

インターネット通信販売価格の把握に関する 諸外国の取組状況

諸外国の取組状況

イギリス

- ・ウェブスクレイピングにより収集したデータを用いたCPI試算(食料品)

オランダ

- ・新しい価格収集方法(スキャナデータ、ウェブスクレイピング、ロボットツール、レジスター)の利用
- ・公式統計の情報源としてインターネットを利用。航空運賃、衣料品等でデータ収集、検証・分析

ノルウェー

- ・国内のネット通販業者からパーソナルケア用品のデータを収集し日次指数を算出
- ・スクレイパーとしてimport.ioを使用

オーストリア

- ・2年間にわたるプロジェクトとしてインターネット自動価格収集を実施
- ・スクレイパーとしてimport.ioを使用
- ・商品及び店舗は、現行の手作業による価格収集手順によって選択
- ・航空運賃の指数作成で作業時間を大幅に削減できることを確認(16時間→2時間)

ニュージーランド

- ・オンラインデータに関する最初の研究として、ウェブスクレイピングにより国内の主要な小売業者から収集した家電の日次ネット価格データを使用して、日次及び月次の非加重の指数を算出

イタリア

- ・データ源としてのインターネットの利用を拡大する方針。実験的活動として家電製品、航空運賃

各国の状況：イギリス

イギリス (ONS (Office for National Statistics), BIG Data Project)

【概要・背景】

- ・33品目 (CPI品目) について毎日約5,000価格を取集
- ・ウェブスクレイピングで取集したデータから実験的な物価指数を作成
 - －商品分類に教師あり機械学習技術、異常検出にクラスター分析と呼ばれる教師なし機械学習技術を使用
 - －「食料及び非アルコール飲料」「アルコール飲料」のGEKSJ指数、日次連鎖指数及び単価指数を作成・比較
 - －ウェブスクレイピングで取得したデータから作成した指数と公表CPIを比較
- ・価格に関する試験プロジェクトは、2014年1月に発足したビッグデータプロジェクトの1つでウェブスクレイピング技術を用いて価格を取集するもの
- ・2015年に関連研究を2本発表
 - －「Trial consumer price indices using web scraped data」では2014年6月～2015年4月の試験的物価指数を提示
 - －「Research indices using web scraped data」では時系列を2015年6月まで延長、GEKSJ算式を用いた指数を導入
- ・様々なデータ源の特性
- ・「データ・ラングリング・プロセス」を含む、データ取集から指数作成までのプロセス
- ・教師あり機械学習を用いた分類

