Distributed Business Service Discovery for Direct Data Driven Business

April 2016

Henrik Biering
Peercraft

Abstract:

Decentralized discovery and resolution has been a key enabler for the resilient and cost effective technical infrastructure of the internet. However, the early internet was dominated by content relating to research, pornography, and personal interests not typically associated with registered activities. Hence significant resources had to be invested by companies taking up the challenge to discover and index this unstructured content appropriately.

Today this situation has changed as the internet has become a primary channel for commerce. In the absence of formal company and personal internet identities, market making has been relying on a number of third party intermediaries to provide mutual discovery and support safe transactions. These third party intermediaries have now proven overly costly and troublesome to business.

Meanwhile many national business authorities have become online allowing instant discovery of businesses based on various criteria. Some authorities even provide their databases for download and enrichment with other data. Hence the author suggests to use the official company registrations as the basis for a decentralized authoritative registration of the commercial services of a company.

For companies such a registry will make their services directly and globally discoverable by search services implementable by anyone with modest resources. Consumers will be able to easily authenticate companies when browsing the web or receiving emails. Also the absence of dominant third party intermediaries facilitates drastically improved control with their use of personal data.

Contents

Background and status for internet based commerce

Up until the industrial revolution most goods and services were produced and consumed within local communities. Sellers and buyers discovered and selected each other based on reputation/trust and/or direct negotiation.
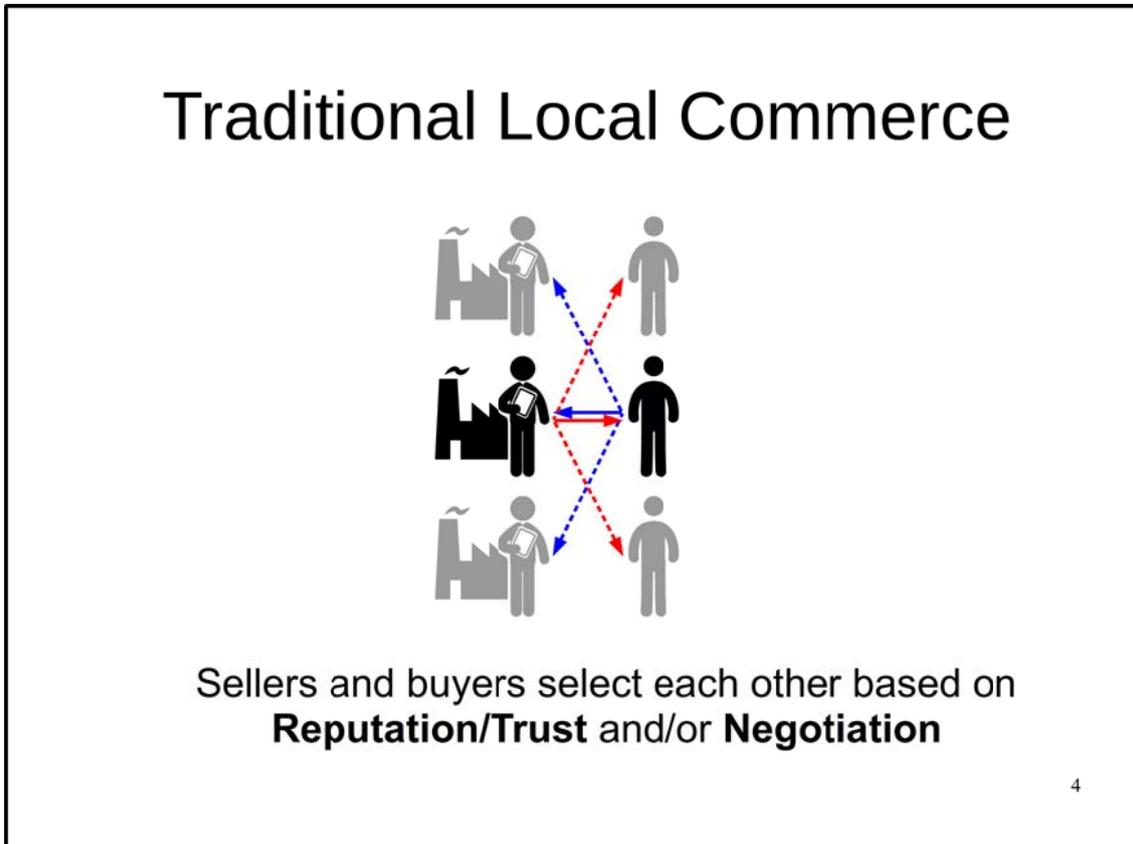
In the industrial age the globalization of commerce activities led to the introduction of supply chains including intermediaries such as wholesalers and retailers. In many cases, however, shipping and logistics was not the main bottleneck for globalization. It was rather bringing remotely produced products and services to the attention and appreciation of the consumer. This led to the emergence of "Marketing" as a business discipline, the purpose of which is to herd potential customers into a purchase funnel where they eventually loose sight of any alternatives. This could include activities such as "branding" or simply paying retailers to promote a specific manufacturers products while persuading them to hide or stop carrying competitors products.

