

第3 調査の概要

1 調査要領

インターネット実態調査フロー

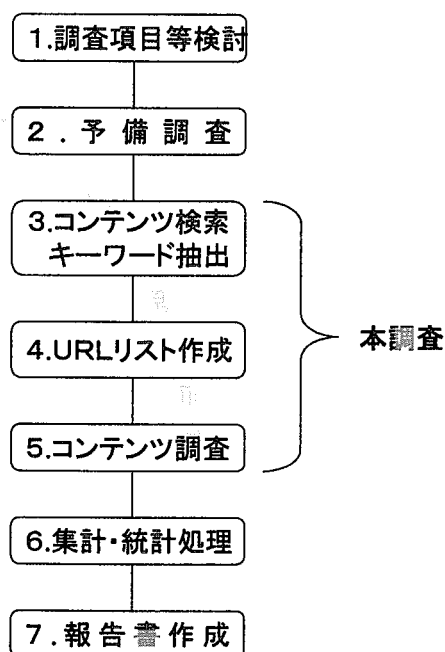


図5 インターネット実態調査フロー

(1) 調査項目等検討

本調査を実施するにあたり、調査対象となる違法・有害情報を次のように定めた。

- ・ **違法情報**：情報の発信自体が違法であるもの及び犯罪が行われている疑いのある情報
(例：わいせつなポルノ、売春の公告、詐欺、悪質商法、マルチの疑いのある公告等)
- ・ **有害情報**：違法情報には該当しないが、犯罪や事件を誘発する恐れのある情報
(例：誹謗中傷、毒劇物・架空銀行口座の販売等)

次に、これら違法・有害情報を大きく8つのカテゴリーに区分した。

1 薬物事犯

覚せい剤・麻薬・向精神薬等の販売等に関するもの

2 銃器

銃器・実包等の販売に関するもの

3 悪質商法事犯

詐欺の疑いのある広告等マルチ商法・ねずみ講の勧誘の疑いのあるものなど

4 賭博、富くじ、ノミ行為

賭博・富くじ・ノミ行為に関するもの

5 著作権違反

著作権法違反・商標法違反等知的所有権侵害事犯に関するもの

6 風俗事犯（性を売りものとするもの）

性風俗特殊営業・テレフォンクラブ等の広告、売春の疑いのある広告・掲示板

7 名誉毀損、脅迫、個人情報

名誉毀損・脅迫・侮辱・誹謗中傷その他個人情報を流布するもの

8 犯罪の誘発等に関するもの

銀行口座番号・クレジットカード番号等の販売、コンピュータウィルスの配布販売、ハッカーツールの配布販売、ハッキング手段の紹介、レイプ・痴漢等の性犯罪手口・体験談の紹介、錠前破り・すり等の手口・体験談の紹介、爆弾・毒物等の製造方法等犯罪を助長する恐れのあるもの、自殺のそそのかし・自殺の方法の教示等

また、各カテゴリーに共通する項目として、

URL関係、注文方法、引き渡し・入手方法、代金決済方法、販売価格、販売・提供の有無、掲示板の有無、リンク集の有無も併せて調査する。

2 予備調査

本調査を実施する前に、効率よく作業を進めるため、また本調査に向けた問題点の抽出等を検討するため、予備調査を実施した。

予備調査では、保秘に留意すべき観点から、独自のサーチエンジンを利用した調査は避け、既存のサーチエンジンを利用した調査に留めることとした。

本調査では、8つのカテゴリーを対象とするが、予備調査では、このうち、薬物事犯を対象にURLの収集と内容調査を行った。

予備調査の方法であるが、薬物事犯の分野についてのキーワードを設定し、そのキーワードを元にサーチエンジンを使用し、URLの抽出を行った。抽出したURLについては、サーチエンジンでのキーワードへの適応度を示すスコア順に、明らかに調査対象に該当しないURLの除去、関連URLの確認などを手動で行い、最終的に、有効となるURLリストの作成を行った。