各国の状況：イギリス

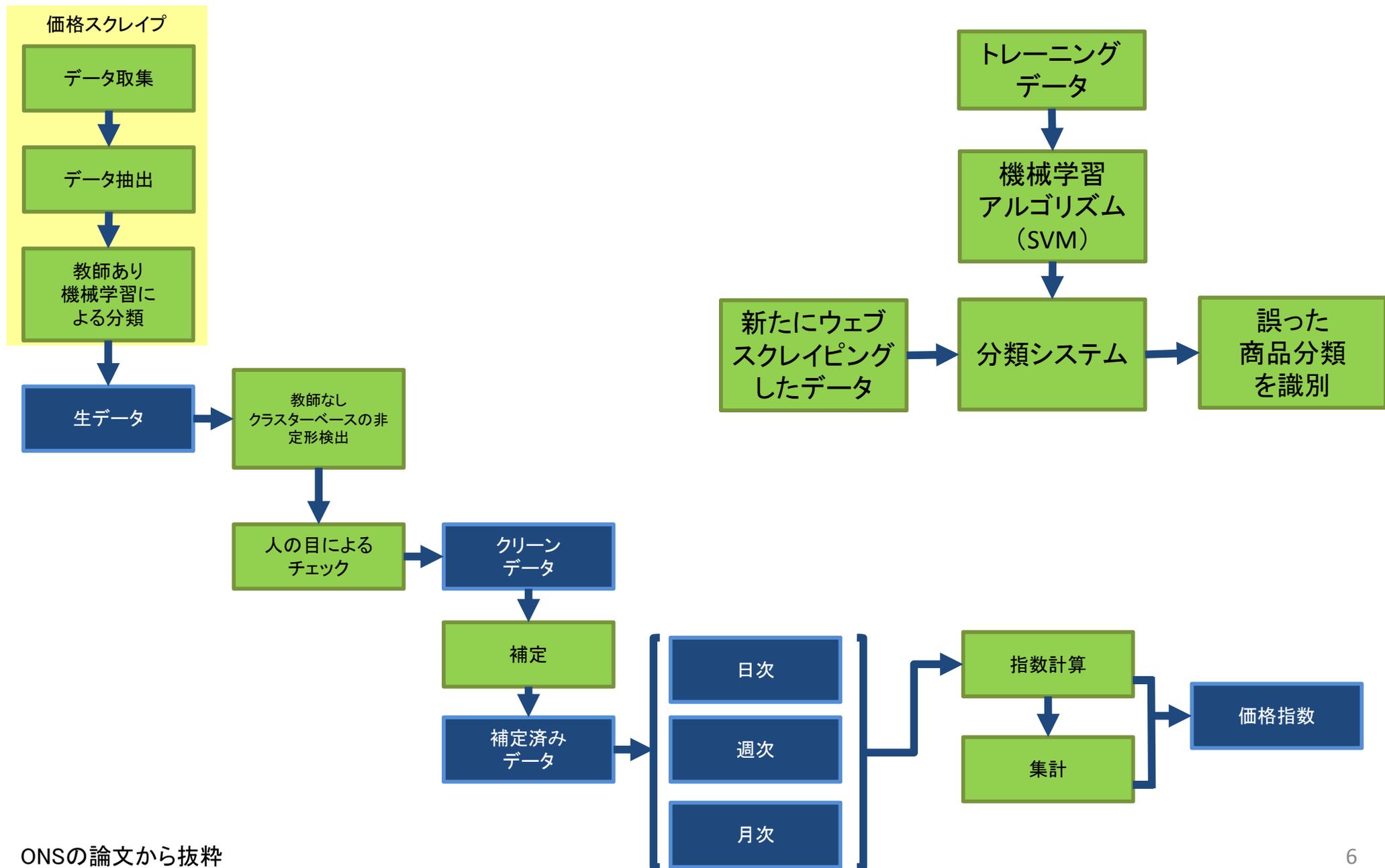
【スクレイピングに係る課題（制限事項）】

- ・ウェブスクレイピングによる価格データは、従来の形式で収集した価格データと異なる。
- ・一般的な制限事項
 - －ブッキング技術使用のサイトもあり、全てのオンライン小売業者から価格をスクレイプできるとは限らない。
 - －どれだけの数量が購入されたか分からないため、全商品を同等に扱う必要がある。
従来の形式では、価格調査員が代表的な品目を選択するため、支出額が少ない商品は考慮されなかった。
 - －価格の収集が毎日実施されるため、公表されるCPIと比較できる範囲が限定される。
 - －商品名が時間の経過で変化する場合がある。代替品を容易に見つけられない。
 - －回転率が高い(商品が出現した後に在庫切れする)商品があり、一部の品目でサンプルサイズが小さくなる。
時系列全体で品目を照合する必要がある単価指数を作成する際に、深刻な問題となる。
 - －データが大量であるため、従来のクリーニング方法が機能しない。
- ・ONSのプロセスの制限事項
 - －価格が大規模店舗3店からしかスクレイピングされていない。CPIでは全国のさまざまな店舗で収集される。
 - －一部のサイトが無限スクロールシステムを有しており、最初のページの24商品に収集が限定された。
(この問題は最近解決した。)
 - －無限スクロール問題と同様に、ウェブサイト構造の変更ならびにスクレイピングに使用するコンピュータ及びサーバーの技術的不具合も「スクレイパー中断」と呼ばれるデータ収集途絶を引き起こす可能性がある。

