Automating Identification of Potentially Problematic Privacy Policies

Greig Paul
Department of Electronic & Electrical Engineering
University of Strathclyde
Glasgow, UK
Email: greig.paul@strath.ac.uk

James Irvine
Department of Electronic & Electrical Engineering
University of Strathclyde
Glasgow, UK
Email: j.m.irvine@strath.ac.uk

***Abstract*—Almost every website, mobile application or cloud service requires users to agree to a privacy policy, or similar terms of service, detailing how the developer or service provider will handle user data, and the purposes for which it will be used. Many past works have criticised these documents on account of their length, excessively complex wording, or the simple fact that users typically do not read or understand them, and that potentially invasive or wide-reaching terms are included in these policies. In this paper, an automated approach and tool to gather and analyse these policies is presented, and some important considerations for these documents are highlighted, specifically those surrounding past legal rulings over the enforceability of some specific and widely-used contract terms — the ability for terms to be changed without directly notifying users (and presumed continued use indicates acceptance), and the protections in place in the event of a sale or acquisition of a company. The concerns these pose to user privacy and choice are highlighted, as well as the extent to which these terms are found in policies and documents from many popular websites. This tool was used to highlight how commonly these terms are found, and the extent of this potential problem, and explore potential solutions to the challenge of regulating user privacy via such contracts in an era where mobile devices contain significant quantities of highly sensitive personal data, which is highly desirable to service operators, as a core valuation asset of their company.**

I. Introduction

Privacy policies, terms and conditions, and other legal documents form a near-universal part of the experience of using connectivity-based services today. Virtually every website, mobile app, and even physical service provider has an agreement of this form, to which users are required to agree, in order to make use of the service. Often, however, these agreements are stated to be implicitly accepted by accessing or using a service, which gives rise to a number of considerations surrounding the validity of these agreements. Online agreements typically take the form of either a click-wrap [1], or browse-wrap [2]. These names are derived from an early form of software-related agreement, referred to shrink-wrap, whereby a user was held to have accepted a software End User License Agreement (EULA) by opening the shrink-wrap seal on the physical packaging itself [3].

The premise of such agreements was originally that it would be impractical for every user of a piece of software to individually negotiate a contract of use with the company providing the software, and record these agreements. As such, the concept of offering a standard agreement, which users accepted by using the software, was established. Such contracts are very common, and when browsing the internet, users are giving implicit consent.

Previous work has considered privacy policies, and protocols for computer-readable representations of such policies, such as P3P [4], allow web browsers to enforce a user's privacy choices. The automated processing of policies for the generation of computer-readable policies has also been carried out [5]. Despite P3P being a formally ratified standard of W3C, it has seen little traction, and has been shown to be abused by website operators [6] to work around restrictions on their sites put in place by user privacy policies. Some basic privacy-related enforcement is carried out in permissions on mobile platforms and web browsers (and these are computer-readable), although these focus simply on restricting access to data, rather than on restricting its use, which was the primary focus of the privacy policies reviewed.

Other work has highlighted a major limitation of P3P version 1, being that policies are only accepted or rejected, without scope for partial acceptance or rejection, or feedback to the service provider [7]. Privacy is a key consideration for future network services and applications, as even the simplest of software becomes increasingly connected, as it expands onto users' mobile devices (themselves holding large quantities of personal information).

In recent years, the rise in the tendency for companies to build their businesses around the prospect of making money as a result of data gathered from the users of an (otherwise) free-to-use service has been clear. Indeed, as Bruce Schneier stated in a conference talk in 2010: "Don't make the mistake of thinking you're Facebook's customer, you're not – you're the product," [8]. With the rise in free (at point of use) services on the internet, designed to encourage users to engage with them for the purpose of gaining a larger user-base, which itself is then used as an asset [9],

This paper firstly explores some legal precedents in contract law which are relevant to the terms encountered in privacy policies. It then explores the presented approach to automated identification of terms of interest within policies, as well as trends identified through an analysis of privacy policies of popular services and websites. Finally, potential points for future discussion are identified, and a set of recommendations for ensuring users' right to privacy is respected when using online and mobile services are made.