In the early days of the web (mid 90's) many expected that the vision of the "The Global Village" already realized in the media world would also govern the development of e-commerce. A common mantra was "disintermediation" or "cutting out the middlemen".

It was overlooked that while the globalisation of public news is easy, a prerequisite for global commerce is mutual trust based on reputation and/or security measures which again rely on globally recognized identity schemes.

It was also overlooked that a global village with billions of consumers and millions of vendors

would require the development of very effective data structures and management tools for mutual discovery, evaluation, and targeting/selection. On the vendor side there has been some discussion of "1 to 1 Customer Relation Management" but a majority of companies still pursue plain single level "loyalty club" schemes.

In the absence of new tools built to enable global market transparency, commercial practices have only experienced a slow evolution over the last decades as compared to the revolution in the underlying technological infrastructure. Hence, vendors still rely on intermediaries and industrial marketing techniques even though there has been a certain shift from vertical intermediaries (e.g. wholesalers) to horizontal intermediaries such as generic search engines and confined market places (also referred to as "platforms"). But fundamentally there has been no change in marketing and sales methodologies.

The missing development of commercial practices matching the needs and possibilities offered by modern information technology has had several negative consequences:

Cost and complexity for vendors

The cost of platform based intermediaries services is extremely high - typically consuming 15-35% of a vendors total revenues (e.g. Amazon, Ebay, Alibaba), or up to their full net profit for auction based ads (e.g. Google). This is a classical consequence of having monopolies blocking a wider market transparency.
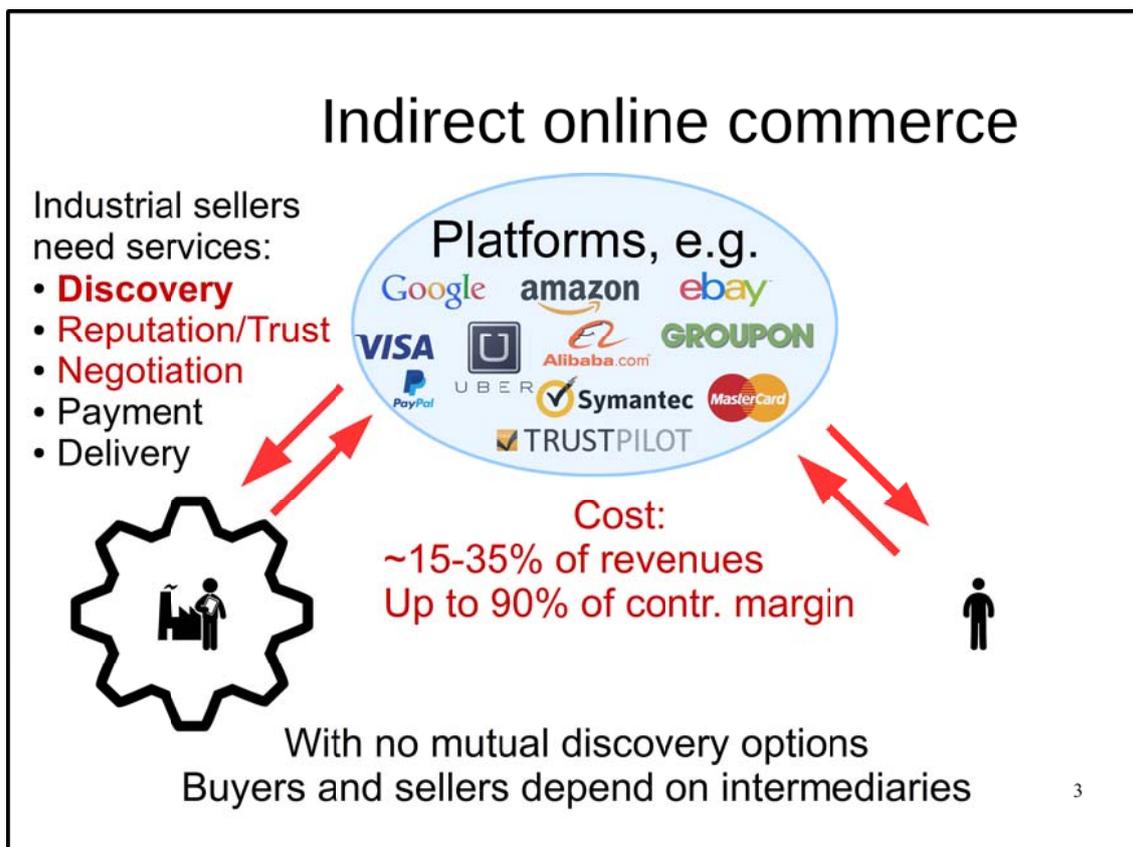
An equally unfortunate situation appears when a number of similar intermediary platforms are present in a specific market segment. In this case vendors have to be present on all of these