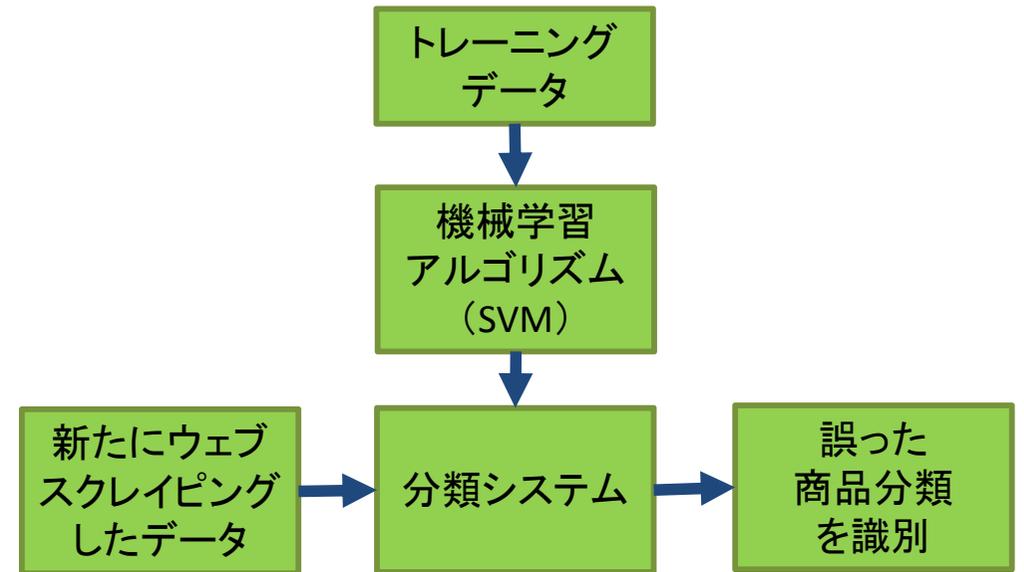
●様々なデータ源の特性

特性	従来型	ウェブスクレイピング	スキャナデータ
収集方法	大部分が現地収集され、ウェブサイト、カタログおよび電話問い合わせによる手作業の中央集中型収集で補完される。	小売業者ウェブサイトの自動ウェブ・スクレイピング。データ収集は、局内収集の場合もあれば、第三者から提供される場合もある。	購入時点で店舗で収集される。小売業者から直接提供される場合もあれば、第三者を介して提供される場合もある。
収集場所	多数の小売業者の全国の多数の事業所(小規模から大規模まで)	オンライン小売業者からの価格。通常、中規模から大規模の小売業者のみ。	取引データを取集している小売業者からの価格。通常、中規模から大規模の小売業者のみ。
タイミング	月のある1日(指数調査日)を中心として価格が取集される。	毎日価格が取集される。	取引のたびに、価格と数量が取集される。集計された形で提供されることもある。
サンプル	価格統計調査員が代表的な商品を選択する。	全商品について価格がスクレイピングされる。	全商品についてすべての取引データが取集される。
チェック	人の目によるチェックおよび自動検証	価格および分類を人の目でチェックするのは、非現実的。機械学習などのデータサイエンス技術を利用する。	価格および分類を人の目でチェックするのは、非現実的。機械学習などのデータサイエンス技術を利用する。
時系列	期間を通して、商品を照合する。	商品の入れ替わりが激しい。	商品の入れ替わりが激しい。
コスト	高コスト	1価格見積当たりのコスト削減の可能性があるが、処理、維持および保存のコストを考慮する必要がある。第三者からデータを取得するコストがかかる可能性もある。	1価格見積当たりのコスト削減の可能性があるが、処理、維持および保存のコストを考慮する必要がある。第三者からデータを取得するコストがかかる可能性もある。
規模	小規模から中規模までの収集	大規模な収集	大規模な収集
適時性	データ収集から約1か月後に発表	ウェブ・スクレイピングが局内で実行される場合、処理速度にもよるが、データ収集とほぼ同時に物価指数を算出することができる。	不明。第三者の供給業者との契約によって異なる。