II. Related Works

A number of previous works have considered privacy policies, their implications, and their design. In particular, the concept of *privacy by design*, originally proposed by Canada's Information & Privacy Commissioner, has since been resolved as a standard for global privacy at the 32nd International Conference of Data Protection and Information Privacy Commissioners. The principles of privacy by design were presented as a set of 7 "foundation principles", designed to be used be applied to products and services as a principle during the development phase. These principles state that;

- Privacy should be built preemptively into products, rather than introduced as a reaction or response to risks or breaches,
- The maximum level of privacy should be enabled by default, without the user having to make a selection to achieve this,
- Privacy should be embedded into the architecture of both technology and business processes, so privacy is inherent throughout, rather than merely reducing functionality when invoked by users,
- Users should not be required to choose between privacy or a feature, as trade-offs should be minimised or eliminated,
- Privacy should be integrated end-to-end, from before data enters the system, to its timely destruction when it is no longer needed,
- Business practices and technology should be transparent, scrutinised, and verified to be operating correctly, so users can place trust based on this verification, and
- Privacy should be the utmost priority, with user-friendly and user-centric privacy options, to make it easy for individuals to protect their privacy. [10]

One important area of previous research which may make practical many of the principles of data protection by default is that of privacy-preserving Attribute-based Credentials (Privacy-ABCs), which allow for attribute-based verification of users, without requiring the full disclosure of identity or the other related attributes [11]. For example, using Privacy-ABCs, it would be possible for a user to show they were a member of a given restricted set or group, without having to disclose details of other sets they belong to, or disclose their individual identity. This technology offers the ability to further many of the principles of minimisation of data, and to reduce the need for storage or processing of sensitive data when Privacy-ABCs could be used to verify an operation without processing sensitive information.

The *Terms of Service; Didn't Read* project also offers users a novel approach to privacy policies, with curated and user-submitted interpretations of policies used to grade policies with a simple letter-grade, along with a selection of key positives and negatives [12]. This system, integrated with browser plugins, provides users with a rapid way to gain an understanding of the implications of a policy, at least in the interpretation of other users.

Previous work has also highlighted the problem of a lack of clear regulation and standard baselines when privacy policies are considered [13], where a lack of standard language, and standard content addressed in them presents effectively different policies, making it difficult to compare privacy policies. Jensen also highlighted the problem surrounding the non-negotiable nature of website policies, given the non-negotiable nature, and the lack of user leverage or voice to request or set new terms. This highlights that the problem of the one-way nature of such agreement is not a new one, and that subsequent years have not led to any significant changes here, as the results of this investigation shall show.

The importance of privacy preservation on the internet has also been previously explored, and likened to the ability to have relative privacy within regular, offline day-to-day life [14]. Culnan and Armstrong highlighted the dependency of businesses on gathering, and processing customer data in large quantities, and the trade-off between value for the company, and privacy for the customer [15]. The use of procedural fairness was proposed, where the presence of fair procedures for protecting individual users' privacy would lead to customers being willing to disclose personal information, and found that where procedural fairness practices were applied, privacy concerns were not the distinguishing factor between customers willing and not willing to be profiled, indicating customer acceptance.

III. Legal Precedents

A. Alterations to Contracts

In Douglas v. US District Court ex rel. Talk America, (Case 06-75424), the Ninth Circuit Court of Appeals found that one party cannot unilaterally alter the terms of a contract. The grounds for this were that modified terms remain an offer to change the terms, until the terms are accepted. The judgement also stated "In California, a contract can be procedurally unconscionable if a service provider has overwhelming bargaining power and presents a "take-it-or-leave-it" contract to a customer — even if the customer has a meaningful choice as to service providers", referring to the 9th Circuit case of Nagrampa v. MailCoups, Inc. (69 F.3d 1257, 1283) in California. This is of interest, as a large number of technology companies are based in Calfornia, and make one of their terms of use that the agreement is covered by US law.

The Douglas ruling specifically highlighted the questions posed where a contract is modified, with the only notice given by posting a revised version. This is of interest with regard to the legal documents and privacy policies found on websites, since of the top 10 websites investigated, all except Amazon contained a clause indicating that continued use of a service indicated acceptance of the policy changes. The original case of Douglas v. Talk America pertained to the enforceability of a change in contract to add a mandatory arbitration clause. The judgement discussed the impracticality of requiring a user to check for updates to contracts or agreements for every service they use, and highlighted that if a user was not aware of the change, the user cannot agree to the offer for change unless they are aware of it, per Samuel Williston & Richard A. Lord, A Treatise on the Law of Contracts.

When reviewing privacy policies of a number of popular websites (as discussed in Section IV), Google was identified as the only service to give a firm commitment to not reduce privacy protections without obtaining "explicit consent" of users. The majority of services made no specific commitment to directly notify users (perhaps using a direct message, or an email) of changes to policies, although many did state they may do it in some cases, although it was found (for the case of Twitter, for example), that this was only if (at their own sole discretion) they decided a change was material.

Clearly there is a balance to be obtained here between notifying users of every single trivial change to a policy, and allowing changes to policies to be made without specifically contacting users. In light of the previously US-based legal rulings covering such changes, however, there is a strong argument to be made that all changes to website terms or contracts should require explicit acknowledgement and consent, to be binding.

