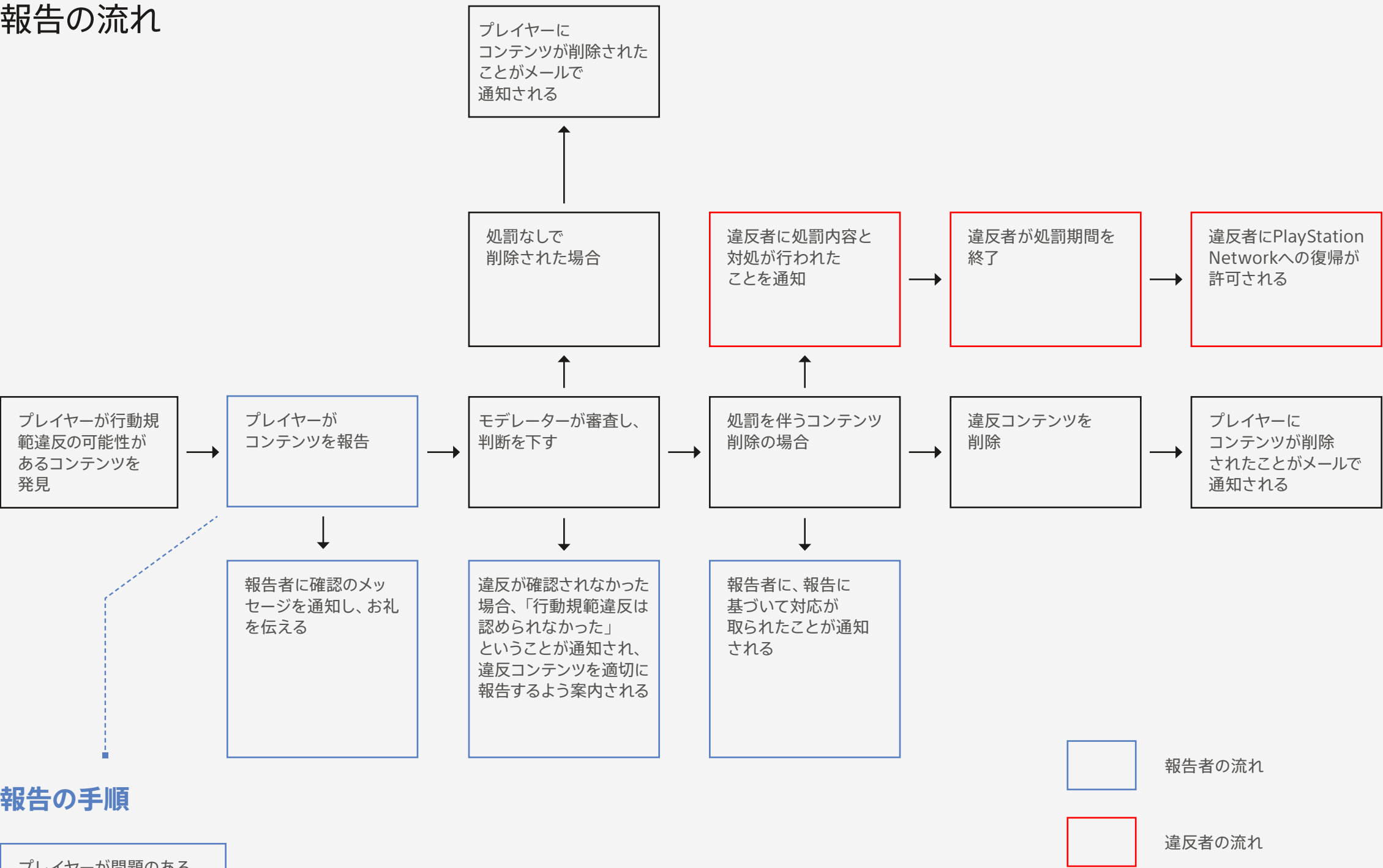
PlayStation Networkにおけるアカウント停止措置

モデレーターの判断は、包括的な教育的枠組みに基づいて行われるため、一貫性と公平性のある行動規範違反への対応ができます。アカウント停止の種類は、違反の重大度やプレイヤーの過去の処分歴に応じて決定されます。たとえば、初回の軽度な違反であれば警告にとどまる場合がありますが、繰り返しの違反や重大な規約違反があった場合は、30日間、60日間、または永久的なアカウント停止が下されることがあります。一般的なモデレーションの決定では、警告、3日間・7日間・30日間・60日間・永久的なアカウント停止という措置が適用されます。また、停止措置がプレイヤー同士のコミュニケーション機能に限定される場合もあります(PS5のみ)。しかし、重大な違反行為が確認された場合は、コンソール自体のネットワーク接続を禁止することもあります。

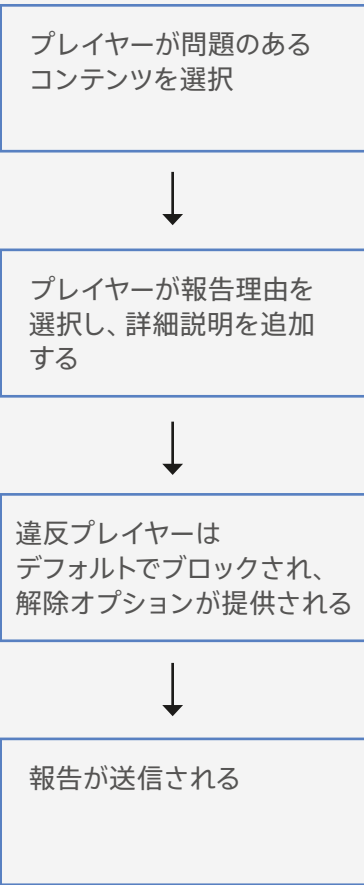
報告書を提出したプレイヤーには、受領の確認が行われ、報告されたコンテンツが行動規範に違反しているかどうかを含め、モデレーションの決定が下された時点で通知が送られます。違反が確認されなかった場合は、今後有効な報告ができるように、教育的リソースが提供されます。