● 食料品価格指数の計算方法



● 教師あり機械学習を用いた分類のフロー



各国の状況:オランダ

オランダ (Statistics Neterlands:オランダ統計局)

【概要・背景】

- ・2002年にスキャナデータ、2012年にウェブスクレイパーによるデータ収集を開始
- ・航空運賃、衣料品等でデータ収集、検証・分析
- ・衣料品のデータ収集の2つの理由
 - －店頭でのデータ収集を可能な限り減らしたい。
 - －品質訂正の必要がなく、プロセスの自動化が容易

■航空運賃(スクレイピング検証のためのケーススタディ)

- ・スクレイピングによる収集と人手による収集を比較し、効率性について検証(作業時間等)
- ・航空会社のサイト変更の頻度と再プログラミングに係る時間等を分析

■衣料品

- ・ロボットが毎日、複数の小売業者のオンライン店舗から、全ての商品情報(分類特性含む)及び価格をウェブスクレイピングにより自動収集
- ・モニタリング、後処理、転送、保存

各国の状況:オランダ

【データ収集手法の比較】

方法	スキャナデータ	ウェブスクレイピング	ロボットツール
データ数	多い	多い	少ない
小売業者の数	少ない	少ない	多い
回答の負荷	あり	なし	なし
数量に関する情報	あり	なし	なし
価格の観察	月の大半	月の大半	1商品提供あたり1回
分類	小売業者が提供するメタデータ	ウェブサイト上の商品説明	局内での商品提供の選択

オランダ統計局の論文から抜粋

【課題】

- ・衣料品市場全体をインターネット・データ収集でカバー可能かを見極め、その方法を見つけ出すことが今後の研究プログラムの1つ。
- ・膨大なデータ処理のため、ルールのメンテナンス、分析を可能な限り自動化しなければならない。
- ・分類ルールのメンテナンスの自動化: 研究部門で機械学習アルゴリズムの利用等を調査
- ・衣料品の写真に有用な追加情報が含まれている場合があることから、写真を利用する研究も開始
- ・すべてのデータが同質とは限らない(収集データに異なるブランドが多数存在)。
- ・ヘッドニック法などの手法を適用する場合、より多くの追加情報(商品情報)を抽出する必要があるが手作業で行うことができないため、より洗練された自動抽出および分類技術が必要となる。