B. Explicit Consent to Alterations

A second matter, however, covers implicit acceptance, where continued use of a service, after its terms are updated, or after a period of time, is taken by the company to indicate agreement with the updated terms. In the case of Nagrampa v. MailCoups, the court highlighted the overwhelming bargaining power of one party over the other, and that the alterations were presented on a "take-it-or-leave-it" basis. Since alterations to website policies are typically carried out as such (without provision made for users unwilling to agree to the new terms), this ruling may prove relevant when considering internet-based services, as the provider holds a position of overwhelming bargaining power (by being able to prevent a user from accessing their own data until they accept new terms). This position may well form the grounds for modifications to be deemed unfair, on account of the imbalance of bargaining power between user and service provider.

This is also inkeeping with UK unfair contracts law guidance [16], which is itself derived from a European directive, meaning that similar terms should exist throughout the European Union. It appears that contracts between users and service providers are inherently designed with an imbalance, given the service provider is able to dictate that a user accessing their own data (held on the service) must accept new and modified terms, with the user having no clear option to access their data without accepting those terms. Indeed, in some cases, users may not need to even use the service, as acceptance was in some cases presumed after a specific period of time. This raises concerns for services where users are not given a clear ability to remove all their data from a service (and terminate their dealings with the company). Another consideration here is whether it is fair or reasonable for a user to no longer have access to a service, simply on account of their objection to a unilateral modification to the contract, which was not in place when they accepted the agreement. Had they been aware of the intention of the provider to make this change, they may never have used it, or may have used an alternative.

C. Transfer of Data Following Acquisition

A common term seen within privacy policies and terms of service covers the ownership (and transfer of) user or customer data, in the event of the sale, acquisition or bankruptcy of the company operating the service. These terms find their origins as a result of a number of cases in the USA, where the Federal Trade Commission (FTC) successfully argued that a company undergoing bankruptcy was bound by its original promises in a privacy policy, such as to not sell private data, or to protect it [17]. Specific examples include that of RadioShack and Toysmart. While there has been considerable legal attention given to the subject, online services are recognised as being largely self-regulated, with regard to the handling of personal data (through privacy policies), which is why the FTC seeks to hold companies to account and ensure they honour their own privacy policies [18].

As a result of this, it is common to see website and service privacy policies clearly state that data may be transferred or sold as part of a sale, bankruptcy or acquisition. By stating this may happen, the company is acting within its own privacy policy, and can likely continue with the sale unhindered. Indeed, previous legal challenges brought forth by the FTC have focused on holding companies to their own privacy policies when attempts at selling user data are made, and in the absence of legally binding protections for private data, it appears acceptable for personal data to be sold as part of a bankruptcy or acquisition, provided such a provision was made in the privacy policy in question.

In order to investigate the prevalence of these kinds of terms, an analysis of a number of privacy policies was carried out, accelerated through the use of the presented toolkit, used to identify potentially interesting contact and policy terms for further analysis. These policies belong to the most popular web and mobile-oriented services.

IV. Automatic Processing

To gain a clear understanding of the extent to which some of the concerns highlighted in Section III are found in policies available today, a toolkit was created for the purpose of analysing these policies. The toolkit accepts a CSV list of top website domain names (the Alexa listing being one widely known example of such a dataset). In order to simplify analysis, at this point only sites using the '.com' and '.co.uk' top-level domains are considered, since these sites typically feature their legal policies in English. In future, detecting website language based on the page content would allow for this to be expanded, and would allow for non-English website policies to be processed, provided suitable policy definitions were created for the language in question.

The identified websites were then accessed on their default "www" subdomain prefix by the automated tool, and the homepage was loaded into an HTML parser, which identified all hyperlinks within the page containing a phrase relevant to legal policies. These links were then followed, and the policy documents gathered and scanned for strings likely to be of interest in investigating the transfer of personal data, or the alteration of policies. At this time, the phrases shown in Table I were used, which were themselves selected from the top 20 websites.