platforms in order to make themselves visible to all customers. And similarly consumers have to query all of the platforms to be certain to discover all relevant vendors. This can be particularly troublesome and expensive if the vendor operates in a market invaded by dozens of Groupon clones or "Get three offers for solving your task" intermediaries.

Lack of consumer choice, privacy, and market innovation

Although the intermediaries aggregate lots of personal information about consumers, they are still lacking the insight – or rather business model – to perform optimal matching of consumer needs to relevant vendors. Instead consumers are typically presented to the vendors that are paying the highest fees to the intermediaries. The uncontrolled sales of personal data from intermediaries to parties that consumers do not trust – or maybe even do not know – is a major privacy problem. And the lack of free flow of personal requests and data to vendors that might be able to offer the consumer new or personalized offerings is detrimental to market innovation.

Fraud & Identity Theft

The lack of market transparency from which the intermediaries thrive also makes it difficult for consumers to distinguish between serious vendors and fraudulent actors when browsing the web or receiving unsolicited emails. As companies do not have a native ID and presence on the internet, a complicated system of domain hosts and certificate authorities has been set up. This is covered in more detail in a subsequent section. However, these entities and their functions remain mostly unknown to ordinary consumers. Also the information associated with the current "certificates" is utterly insufficient for any consumer to make an appropriate risk assessment of the company associated with the certificate.

Direct Data Driven Commerce

Progression from the current industrial business layer on the internet to a globally transparent data driven commerce environment requires

- that companies and their services are easily and freely discoverable

- that consumers have tools allowing them to discover, process, and select company services

- that proper trust can be established between a company and its potential customers

Additionally, privacy concerns may dictate that consumers use tools or agents providing them anonymity or pseudonymity when sharing their data or otherwise interacting with vendors.

The emergence of a data driven market where company and product data travel freely to consumers naturally leads to a situation where much of the data analytics (e.g. purchasing decisions) will take place on the consumers side. It will also provide the consumers a better background for judging which companies they should allow to access their data for purposes that are better served with data analytics taking place on the vendor side (e.g. product innovation).

The possibility to choose between a very large number of product and service vendors also requires consumers to make use of filters based on their individual preferences. Consumers will never be able to manually browse through all potentially interesting commerce websites when searching for commonly available products and services.
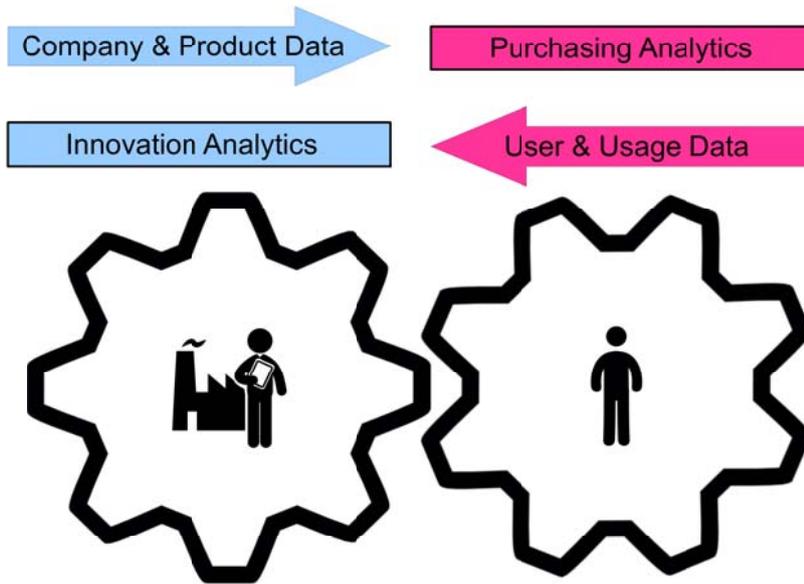
In the absence of a priori trust, vendors will often have to submit third party verified claims about their product and service qualities to pass such filters. This may be anything from generic industry trust marks, to specific competencies (e.g. welding certificates) or qualities (e.g. website certified as accessible for visually impaired persons).