●インターネット通信販売価格の把握に関する諸外国の取組状況

対象国	対象品目	主な取組(今後の予定含む)	手法等	スクレイピングの利点(プラス側面)	スクレイピングの欠点(マイナス側面)	CPiへの言及	(CPiへの)適用状況	その他	参考文献
イギリス (ONS (Office for National Statistics) : Big Data Project)	食料品 (食料及び非アルコール飲料、アルコール飲料)	<ul style="list-style-type: none"> 食料品の価格データから実験的価格指数を作成 機械学習技術を分類及びクリーニングに利用(教師あり機械学習技術及び教師なし機械学習技術) 欠測による影響を減らす手段として補定についての分析 今後の取組み <ul style="list-style-type: none"> 「データ・ラングリング・プロセス」(収集データを集計に利用できる状態に加工するプロセス)の合理化 クリーニング及び分類アルゴリズムの向上 スクレイパーの改善 よりタイムリーな新指数の公表 指数集計のための新しい方法論 ウェブスクレイピング用の新たなデータ源などについて 	<ul style="list-style-type: none"> ONS試験プロジェクトで、3つのオンライン・スーパーマーケットチェーンに対してウェブスクレイパーを開発(Pythonでプログラム) 毎朝午前5時にウェブスクレイパーがCPiバスケット内33品目の価格を自動抽出。 平均的に、1日当たり約5,000価格を抽出(1か月当たり150,000価格) 従来手法は指定33品目について1か月当たり約6,800価格) ONSの2地点からスクレイピング(1地点の時系列に中断があっても、別の地点の収集データをマージ可能にするため) 以前は単純な検索語と除外語を用いた手法(「りんご」で検索、「りんごあめ」を除くために「アメ菓子」という除外語を指定など)はうまく機能するが手間。効率的な手法として教師あり学習を利用した手法を利用(アルゴリズムとしてSVM) 分類後のデータチェックとして教師なしクラスターベースの異常検出 スクレイパー中断等による欠測価格が生じた際の補正処理 	<ul style="list-style-type: none"> 大規模な収集。 データ収集とほぼ同時に物価指数を算出できる。 CPiの質及び効率を大幅に向上させる可能性を持っている。 	<ul style="list-style-type: none"> 価格及び分類を人の目でチェックするのは非現実的。機械学習などのデータサイエンス技術を利用 プログラミング技術により、全てのオンライン小売業者から価格をスクレイプできるとは限らない。 支出額に関わらず全ての価格が取得される。 価格収集が調査日ではないため、公表CPiと比較できる範囲が限定される。 在庫切れの問題。 大量データであるため、従来のクリーニング方法が機能しない。 	<ul style="list-style-type: none"> 2015年に日次・週次・隔週次・月次の頻度の試験的物価指数を提供(日次連鎖物価指数、単価指数、可能な限りCPi方法論に従った固定基準指数)。同年にGEKS-Jevons算式(GEKSJ)を用いて作成した指数を導入。 多数の研究者と共同で価格変動の新たな計算方法の開発にも取り組んでいる。 	<ul style="list-style-type: none"> 未適用(「公表CPiに可能な限り比較可能な方法で作成された指数について考察する。」) 	<ul style="list-style-type: none"> 700の財及びサービスのうち、約520品目の価格は調査員によって全国の店舗から調査、残りの価格はウェブサイトやカタログ、電話で中央集約的に収集(従来型価格収集と呼ばれる。) 小売業者のほとんどがスキャナデータの提供に消極的。進展が見られないため、ウェブスクレイピング利用の調査を継続。 一部サイトで「無限スクロール」システムを有し、24品目までしか取得できなかった(現在は解決済)。 	<ul style="list-style-type: none"> Research indices using web scraped data: May 2016 update, Robert Breton, Tanya Flower, Matthew Mayhew, Elizabeth Metcalfe, Natasha Milliken, Christopher Payne, Thomas Smith, Joe Winton and Ainslie Woods (ONS Website)