V. Challenges in Automated Processing

In the process of carrying out this automated processing, a number of challenges were identified when attempting to

TABLE I. QUERY TERMS

Hyperlink Query String	Page Content String
terms	transfer
privacy	acqui
legal	merge
	amend
	modif
	notif

automate the process of extracting key privacy policy information. Firstly, as discussed previously in Section IV, websites written in languages other than English were not considered at this time, although the process could be applied to websites of any language provided suitable experience was available for identifying suitable query strings. Secondly, the challenge of ambiguity in links was also identified, where a website may contain more than one link mentioning "privacy" or "terms". Two potential approaches to handle this scenario were identified. One approach was to use the link with the shortest visible text, since articles mentioning the query strings or other, more generic pages about privacy tended to be qualified with other words, such as "Major privacy scandal over...". The other approach was to use the link which was found last on the page, since all websites investigated listed their privacy policy at the footer of their page, which is conventionally listed towards the end of the HTML in a page. One other potential challenge identified (but not encountered on any of the most popular websites investigated) was the use of javascript-based hyperlinks to show a privacy policy. Since these would require javascript code to execute in order for the policy to display, retrieval of the destination of the hyperlink would not show the policy. While this was not encountered on any websites processed, it would potentially lead to the tool not detecting any relevant policies on the page. For this reason, manual supervision of the acquisition process remains necessary at this point, to ensure the policy is identified correctly.

After identifying the relevant hyperlinks within a website, a number of challenges were identified in attempting to automatically process the policies themselves. While the process of attempting to extract context of a query term by retrieving the full HTML component within which the term appeared was relatively successful, it was not always perfect. Specifically, if a sentence was split (as is common when breaking up lengthy legal statements using bullet points, or sub-sections), the subsequent information was lost. To ensure that findings were accurate for this work, the relevant terms of policies were manually reviewed (within the regular context of the full policy). By also taking into account the variations in use of terminology, fully automated processing of these privacy policies and other legal documents remains a significant hurdle. Some relatively common words (such as merge, or modify), are by definition necessary to highlight for the purpose of identifying policy terms of interest, and are also likely to be found in other contexts (such as users being able to modify existing content), or indeed within in-page javascript content to merge multiple objects together when rendering the web page. The latter of these were able to be removed using length filtering (to remove infeasibly long lines of content), although parsing the string as javascript to verify syntax also allowed these false positives to be removed.

VI. HANDLING OF CHANGES TO POLICIES

Two factors shall be considered in how services stated they would handle changes to policies. Firstly, whether or not users would be notified of change beyond a simple update to the policy was considered. Secondly, whether or not a service states that continued use of the service implies acceptance of updated terms will be considered — this is to identify whether or not users are giving informed consent to the changes, or whether they are perhaps unaware of what they are agreeing to. Given the OECD privacy guidelines focus on consent as one of their key principles, this is an important means of verifying if privacy protections in place are sufficient to protect users [19].

Google's privacy policy stated that no changes will reduce user rights under their privacy policy without explicit consent, but did not specifically require itself to directly notify users on each change — for "significant" changes, prominent notice would be given, but this was not defined, and the meaning of prominent and significant were not

> Our Privacy Policy may change from time to time. We will not reduce your rights under this Privacy Policy without your explicit consent. We will post any privacy policy changes on this page and, if the changes are significant, we will provide a more prominent notice (including, for certain services, email notification of privacy policy changes). We will also keep prior versions of this Privacy Policy in an archive for your review. [20]

While Google's policy did not explicitly state whether it presumes implied consent when they alter their privacy policy, the email they circulated during the merge of their privacy policies in January 2012 stated, "If you choose to keep using Google once the change occurs, you will be doing so under the new Privacy Policy and Terms of Service". This would indicate that Google believes continued use of their services indicates consent to the changes.

The Facebook privacy policy stated that "We'll notify you before we make changes to this policy and give you the opportunity to review and comment on the revised policy before continuing to use our Services", indicating that notification would be carried out. While the ability to comment on revised policies is one not commonly seen, their terms of service state that consent is presumed if users continue to use Facebook; "Your continued use of the Facebook Services, following notice of the changes to our terms, policies or guidelines, constitutes your acceptance of our amended terms, policies or guidelines" [21].

Yahoo's terms of service state that "if you do not agree to the changes in the Terms, you should stop using the Yahoo Services and Software", indicating that they also view consent as implicit if a user continues to use their service, and that users have no means to opt out of changes to policies. Yahoo did state in its privacy policy that "if we make any substantial changes in the way we use your information we will notify you by sending a notice to the primary email address specified in your Yahoo account or by posting a prominent notice on our pages" [22], indicating that notification of changes should take place in some situations.

Twitter's privacy policy stated that "If we make a change to this policy that, in our sole discretion, is material, we will notify you via an @Twitter update or email to the email address associated with your account. By continuing to access or use the Services after those changes become effective, you agree to be bound by the revised Privacy Policy" [23]. This indicates that Twitter may in some situations choose to alert users to changes, and that they presume consent if a user accesses the service after these changes took effect. It is also worth noting that the *@Twitter* account is a high-activity account which regularly tweets and re-tweets content, indicating that privacy update notices may not be highly visible if only sent from that account.