In such a data driven market many resources can be transferred from promotional marketing to innovation and certification. Example: If a company today has had a website not suited for visually impaired persons it would require a marketing campaign costing in the order of 1M€ pr. country to profile it to the relevant audience once a new compliant website is launched. Contrary, in a data driven market, the website will automatically pass the filter to the desired audience upon being certified by a trusted third party.

Removal of the dependence on intermediary platforms simultaneously reduces the current "network effect", which means that new innovative commerce services must acquire large numbers of vendors as well as consumers to become competitive in relation to existing players in the market.

In a data driven market it will be easier for vendors to implement 1-to-1 relationship management based on specific reputation related information received directly from customers trusting them.

For a long time we have benefitted from the advantages of a 4-party model (vendor – vendors agent – consumers agent – consumer) when transferring money or communicating with each other by phone or email. Now the time has come to mature the e-commerce market in a similar way.

The need for a decentralized business service registry

Many of the principles powering todays internet were laid down more than 20 years ago. A major concern when designing the protocols for the technological infrastructure was resilience and independence from any specific network operator.

An example of this is the domain name system (DNS) which is used to relate computers to network (IP) adresses. The intention with the hierarchical domain name system was originally to create a categorization based on sector (company, education, government) or location (country). But today most registries are open to registrations from companies and persons from anywhere. Adding to the resulting distrust, ICANN that administers the root domain, in recent years has launched almost 1000 new "branding" top level domains. In this context the latest efforts to strengthen the validation of domain registrants can be considered a joke: e.g. a Danish person now has to use his 2-factor national eID to register a .dk domain. Whereas anybody from the rest of the world may register a .dk domain without any proper validation.

Browser vendors and so called Certificate Authorities have tried to substitute the missing trust in the domain name system by offering "Extended Validation" in connection with the issuance of server certificates. But this does not solve the above mentioned problem of cross country validations. Nor does the fact that any Certificate Authority registered by a specific browser may issue certificates in the name of any company in the world. There are many projects under way to mitigate the resulting risks, but none of these are without drawbacks.

Even if such solutions gain wide acceptance there are other reasons that makes it impossible to assess the proper legal identity of a company based on the domain associated with web-pages or email from the company. Today it is common practice to host company pages at complex URL's on markets, social media, or messaging platforms. And company identities may be even more consealed when consumers purchase services via agents e.g. as typically seen for airline tickets.