オランダ	航空運賃 衣料品	<ul style="list-style-type: none"> 衣料品のネット価格指数作成、店頭結果と比較 研究部門は機械学習アルゴリズムの利用などで分類ルールメンテナンスを自動化する方法を調査(高度なテキストマイニング手法を利用して実現可能となっている。) 衣料品の写真に有用な追加情報が含まれている場合があることから写真を利用する研究も開始 	<ul style="list-style-type: none"> 2種類の使用法 <ol style="list-style-type: none"> 多くの品目があるサイトからのデータ収集(電気製品のオンライン店舗のテレビなど) → 毎週、毎日作動のインターネット・ロボットで取組む「ロボット・フレームワーク」を開発(衣料品小売業者のウェブサイトの構造及び商品メタデータに類似点があるという事実を利用。ID、種類、名称、特徴、価格を最低限収集。) わずかな品目しかないネット上の多くのデータ源からの収集(映画館のウェブサイトの映画チケットの価格など) → インターネット・ロボット・ツールで取組む。 2016年1月現在、CPiに3つの小売業者のネット価格データを利用、他の19の小売業者から16のロボットが収集、すべてのロボットは衣料用に使用 <ul style="list-style-type: none"> コスト上の理由(店頭収集はコストがかかる) 方法論上の理由(衣料品は品質訂正の必要がほとんどなく、プロセスの自動化が容易) 	<ul style="list-style-type: none"> 店頭で収集するデータよりもはるかにすばやく収集できる。 小売業者の回答への負担が減少する。 コスト削減により統計局内の効率が高まる。 	<ul style="list-style-type: none"> 動的ウェブを背景としたロボットのメンテナンス(しかし店頭データ収集よりはコストは安い) 	<ul style="list-style-type: none"> 2018年には20~30のウェブサイトから取得したデータをCPiに取り込む予定 商品グループ内の同質性をできる限り保証するために、取扱品目が同一の区分に属する少数のブランドを有する小売業者を選択した。 	<ul style="list-style-type: none"> 公式統計の情報源としてインターネットを使用している(CPiに小売業者のネット価格データを利用)。 2018年までのプログラムの目標の一つがインターネット・ロボットをCPi作成に採用すること。 	<ul style="list-style-type: none"> 2012年にCPi情報源としてインターネットの利用を開始(ウェブスクレイパーによるデータ収集を開始) オランダCPiが書籍や電気製品といった他の商品種類のインターネット・データに進む可能性は低い(スキャナデータの方が望ましいデータ源である。) 衣料品については店舗調査をやめたいと考えている。 データの妥当性確認を可能な限り自動化しなければならない。 外れ値の人の目による検査は、値が何らかの境界値から外れたときのみ行うべきである。 自動化されたシステムは、エラーを検出・修正するだけでなく、それに関する処理情報も生成すべきである。 	<ul style="list-style-type: none"> On the use of Internet data for the Dutch CPI, Robert Griffioen, Olav Ten Bosch, Statistics Netherlands (Meeting of the Group of Experts on CPI 2016)
ノルウェー	パーソナルケア製品	<ul style="list-style-type: none"> ノルウェー統計局のオンラインデータプロジェクトの2つの目的: ① ノルウェーのCPi/HICPiにおけるオンライン価格の影響を大きくしたい。② オンライン収集を可能な限り効率化したい。 航空運賃など、1990年代中頃以降、インターネットから価格を収集して使用。今日ではインターネットから収集される価格がCPiウエイト・シェアで測定された財及びサービスのサンプル合計の約18%を占めている(大部分がサービス)。 航空運賃のデータ収集も調査。航空運賃に着目したのは主に効率のため(現在、人手がかかっている)。 	<ul style="list-style-type: none"> import.ioが提供するサービスを利用(他の統計機関もテストしている) 抽出したデータをソフトウェア(SAS)にロードして分析、計算及び保存 技術的な容易さから、様々なオンライン店舗(中でも大手オンライン店舗)から直接データを抽出することにした。 データを毎日収集(CPiの日次指数を得ることが目的ではないが、短期変動を確認することは有益) 出発点としてパーソナルケア製品、家電などオンライン購入が高いシェアを占めている消費財グループに注目(パーソナルケア製品のオンライン購入は売上高総額の7%と推定されている) 売上高上位の大手オンライン店舗4店からデータ収集 	<ul style="list-style-type: none"> 利用できるデータ量が膨大であること。 価格観察の量を大幅に増加することができる。 	<ul style="list-style-type: none"> 販売数量情報が得られない。 商品が対象期間中に実際に購入されたか分からない(しかしこれは物価統計担当者にとって新たな課題ではない。) オンライン小売業者は、クッキーなど、個々の消費者のブラウザ履歴に応じて、また所在地に応じてでも価格を変更することができ、この種の動的な価格設定(価格差別)はCPiにとって大きな課題になるかもしれない。 	<ul style="list-style-type: none"> パーソナルケア製品グループの日次連鎖指数のテスト計算でインターネットから収集した価格が実店舗の価格に比べて変動しやすいつながりが分かっていて。 	<ul style="list-style-type: none"> これまでCPi/HICPiデータの多くを紙の質問票で収集していた。