Amazon offered a stronger committment to privacy policy changes, stating that they will not change their policies to offer less protection to previously-collected data without consent, although there was no statement that customers would be allowed to refuse these changes. Amazon did however state that users should frequently check their website to see changes to policies, an approach which has been viewed as unrealistic by US courts in the case of Douglas v. US District Court ex rel. Talk America, (Case 06-75424), as referred to above.

> Our business changes constantly and our Privacy Notice and the Conditions of Use and Sale will change also. We may e-mail periodic reminders of our notices and conditions, unless you have instructed us not to, but you should check our website frequently to see recent changes. Unless stated otherwise, our current Privacy Notice applies to all information that we have about you and your account. We stand behind the promises we make, however, and will never materially change our policies and practices to make them less protective of customer information collected in the past without the consent of affected customers [24].

The Wikipedia privacy policy stated that "your continued use of the Wikimedia Sites after this Privacy Policy becomes effective constitutes acceptance of this Privacy Policy on your part" [25], indicating that they consider users implicitly consenting to their privacy policy, and stated that notification would be carried out for substantial changes, which would not include changes such as "grammatical fixes, administrative or legal changes, or corrections of inaccurate statements";

> In the event of substantial changes, we will provide the proposed changes to our users in at least three (3) languages (selected at our discretion) for open comment period lasting at least thirty (30) calendar days. Prior to the start of any comment period, we will provide notice of such changes and the opportunity to comment via the Wikimedia Sites, and via a notification on WikimediaAnnounce-L or a similar mailing list [25]

Microsoft's privacy policy stated that it would notify users in the event of significant changes, although did not specifically state how consent was handled when changes are introduced;

> We will update this privacy statement when necessary to reflect customer feedback and changes in our services. When we post changes to this statement, we will revise the "last updated" date at the top of the statement. If there are material changes to the statement or in how Microsoft will use your personal data, we will notify you either by prominently posting a notice of such changes before they take effect or by directly sending you a notification. We encourage you to periodically review this privacy statement to learn how Microsoft is protecting your information [26].

Their terms and conditions stated, however, that "Microsoft reserves the right to update the TOU at any time without notice to you", indicating that it is likely Microsoft views their policies as being changeable without user consent or knowledge, in some cases [27]. This is backed up by their use of the passive description of changes taking effect in the privacy policy quoted above, rather than of users actively agreeing to them.

The LinkedIn privacy policy stated that material modifications would be notified to users through the service or another means, but that using the service after a notice is given, consent is implied to the changes;

> We may modify this Privacy Policy from time to time, and if we make material changes to it, we will provide notice through our Service, or by other means so that you may review the changes before you continue to use our Services. If you object to any changes, you may close your account. Continuing to use our Services after we publish or communicate a notice about any changes to this Privacy Policy means that you are consenting to the changes [28].

The eBay privacy policy stated that changes may be made by adjusting the policy, and that significant changes would be notified, with changes taking effect after 30 days;

> We may amend this Privacy Notice at any time by posting the amended terms on this site. All amended terms automatically take effect 30 days after they are posted. We will announce any material changes to this Privacy Notice through the eBay Message Centre and/or via email [29].

The Instagram privacy policy likewise stated that changes may be made to the policy, and that users should review it regularly, with optional notice being given for more significant cahnges. There was also presumed consent ot changes if a user accessed Instagram after the changes;

> Instagram may modify or update this Privacy Policy from time to time, so please review it periodically. We may provide you additional forms of notice of modifications or updates as appropriate under the circumstances. Your continued use of Instagram or the Service after any modification to this Privacy Policy will constitute your acceptance of such modification [30].

VII. Discussion

A number of common trends were identified within these policies, raising a number of questions. Firstly, not all services explicitly stated the duration of time between changes

being made available, and changes taking effect. Specifically, Instagram, LinkedIn, Amazon, Yahoo and Google did not state explicitly that a notification period would be given. Microsoft, Twitter and Facebook mentioned a period of time but did not specify it, or implied that changes did not take effect immediately upon posting. eBay and Wikipedia expressly stated a specific period of time which would be given to review changes.

Secondly, the question of validity of consent was highlighted — Google, for example, promised not to reduce privacy protections without explicit user consent, although did not state that users would be offered any means to object to this change, or to otherwise avoid it. Google has previously used implied consent to significant privacy changes [31], as discussed earlier with regard to their email to users in 2012. Of the services investigated, Google, Facebook, Yahoo, Twitter, Wikipedia, Microsoft, LinkedIn, eBay and Instagram all made statements which indicated that consent to changes was implied, or which stated that changes may be made and would take affect. Amazon did not explicitly state this, although gave no details as to how consent would be obtained, although their promise to not reduce the protections on existing data may indicate that express consent would be sought.