So the major problem remaining more than 20 years after the conception of todays internet is still that neither companies nor persons are natural entities able to claim identities on the internet as they can when being physically present.
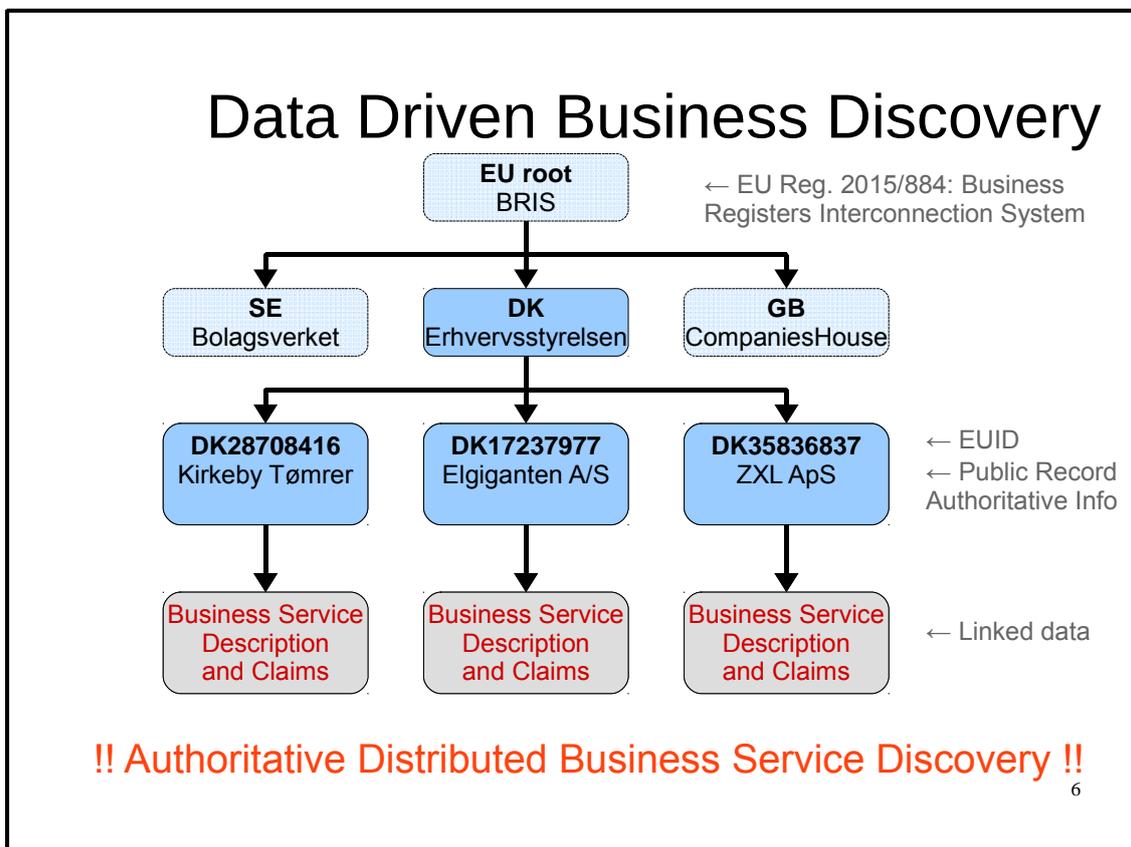
And while a lot of efforts have been directed towards providing individual persons with online identities, almost nothing (as seen above) has been done to assure the identity of companies online. The consequence has been an ever increasing level of phishing, fraud, and resulting cybersecurity issues.

Meanwhile – particularly in the EU – major efforts have been undertaken to convert the traditional paper based national business registries into modern registries allowing online registration for companies as well as data access to the public. Also there is now agreement on an EU wide company identifier, which could easily be extended for global use.

Some business registries have even started to register and list digital adresses such as email and web addresses. However, this practice is not recommendable as a company will typically possess lots of such entities for various purposes. Also the digital presence of a company and its services will generally be much more multifaceted in terms of both endpoints, services, and qualifications.

Hence, rather than extending the official business registries with such detailed information, it is suggested to add a single pointer to a freely chosen location, where the company may place all relevant information about itself and its services. This will enable a distributed authoritative business service registry in the same way as domain registrants can freely decide where to host the DNS records for technical services associated with their registered domains.

The resolution of company ID's would also follow a similar hierarchical scheme as the DNS resolution.

Data Driven Business Discovery

EU root
BRIS
← EU Reg. 2015/884: Business Registers Interconnection System

SE
Bolagsverket

DK
Erhvervsstyrelsen

GB
CompaniesHouse

DK28708416
Kirkeby Tømrer

DK17237977
Elgiganten A/S

DK35836837
ZXL ApS

← EUID
← Public Record Authoritative Info

Business Service Description and Claims

Business Service Description and Claims

Business Service Description and Claims

← Linked data

!! Authoritative Distributed Business Service Discovery !!

6

By inserting its identifier in the header section of emails and web pages as well as in their business service registry, a company can make it easy for web and email clients to verify the content provider via a lookup in the latter. Such verifications may function even when the content is being delivered via a third party website (e.g. market place, social media or agent) authorized by the original service provider in its service registry.