今では、ウェブ・アンケート、スキャナデータ及びオンラインデータが重要なデータ源として取って代わっている。 	<ul style="list-style-type: none"> 2014年にオンライン価格の利用及びCPi/HICPiにおけるオンライン価格の利用を拡大する可能性に向けた作業を開始 ノルウェー人のオンライン購入の要因は高い生活水準、消費者パワーの強さ、ネット普及率が高いこと、ノルウェー国内の電子商取引で最も購入高の多いカテゴリは休暇及びレジャー旅行 オンラインで購入する商品の例は家電、衣料品、書籍、パーソナルケア製品 レジャー旅行産業は、デジタル化及びオンライン・ショッピングが幅広い変化をもたらした産業のひとつである。 ノルウェーの消費者の大多数がオンラインで航空券を購入 	<ul style="list-style-type: none"> The use of online prices in the Norwegian Consumer Price Index, Ragnhild Nygaard (Statistics Norway) (Ottawa Group 2015)
オーストリア	航空運賃	<ul style="list-style-type: none"> オーストリア統計局ウェブ・クロウリング・プロジェクトを実施。 ソフトウェアの選定 <ul style="list-style-type: none"> 法的分析 ソフトウェア・ITインフラの導入及び保守 商品グループ及びオンライン小売店の選定 選択したソフトウェアを用いた自動価格収集プロセスの開発 品質保証方法の開発 自動価格収集を様々な物価統計に利用すること 	<ul style="list-style-type: none"> ウェブ・クロウリング・プロジェクトは、外部企業からの既存の商用ビジュアル/ポイント&クリック型スクレイパーを利用。 ビジュアル/ポイント&クリック型スクレイパーの利用は以下の点を考慮して決定された。 <ul style="list-style-type: none"> 非常に低価格又は完全に無料で提供されていること 導入コストが低いこと 統計部局内及び組織全体の負担が少ないこと 融通性が高まること データ・サイエンティストが生まれること スクレイパーの中でもimport.ioを使用している。ただし、サービス拒否・事業廃止などの可能性があるため、単一のソフトウェア提供者への依存を避けるべき。 スクレイパーを用いた価格収集手順は少なくとも2段階で構成。 <ol style="list-style-type: none"> データを抽出する段階、② 抽出されて妥当性検証された価格データをデータベースにインポートする段階。 価格収集の後には、データのクリーニング、編集及び以前の価格収集期間の価格指数とのマッチングプロセスが行われる。 	<ul style="list-style-type: none"> スクレイパーは高品質の物価指数の作成に利用できる大量の価格の観察を生み出すことができる。 ビジュアル/ポイント&クリック型スクレイパーは、価格収集チーム自身で開発・維持することができる。 ウェブスクレイピングにより、航空運賃の指数作成に必要な価格収集に費やす月間作業時間を16時間から2時間に大幅に削減することができる。 データインポート、クリーニング、チェックなど様々な作業が含まれている。 	<ul style="list-style-type: none"> 比較的頻繁にウェブサイトが変更されるため、熟練したITプログラマーによるスクレイパーの再プログラミングが必要になる。 情報を抽出するにはスクレイパーにウェブサイト内をどのように検索するか、必要なデータをどのように見つけるか教える必要がある。そのためウェブサイトが有する個別のアーキテクチャに加え「無限スクロール」やJavaScriptなど特殊な機能を考慮に入れなければならない。 物価指数作成プロセスの作業は価格収集だけでなく、重み付けが必要な場合もある。例えばオーストリア国内発の全ての航空便の価格データはスクレイパーを利用すれば容易に入手できるが、全ての航空便価格情報を利用するためには、目的地ごとの信頼性の高い重み付けを最新の空港統計を用いて作成する必要があるため、物価指数作成の別の手順の作業負荷が大幅に増加する。 自動ウェブスクレイピングはオーストリア統計局のITシステムにとって情報漏洩の可能性があるとされている。 	<ul style="list-style-type: none"> 自動価格収集への投資の成果を最大化するために、プロジェクトの活動は、できるだけ多くの物価統計を含めることを目指す予定である。 全ての物価統計プロジェクトは、とりわけHICPi(ユーロ圏統一消費者物価指数)、PPP(購買力平価)のほか、生産者耐久財物価指数などの他の物価統計の開発について協力する予定である。 	<ul style="list-style-type: none"> 航空運賃の指数作成に必要な価格収集について、手作業から自動収集に既に完全に置き換えられている。 	<ul style="list-style-type: none"> ネット上での自動価格収集の利用はオーストリア統計局の2年間にわたるプロジェクト。EU統計局の助成金による支援を受けており、物価統計の近代化活動の一環 ウェブ・スクレイパーは、物価指数部職員によって開発・保守されるべき。物価指数部職員は、物価観察の正しい収集に関する様々な規則を適用して、ウェブサイトの頻繁な変更を評価して対応できる最適者である。 	<ul style="list-style-type: none"> Automatic data collection on the Internet (web scraping), Ingolf Boettcher (Statistic Austria) (Ottawa Group 2015)