Thirdly, while the explicit or implicit nature of consent has been considered, an equally, if not more important, consideration is what happens in the event of a user not agreeing to a change in policies. There are a variety of scenarios where this raises a number of questions — a user may be travelling abroad for an extended period of time without access to the internet by choice, meaning they are unaware of changes being imposed on policies to which they have agreed. Based upon the basic premise of agreements such as these, these changes would be considered as an offer to variation, and it would be unreasonable to apply them to an unaware user [32]. Another scenario where this may be problematic would be where a user is incapacitated or hospitalised, and not in a state to view length privacy policies or migrate data outwith services.

VIII. Key Observations

As a result of carrying out this investigation, a number of key findings were identified, which should be considered in line with the discussion regarding enforcement of terms in Section III. Specifically, these findings lead to a belief that;

- Almost all online services deem continued use of a service as acceptance of new terms or policies, irrespective of whether the user is made aware of the changes, which is potentially not in-keeping with legal precedent.
- Very few services promise to directly contact and inform individual users of changes to policies. Most say they may do this, but some services may only post a message on their service website, or perhaps only even update the policy.
- Some services say they will notify users of major changes to their policy, although this is as determined by the company itself, and in some cases the terms state that only the company may determine if a change is major.
- Service operators may in a position such that rights in a policy or contract could be found significantly imbalanced — a user must agree to arbitrary terms to continue to access their own data on a service, and their access to that data can be terminated if they refuse to consent to the new terms.
- Many services state in their terms that user data may be transferred to another company in the event of an acquisition, merger or sale. The privacy protections experienced here may well differ from what users expect, or previously were in force.

In the case of policy updates, it should also be considered that, on account of the wide use of smartphone applications, not all users will encounter a notice of a change placed on the homepage of a service. For example, many users of Google services on Android will have had an account created during the Android setup process, potentially within a retail store, and may not ever use (or know their password for) the web interface for these services. Additionally, especially on mobile devices, users are often prompted to accept policies without necessarily having even seen the policy - due to screen space limitations and the lengths of these policies, the policy was often made available by way of a hyperlink, which upon clicking would open a new screen containing the policy in question.

A. *Safe Harbor Legislation*

Another major consideration when reading privacy policies and other legal documents, specifically pertaining to the usage of internet services, is the jurisdiction under which contract is governed. While most policies investigated stated this, there can be considerable variation in legal protections for users between jurisdictions. This is particularly clear when US and European Union data protection laws are considered. The *Safe Harbor* process previously allowed for companies to self-declare their compliance with European law. This aimed to remove the challenges faced with the EU's comparatively heavily regulated handling of private data (which is established as a "fundamental right" [33]), in comparison with the relatively hands-off approach taken in the US, where privacy protections are typically implemented voluntarily by companies, in order to prevent potential lawsuits [33]. Despite this, a European court ruling recently led to the effective suspension of the Safe Harbor provision [34]. In this case, Schrems brought a case against Facebook Ireland Ltd. on account of its transfer of EU citizens' personal data to the United States, and its storage there. The case centred around Schrems' complaint that Facebook's transfer of his personal data to the United States did not offer sufficient protections, and that he wished to exercise his right to prohibit the transfer of such data, on account of surveillance activities being carried out by public authorities within the United States.

The Irish High Court found that, since US law did not give EU citizens a right to be heard, and that oversight of intelligence services was carried out through secret procedures, it did not offer adequate protection. The end result was that Decision 2000/520, which stated the European Commission "may find that a third country ensures an adequate level of protection", and established that the Safe Harbor program

offered such protections, was reversed. This meant that data transfers under Safe Harbor are no longer automatically valid, and that legal challenge may be made to data being transferred to the United States by European companies [34].

It is therefore clear that the treatment of personal data may vary significantly between different countries and regulatory systems — within the context of privacy policies, it is possible that users may be asked to agree to the transfer of data to other countries whose protections may not be sufficient, as was seen in the case above regarding Facebook.

IX. Recommendations for Privacy

As a result of these observations, some recommendations are presented, to improve the fairness of privacy policies for end-users. Firstly, all users should be made aware of all changes to a privacy policy or other legal document governing the use of provision of a service. This notification should be made directly to the user, in plain language, and in an honest and understandable way. There is a clear trade-off between annoying users and ensuring they are informed of changes, but given the legal precedents discussed in Section III, it is necessary to ensure users are definitely made aware of changes. In the meantime, use of a service such as *TOSBack* [35] may help users to monitor services' policies for changes, although the addition of notifications for changes would be beneficial for ensuring that users become aware of policy changes quickly, such that they may react if they do not agree to them.