Depending on the information publicly available from the relevant business authority, clients will will also be able to provide a facts backed visual indication of the verification in a similar way as current browser extensions can use informations from social sites, e.g. as the rapportive.com extension shows LinkedIn data for senders next to the mails sent to you.

UDDI failed! Repeating a mistake?

Between 2001 and 2005 several large IT companies led an effort to create a universal structure for registering and locating web services. The service discovery challenge addressed by UDDI was similar to the proposal outlined here. But there are major differences in the scope and approach.

UDDI:

- was mainly aimed at B2B commerce
- required a set of independent replicated registries supposed to be run by large corporations
- was relying on SOAP/WSDL rather than more scalable restful protocols
- involved a very complex specification

In contrast, this proposal is focusing at B2C, relies directly on the existing authoritative legal registries, is generically distributed, and should be build as simple as possible using appropriate protocols and data formats for specific use cases. Companies should be able to specify their services in a variety of ways as long as they are easily machinereadable. Such data format flexibility is already offered by other service providers today, e.g. price comparison services, that accept data from product vendors in various structured formats from csv to xml.

It will be the task of anyone wanting to gather and use various business service data to implement pre-processing to format the data as appropriate for their specific analytics purposes.

Implementation overview

Implementation af this proposal requires efforts from:

- the various EU companies registers
- vendors, i.e. companies wanting to exhibit their services and qualifications
- service providers seeking to gather and analyze company and business service data

The efforts required from service providers will have major elements in common. An initial effort to develop open source solutions and common schemas for the most important generic use cases would therefore be beneficial to a successful implementation of this proposal.

Kickstarting effort

As many service providers will need similar systems to aggregate, filter, and analyze the combined data from the business registers and the companies business service records, the availability of an open source generic program providing the basis for this processing would be a significant catalyzer for achieving overall success.

Another significant task will be to standardize schemas for exchanging information in machine readable formats, but this activity would eventually be necessary irrespective of how mutual discovery of vendors and consumers may be implemented. Specifically the practices regarding company identifiers and semantics used by the European companies registers were not previously coordinated. Fortunately such a coordination is currently secured on the basis of Directive 2012/17/EU concerning cross-border company groupings and mergers.

Smaller tasks would be the provision of easy to use business record editors as well as extensions for browsers and email clients enabling consumers to easily authenticate and assess vendors.

Companies registers

The only basic requirements to the national or local companies registers is that:

- The person(s) allowed to update a companies records shall be able add a URL pointing to the companies business service record. This could either be possible by an explicit manuel

entry – or implicitly by authorizing a registration request from a business service provider.

- Companies records – at least for companies that have chosen to be discoverable - must be freely discoverable and accessible. (which is currently the case in at least Denmark, Norway, Belgium, Romania, Ireland, and United Kingdom. Source: OpenCorporates)

Vendors

The only thing companies have to do to become discoverable on the internet is to self-host or have a service provider to host a business service record that explains the nature of their services, any relevant qualifications, and preferred access methods. Non technical users may do this by using an open source business record editor or an editor made available by their chosen service provider. By further adding their company ID in the header of webpages and signing messages with a private key matching a public key placed in their service record, they can furthermore reduce (or next to eliminate) the risk of their customers becoming victims to phishing attacks.

Service providers

Typical service providers include Personal Information Management (PIM) solutions for consumers, Business Intelligence (BI) solutions for banks, investors, and competitor benchmarking, as well as Trust Services providing various types of certifications and reputation services that may be used as filters by PIM and BI solutions.

Service providers in this new ecosystem have the most comprehensive task of aggregating and maintaining up-to-date data from both the public registries as well as from companies associated business service records. Depending on the purpose of the service providers application, the required amounts of collected data as well as refresh cycles may vary significantly.

Hence, the emergence of a broad community of SME services would be greatly facilitated if a kickstart initiative did not only include the development of open source software, but also running one or more instances of a shared company and business service aggregation and filtering engine.

Benefits

Vendors, consumers and various types of service providers for either of these will all benefit from this proposal. The benefits fall in two major categories:

- Authentication and Trust
- Service Discoverability

Authentication and Trust

By enabling companies to assert their identity directly based on their legal records, they not only save time and money. They can also build trust, not only making customers more efficiently aware about phishing attacks, but also by providing prospective customers up-to-date verified business information about themselves.