ニュージーランド	家電商品	<ul style="list-style-type: none"> 家電製品について、オンラインデータから価格指数を作成 	<ul style="list-style-type: none"> 「Billion Price Project」からニュージーランド統計局に提供されたニュージーランドの大手小売業者に関するウェブスクレイピングオンライン日次データ(15か月分)を用いて、3つの家電商品(デジタルカメラ、携帯電話、テレビ)の日次連鎖Jevons指数と(非重み付け)日次FEWS(Fixed Effects Window-Splice)指数の実績を比較 	<ul style="list-style-type: none"> 自動ウェブスクレイピングを用いてあらゆる頻度で価格を収集できる可能性がある。 オンラインデータは、小売業者がオンラインの世界で重要な位置を占める場合に非常にタイムリーで高頻度の物価指数の可能性を提供する。 	<ul style="list-style-type: none"> 販売数量に関する情報が得られない。 数量の欠如によってオンライン指数間に乖離が生じる可能性及びカバレッジの相違 ヘドニックモデルなど明白な品質調整手法のために用いる商品の特性情報がほとんど(あるいは全く)ない。 	<ul style="list-style-type: none"> オンラインデータから作成したFEWS指数 日次連鎖Jevons指数と(非重み付け)日次FEWS指数を作成し、比較。 公的な統計機関の観点から見れば、オンラインデータの利用は非常に期待が持てる。 	-	<ul style="list-style-type: none"> スキャナデータによる分析(8つの家電商品、米国のスーパーマーケット商品)も実施 スキャナデータからの物価測定は、過去5年間にわたって活発に助成金を行ってきた分野 	<ul style="list-style-type: none"> Price indexes from online data using the fixed-effects window-splice (FEWS) index, Frances Krsinic (Statistics New Zealand) (Ottawa Group 2015)
イタリア	(試験的活動) 家電製品 航空運賃	<ul style="list-style-type: none"> ウェブスクレイピングに関して、家電製品と航空運賃の分野でデータ収集を自動的に再現するツールに置き換える試験的活動 <ul style="list-style-type: none"> 家電製品については作業時間が減少し、精度が向上した。 航空運賃については不確かな結果となり、限定的な改善にとどまった。 今後、スクレイピングで取得した大量データの管理を開始 	<ul style="list-style-type: none"> スクレイピングデータの処理及び分析のための専用ツール等を導入する。 	<ul style="list-style-type: none"> データ収集の負担を削減できる。 ビッグデータの論理において、商品流通のウェブ販路に関する調査設計を変える。 大量のデータを収集するロボットの試験の開始により商取引の世界をより適切にカバーできるようになる。 ビッグデータ技術によるデータ処理の効率化 価格データの包括的な過去のデータウェアハウスの構築 	-	<ul style="list-style-type: none"> CPi作成のため、2018年にビッグデータ(スキャナデータ及びウェブスクレイピングデータ)の利用を開始する。 サンプリング及びデータ収集等の異なる手法を取り込んで、消費者物価の調査を再設計する。 	<ul style="list-style-type: none"> 2015年にイタリア統計局が毎月収集する価格は95,600であり、このうち約13,000件はインターネット上の消費者向け電子製品の価格収集のためのウェブスクレイピング技術を使用して収集されている。(イタリア統計局HP) 	<ul style="list-style-type: none"> 2015年については、全国指数の場合、ローカルに収集された商品のウエイトは76.9%で、ISTATが直接収集する価格は23.1%。(イタリア統計局HP) 	<ul style="list-style-type: none"> Italian experience and perspective of using big data to estimate inflation., Federico Polidoro, Antonino Virgillito, Italian National Institute for Statistics (Meeting of the Group of Experts on CPI 2016)