As an extension of the above, services should also seek explicit consent when updated terms or policies are presented. It should not be necessary for users to depend on third party community services to ensure policies are not changed without their knowledge. Ensuring such consent is given ensures that users have directly consented to the variations in terms, and that there is no presumption of consent. If a user no longer uses a service, for example, their consent may be presumed when the user was unaware of the alteration (and unwilling to agree to it). By requiring explicit consent to all changes (perhaps at subsequent logins, as some websites implement), service operators know which users have consented directly to the changes, and can separate user data which is not able to be used under the updated policies..

Similarly, the tie between agreeing to new terms, and being able to continue to use a service which has already been available presents a question and challenge for further discussion. If a user is prevented from accessing their data until they agree to a new set of policies, they are arguably being unfairly pressured into agreeing to a change in policies or contract. This could well be the equivalent of being forced to sign an agreement under duress, which would ordinarily invalidate it. Nonetheless, this topic seems to require further discussion, as it highlights and underpins the differing needs of service users and providers. While service providers typically wish to protect themselves from legal challenge over the use of user data, and ensure their users understand the ways data may be used by the provider, the user of a service wishes to understand how their personal information is used, and to exercise control over this to prevent use which they do not consent to. By being able to balance these needs (perhaps through providing access to the previously-offered service for compatibility), a satisfactory outcome should be able to be achieved, such that service providers may change policies, but users may continue to use services which they are reliant upon, without being coerced or forced into agreeing to something they are unhappy with. When personal data is being handled, the ability to challenge a change after the fact may not necessarily offer a satisfactory remedy, since the personal data may already have been shared or sold or otherwise used without the user's consent, and it cannot easily be "taken back" after the fact.

Finally, existing standards (such as P3P) for the machine-readable representation of privacy policies should be revisited, as they offer significant benefit for those wishing to compare privacy policies between websites. For these to be effective, however, it would be necessary that service operators and websites be bound by their P3P-stated policy, such that users may allow their browser to make decisions based on website policies. If service operators believed that their human-readable policies were the canonical definition of their policies, this would potentially allow for abuse, where sites would place incorrect P3P policies on their site to mislead users into disclosing information under false pretences.

X. Conclusion

This paper has discussed some of the legal and user-facing considerations of privacy policies online, as well as an approach to automatically locating them on websites, and attempting to highlight important portions of these policies for further review. This approach has been used to automatically flag sections of privacy policies for further review, which appear to present clauses which act against the legal precedents explored. A set of observations across an analysis of many of these policies from some of the most popular websites on the internet was presented, and recommendations were made with regard to specific concerns surrounding presumed or implied consent to updated policies, and of the notification of users to changes in such policies. Finally, the issue of the all-or-nothing approach to the use of services has been highlighted, and the potential risks of data lock-in being used to force users to agree to changes to privacy policies or other legal documents under what could be perceived as duress.

Acknowledgment

This work was funded by EPSRC Doctoral Training Grant EP/K503174/1.

References

[1] C. L. Kunz, M. F. Del Duca, H. Thayer, and J. Debrow, "Click-through agreements: Strategies for avoiding disputes on validity of assent," *The Business Lawyer*, pp. 401–429, 2001.

[2] C. L. Kunz, J. E. Ottaviani, E. D. Ziff, J. M. Moringiello, K. M. Porter, and J. C. Debrow, "Browse-wrap agreements: Validity of implied assent in electronic form agreements," *The Business Lawyer*, pp. 279–312, 2003.

[3] R. H. Stern, "Shrink-wrap licenses of mass marketed software: Enforceable contracts or whistling in the dark," *Rutgers Computer & Tech. LJ*, vol. 11, p. 51, 1985.

[4] L. Cranor, M. Langheinrich, and M. Marchiori, "A P3P preference exchange language 1.0 (APPEL1.0)," *W3C working draft*, vol. 15, 2002.

[5] C. A. Brodie, C.-M. Karat, and J. Karat, "An empirical study of natural language parsing of privacy policy rules using the SPARCLE policy workbench," in *Proceedings of the second symposium on Usable privacy and security*. ACM, 2006, pp. 8–19.

[6] P. G. Leon, L. F. Cranor, A. M. McDonald, and R. McGuire, "Token attempt: the misrepresentation of website privacy policies through the misuse of P3P compact policy tokens," in *Proceedings of the 9th annual ACM workshop on Privacy in the electronic society*. ACM, 2010, pp. 93–104.