This eliminates the need for dubious validations by certificate authorities around the globe that do not possess authoritative information about the company nor provide the means for companies to generically bind business service records to their identity assertions. At the same time consumers will have significantly better options for not becoming victims to phishing sites.

Service Discoverability

For *companies* offering products and services there are both practical and financial advantages of the global discoverability offered by the proposal:

- Write and post a business service record in *one* single place. No more need to register and supply company and product information individually to many recipients.

- Zero cost for vendors to to have themselves and their services discovered by consumers – as opposed to companies typically spending 30% of their total revenues for this purpose today.

- Address the increasing amount of consumers, that opt out of push marketing by using add-blockers and tracking-blockers.

Analytics service providers offering services to consumers (e.g. PIM's), vendors (e.g. CRM), banks, insurance companies, potential investors and news agencies will all benefit from generic access to richer and more consistent data about companies.

- Any new PIM solution will be able to discover all companies and convey purchasing intents to any relevant vendor on behalf of its connected consumers. Hence there will be much less "network effect" to overcome for new PIM solutions than in the current market landscape.

- New and existing price comparison sites and similar services can easier develop into more advanced services that enable filtering or differentiation on other vendor and product aspects than price.

- Doing "Big Data" analytics or creative apps involving data from business registries combined with business service data from large numbers of companies becomes significantly easier and less error-prone than today.

By stimulating the development of user centric service providers such as PIM's, this proposal will *indirectly help consumers* to gain efficiency and more leverage in interactions with vendors.

- Consumers will be not be subjected to various manual eyeball/finger interfaces to each vendor but will be able to discover and interact much more efficiently with businesses through agents such as PIM's or next generation price comparison sites.

- Consumers will only be directed to sites that meet their preferences in various aspects. Very importantly, this will turn the much debated "consent"-negotiation question upside down, as vendors not meeting a consumers preferences will be deselected by the users agent.

- User centered services as facilitated by the proposal may provide consumers with several alternative models for managing their privacy while appearing trustworthy to vendors. The users agent might act as a traditional butler having knowledge to act intelligently on behalf of a user. It might also be implemented in a way where it acts as a plain router with zero knowledge about the information being transferred. Or any model in between including self hosted services and advanced services using calculations on encrypted data.

- Via PIM's consumers may replace numerous vendor loyalty clubs with their own portable "Loyalty Club of One" that facilitates 1-to-1 relationships with any vendor they choose.

Conclusion

The internet was designed as an efficient technological communications platform between servers and people knowing these. However, it was never designed for many of todays use cases, e.g. global e-commerce as discussed in this paper. The relationship between servers and companies was poorly adressed while discovery of business services was not adressed at all.

Fortunately companies like Google, Amazon and lately Uber stepped in to create partial patches to the discovery problem, while others such as Verisign and Comodo stepped in to facilitate some trust in the binding between servers and legal companies.

However, several symptoms are now emerging that indicate severe deficiencies of this operating model for the internet in relation to globalization of commerce: Vendors experience increasing "middle man" costs. Consumers are starting to block ads and worry about their privacy. Start-ups wanting to innovate and challenge existing market players are strongly up against "the network effect". And this all takes place in an environment with an increasing number of skilled fraudolous actors.

Until now governments and various interest groups have been focusing on addressing these various symptoms individually. This has been without much success, apart from generating a negative spirit between certain government agencies, some of the prevailing discovery platforms, and citizens concerned about freedom of speech as well as privacy.

Changing the focus to address the underlying problems of indirect company identity assessment and centralized service discovery is a much more likely route to obtain an internet business layer that facilitates a reduction or elimination of the current problems.

The author suggests a simple method to establish a distributed but authoritative business service directory in a similar way that the DNS system supports the technical layer of the internet.

The huge benefits for all parties concerned (vendors, consumers, analytics services) may hopefully lead to a collaborative design and implementation effort, rather than the current controversy between big data analytics and privacy proponents that thrive under the current market conditions.

In particular as regards the European Union, the current proposal can be seen as a natural succession to the efforts that have already been undertaken to align the national business registers and provide open access to corporate data. At least 6 countries in the European Union are already technically able to support the proposal with minimal additional efforts.

NOTE: While this paper is focused specifically on (what is currently considered) e-commerce, a similar distributed discovery method would likely also be of benefit in other sectors. E.g. as a basis for discovery and trusted interactions between patients, doctors, hospitals, and the pharmaceutical industry.