行動規範違反が認定されたプレイヤーには、メールで通知されます。

報告の流れ



報告の手順



PlayStation Network における永久アカウント停止への異議申し立て方法

永久アカウント停止処分を受けたプレイヤーには、専用のWebフォームを通じて異議申し立てを行う機会が提供されます。プレイヤーが異議申し立てを提出すると、特別な研修を受けたモデレーターがプレイヤーの違反状況や履歴を審査します。モデレーターが異議申し立ての結果を決定した後、プレイヤーに審査結果が通知されます。

PlayStation Networkにおけるアカウント停止への異議申し立て

プレイヤーの安全と犯罪関連の

エスカレーション に対する当社の取り組み

当社のモデレーションチームは、違反者のアカウントへの対処に加え、違反内容が命への脅威、プレイヤーの安全、またはその他の違法行為に関連する場合は、さらなる審査のためにエスカレーションする権限を持っています。必要に応じて、法執行機関、関連する政府機関、または適切な当局へ通報する場合があります。

私たちは、CSEA（児童性的搾取・虐待）に対してゼロ・トレランス（いかなる違反も許容しない）方針を貫きます。特定されたすべてのCSEA事案は、NCMECおよび法執行機関に同時に報告されます。

人や財産に対する脅迫行為や自傷行為の意図が含まれる報告は、複数の段階にわたる審査プロセスを経て精査されます。脅迫の信憑性が高いと判断された場合は、直ちに法執行機関へ通報されます。

CSEAの検出・削除・報告件数（年間）

年	CyberTips（通報）件数
2021	2,071
2022	4,102
2023	3,974
2021年から2022年と2023年にかけて、SIEによる通報件数が増加したのは、搾取的なコンテンツの検出技術を強化したことによるものです。	

より分かりやすくなった 行動規範

行動規範は、プレイヤーがPlayStation Network上でポジティブで敬意あるゲームコミュニティを維持するための指針となるものです。このガイドラインを、より理解しやすく、世界中で一貫性のあるものとし、すべてのプレイヤーがアクセスしやすいよう、単純化して更新しました。オンラインプレイを始める前に、すべてのユーザーが最新のルールを確認し、同意する必要があります。また、プレイヤーの皆様には、定期的に行動規範とヘイトスピーチに対するポリシーを見直し、適切なゲーム内行動について理解を深めることを推奨しています。

ヘイトスピーチに対するポリシー

PlayStationでは、ゲーム内およびオンラインの両方で、包摂性（インクルージョン）と個性の尊重を推奨しています。

PlayStation Networkは、多様な背景を持つプレイヤーが集うコミュニティであり、楽しいゲーム体験を提供する場です。

プレイヤー同士が自由に自己表現をし、意見を交わすことは歓迎されますが、ヘイトスピーチや憎悪を助長する行為は一切許容されません。プレイヤーは、アカウントを作成する際に、PlayStation Network上での適切な振る舞いについてまとめた「行動規範」に従うことに同意します。プレイヤーの皆様は、互いを受け入れ合い、適切な行動を心がけ、不適切な行為をSIEに報告することで、コミュニティをサポートすることができます。PlayStation Networkでは、どのようなヘイトスピーチや差別的な表現も許容されません。たとえ軽い気持ちで、真剣ではなく、あるいはリアプロプリエーション（意味の取り戻し）でも例外ではありません。一部のプレイヤーが「友人同士でふざけた」、「冗談で罵りあった」だけで問題ないと考えたとしても、PlayStation Network上では不適切なコンテンツや表現は処罰の対象となり、行動規範が適用されることを理解してください。



プレイヤーとモデレーターのサポートおよびウェルネス

私たちは、PlayStationコミュニティのウェルネスに深く取り組んでいます。そのため、プレイヤーとのつながりを強化し、あらゆるサポートを提供できる、新たな機会を常に模索しています。

モデレーションチームは、あるプレイヤーが困っているようだという報告を受けた場合、そのプレイヤーにメッセージを送り、具体的な支援策を案内します。

また、私たちは「Crisis Text Line」、「Shout」、「あなたのいばしょ」、グローバル対応の「ThroughLine」と提携し、24時間365日、専門のボランティアによるライブメンタルヘルスサポートおよび緊急時対応を提供しています。

また、セーフティモデレーションチームからのメッセージがなくても、プレイヤーコミュニティがこれらのパートナーを利用し、必要な時、困った時に専門家のカウンセラーに相談することも可能です。

さらに、米国および英国のプレイヤーは、専用キーワード「CIRCLE」を使用することで、モデレーターを介せずに「Crisis Text Line」や「Shout」へ直接連絡することができます。キーワード「CIRCLE」は、PlayStationの象徴であるだけでなく、サポートの輪を表しています。

提携するサポートパートナーはすべて、完全な機密保持を徹底しており、プレイヤーや他者の安全を確保するために必要不可欠な場合のみ、緊急サービスに情報を伝えます。これらのサービスに関する詳細や個人情報の取り扱いについては、各パートナーの利用規約をご確認ください。

モデレーターのウェルネスも、SIEにとって重要な優先事項です。2024年2月、SIEのオンラインセーフティチームは、モデレーターのウェルネスと安全対策における取り組みが評価され、「Great British Workplace Wellbeing Award」の最優秀戦略賞を受賞しました。