[7] R. Grimm and A. Rossnagel, "Can P3P help to protect privacy worldwide?" in *Proceedings of the 2000 ACM Workshops on Multimedia*, ser. MULTIMEDIA '00. New York, NY, USA: ACM, 2000, pp. 157–160. [Online]. Available: http://doi.acm.org/10.1145/357744.357917

[8] D. Shane. Facebook is "deliberately killing privacy", says Schneier. [Online]. Available: http://www.information-age.com/technology/security/1290603/facebook-is-%22deliberately-killing-privacy%22-says-schneier

[9] T. Louis. (2013, 31 August) How much is a user worth? [Online]. Available: http://www.forbes.com/sites/tristanlouis/2013/08/31/how-much-is-a-user-worth/

[10] A. Cavoukian *et al.*, "Privacy by design: The 7 foundational principles," *Information and Privacy Commissioner of Ontario, Canada*, 2009.

[11] K. Rannenberg, J. Camenisch, and A. Sabouri, *Attribute-based Credentials for Trust: Identity in the Information Society*. Springer, 2014.

[12] (2016) Terms of service; didn't read. [Online]. Available: https://tosdr.org

[13] C. Jensen and C. Potts, "Privacy policies as decision-making tools: an evaluation of online privacy notices," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 2004, pp. 471–478.

[14] I. Goldberg, D. Wagner, and E. Brewer, "Privacy-enhancing technologies for the internet," DTIC Document, Tech. Rep., 1997.

[15] M. J. Culnan and P. K. Armstrong, "Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation," *Organization science*, vol. 10, no. 1, pp. 104–115, 1999.

[16] Competition and Markets Authority. (2014, 5 June) Unfair terms in consumer contracts regulation explained. [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/317667/UTCCR-explained-an-overview-in-an-easy-to-digest-format.PDF

[17] M. P. Heuga, G. Manglik, and C. Meyer. (2015, 13 August) Transferring customer data in an asset sale. [Online]. Available: http://www.jdsupra.com/legalnews/transferring-customer-data-in-an-asset-47838/

[18] W. E. Agin, "The new regime for treatment of customer data in bankruptcy cases," *J. BANKR. LAW & PRAC.*, vol. 10, p. 365, 2001.

[19] OECD. (2013, July) 2013 oecd privacy guidelines. [Online]. Available: https://www.oecd.org/internet/ieconomy/privacy-guidelines.htm

[20] Google Inc. (2015, August) Welcome to the google privacy policy. [Online]. Available: https://www.google.com/intl/en/policies/privacy/

[21] Facebook. (2015, January) Statement of rights and responsibilities. [Online]. Available: https://www.facebook.com/legal/terms/update

[22] Yahoo. (2014, January) Yahoo privacy centre. [Online]. Available: https://policies.yahoo.com/ie/en/yahoo/privacy/index.htm

[23] Twitter. (2015, May) Twitter privacy policy. [Online]. Available: https://twitter.com/privacy?lang=en

[24] Amazon.co.uk. (2015, April) Amazon.co.uk privacy notice. [Online]. Available: https://www.amazon.co.uk/gp/help/customer/display.html/ref=footer_privacy/275-7312912-5386142?ie=UTF8&nodeId=502584

[25] Wikimedia Foundation. (2014, April) Privacy policy. [Online]. Available: https://wikimediafoundation.org/wiki/Privacy_policy

[26] Microsoft. (2015, October) Microsoft privacy statement. [Online]. Available: https://www.microsoft.com/en-gb/privacystatement/default.aspx

[27] ——. (2015, June) Microsoft terms of use. [Online]. Available: https://www.microsoft.com/en-us/legal/intellectualproperty/copyright/default.aspx

[28] LinkedIn. (2014, October) Your privacy matters. [Online]. Available: https://www.linkedin.com/legal/privacy-policy

[29] eBay, "User privacy notice," May 2015. [Online]. Available: http://pages.ebay.co.uk/help/policies/privacy-policy.html

[30] Instagram. (2013, January) Privacy policy. [Online]. Available: https://help.instagram.com/155833707900388

[31] N. Mediati. (2012, February) Google privacy checklist: What to do before Google's privacy policy changes on march 1. [Online]. Available: http://www.pcworld.com/article/250950/google_privacy_checklist_what_to_do_before_googles_privacy_policy_changes_on_march_1.html

[32] S. Williston and W. H. E. Jaeger, *A Treatise on the Law of Contracts*. Baker, Voorhis, 1957, vol. 6.

[33] W. J. Long and M. P. Quek, "Personal data privacy protection in an age of globalization: the US-EU safe harbor compromise," *Journal of European Public Policy*, vol. 9, no. 3, pp. 325–344, 2002.

[34] "C362/14, Judgment of the Court (Grand Chamber) of 6 October 2015. Maximillian Schrems v Data Protection Commissioner," ECLI:EU:C:2015:650.

[35] (2016) Tosback - the terms-of-service tracker. [Online]. Available: https://tosback.org