なお、対象URLとして挙げるのかどうかの判断については、違法行為を行う、もしくは違法物品のやりとりの支援を行いうるページかどうかを判断基準とした。

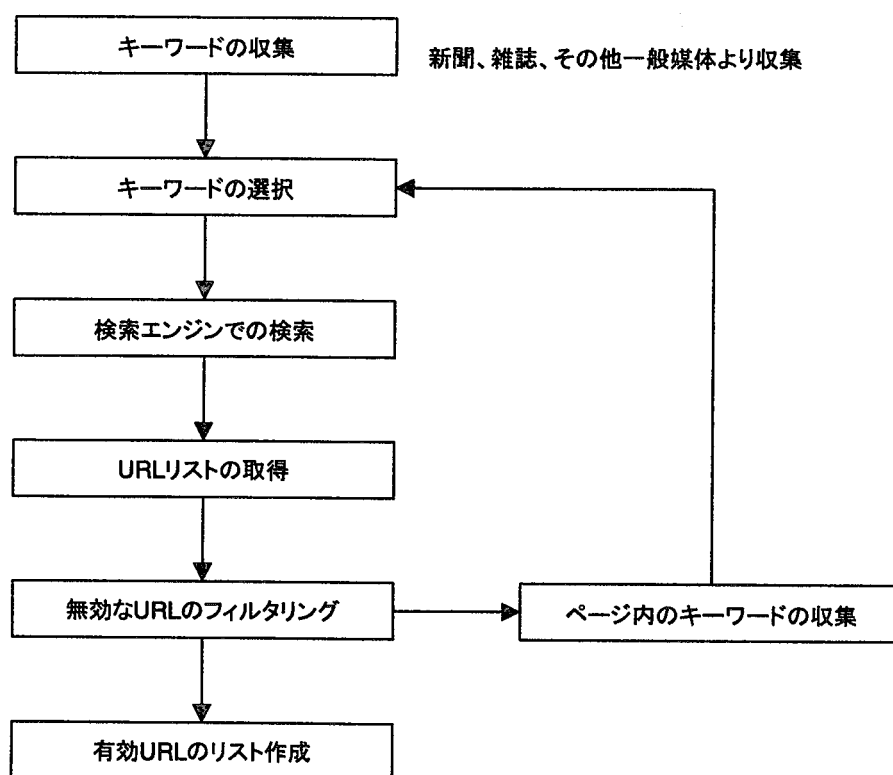


図6 予備調査フロー

薬物事犯として、「コカイン」をキーワードとした場合の具体例を挙げる。

サーチエンジンを利用した場合、「コカイン」などのキーワードのみを利用して検索を行った場合、その多くは薬物そのものの特徴の説明、社会的な現状、薬物関連のゴシップ記事などのページが高スコアページとしてヒットした。図7にサーチエンジン infoseek^{*1}を利用して検索を行った結果を示す。

「コカイン」のみを検索キーワードとしたとき、上述の infoseek で1,966件、その他のサーチエンジンでは goo^{*2}で1,211件、lycos^{*3}で2,055件、excite^{*4}で1,541件がヒットした。また、infoseekによって検索を行い、表示された結果のうち検索の高スコア100件の内容について調査を行ったところ、ほとんどがまったく危険性のない芸能ゴシップ、社会問題として取り扱ったもの、薬物そのものについての学術的説明を示したページであった。結果として、違法・有害情報を含むURLとして該当するページは、「合法ドラッグ」「マッシュルーム」などの輸入販売の代行を行っている業者などが3件あっただけであった。



図7 infoseekを利用した検索結果

単純なキーワードによる検索の場合、違法性の高いページの該当率は約3%程度であった。最も多く見つかったページは公共機関・団体等による情報提供であり、その割合は全体の約2割程度だった。以上のことから、違法・有害情報を含むURLの取得には、単純なキーワードを選択して機械的に検索する方法だけでは不十分であることが分かった。このことは、一般の人々がサーチエンジンを利用して、直接的に違法性の高いページを発見することが難しいことを示している。

この原因としては、サーチエンジンから取得可能なURLとその詳細情報だけでは、当該ページの内容の把握が十分でないことが挙げられる。通常、情報発信者側では、サーチエンジンによるURLの取得を行わないように指定することが可能であったり、また、違法・有害コンテンツでは、あからさまな文言を使用せず、伏せ字や俗語を用いるため、通常のサーチエンジンでは、該当ページ自体を抽出することが難しい。このため、本調査に有効な違法・有害情報を含むURLの取得のためには、より有効なキーワードの選択が重要であると同時に、効率よく有効なページの取得が必要となる。このため、

目視による内容の評価も行う必要が生ずる。

目視による評価は、1 キーワードあたり 300 から 500 ページの内容を確認しなければならぬため、調査結果の報告を速めるためにも、処理の迅速化を行う必要がある。

また、予備調査では、取得した URL の内容を確認すると関連するページ相互にリンクが張ってあったり、当該ページ自体は違法・有害情報でないものの、問題を含んでいる恐れのある URL を集めたリンク集が多く存在することが分かった。これらリンク集のページを基点として検索ロボットを利用した調査を行うことで、よりの絞った調査が行えるものと考えられる。

上記方法で薬物関連について試行した結果、通常のサーチエンジンによる検索では発見が難しい 8 つの URL *⁶ の取得を行うことができた。

本調査では、本調査用に最適化した独自のロボット型サーチエンジンを利用するため、より効率的な URL の取得が可能であると思われる。

また、当該エンジンを利用することにより、直接的な違法性はないが、その分野のいわゆるアンダーグラウンドな総合情報、用語辞典的なページも発見でき、これらのページから、さらに、より有効なキーワードの抽出を行うことができる。

(参考 URL)

*1 : <http://www.infoseek.co.jp/>

*2 : <http://www.goo.ne.jp/>

*3 : <http://www.lycos.co.jp/>

*4 : <http://www.excite.co.jp/>

*5 : <http://www.globe-jp.com/2nd-tripper/index01.htm>

http://plaza6.mbn.or.jp/~top/index_03.html

<http://www4.osk.3web.ne.jp/~arkcg/index.html>

*6 : <http://www.bekkoame.ne.jp/~taimado/index.html>

<http://www.broadway.or.jp/~mk2/index.html>

http://home4.highway.ne.jp/tk-net/p_sickdrug_end_TK.html

<http://soma.system.to/>

<http://www.INSIDETheWeb.com/mbs.cgi/mb630542>

http://members.tripod.com/flashback_jp/fuck_the_system.htm

<http://www.bekkoame.or.jp/~alteredim/>

<http://lavender.fortunecity.com/shorty/153/med.html>