Trust Models Go to the Web: Learning How to Trust Strangers

PASQUALE DE MEO, DICAM, University of Messina, Messina, Italy
YLLI PRIFTI, CMS, Birkbeck - University of London, London, United Kingdom
ALESSANDRO PROVETTI, SPS, University of Milan, Milan, Italy and CMS, Birkbeck - University of London, London, United Kingdom

We study emerging traits of interpersonal and social trust in online social networks of needs (OSNNs), where trust interactions start online and evolve into in-person meetings. We present a lightweight web scraping solution to harness data from online social networks; thanks to it we were able to monitor a nation-wide portal for childcare and see the evolution of online reviews from both families and carers. We analysed the data by first considering topological information to test centrality metrics as proxies for trustworthiness. Next, we focused on features/profile analysis and tested the Castelfranchi–Falcone trust model from psychology (CF-T), fitting it to online reviews of childcare services. Even though such reviews are relatively scarce and seemingly skewed, we feature-engineered the CF-T model to predict the evolution of reviews, treated as proxies for trust. By aggregating CF-T scores at the regional level, we discovered a strong correlation with per capita GDP, which suggests that high levels of trust in social networks of needs reflect social capital.

CCS Concepts: • **Applied computing** → *Computers in other domains*; • **Information systems** → **Social networks**; **Social recommendation**; **Personalization;**

Additional Key Words and Phrases: Security, trust, and privacy in social networks and media, social networks design and analysis, online reviews, user modelling

ACM Reference Format:
Pasquale De Meo, Ylli Prifti, and Alessandro Provetti. 2025. Trust Models Go to the Web: Learning How to Trust Strangers. *ACM Trans. Web* 19, 2, Article 12 (March 2025), 26 pages. https://doi.org/10.1145/3715882

1 Introduction

Trust is a concept that spans a multitude of disciplines. A common, and fundamental, question in the literature is how to define it and how the various definitions connect across the different disciplines. The quest for a general notion of trust becomes even more complex when considering the two relatively recent scenarios of computational social science and intelligent agents in an artificial intelligence context. The sheer volume of interactions and the blurring of what 'interaction' between online social cognitive entities means raise another question: 'how do online interactions correlate to interpersonal and social trust?' This is the core question when trying to fit a trust model

Author's Contact Information: Pasquale De Meo, DICAM, University of Messina, Messina, Italy; e-mail: pdemeo@unime.it; Ylli Prifti, CMS, Birkbeck - University of London, London, United Kingdom; e-mail: ylli@prifti.us; Alessandro Provetti, SPS, University of Milan, Milan, Italy and CMS, Birkbeck - University of London, London, United Kingdom; e-mail: a.provetti@bbk.ac.uk.

ACM 1559-1131/2025/03-ART12
https://doi.org/10.1145/3715882

to data representing interactions on online social networks, online peer-to-peer services, electronic commerce and many other online exchanges. We begin to address this question by reformulating it, without loss of generality, as follows: *What 'ingredients' of online interactions must be retained to capture interpersonal and social trust?* In addressing it, we need both models and experiments to reveal what are the aspects of online interactions that underpin interpersonal and social trust.

The definition and categorisation of web communities and social networks that retain interpersonal and social trust characteristics is the first significant contribution of this research. It will clarify when the application of trust models is appropriate to data retrieved from the web.

1.1 A Question of Data

Trust models discussed in the literature, especially in computational social science, are often either strictly theoretical or empirical. Empirical models, however, are often based on simulations or backed by data from one single source, which in turn may not retain the fundamental ingredients of social trust sought by the theoretical models.

On the web, the empirical approach seems natural: once a set of websites (portals, online social networks, etc.) that encourage, capture and display data representing social interaction becomes available, we can start attacking multiple research problems: the validation of existing models, the derivation of new models and—also importantly—the creation of unbiased, privacy-preserving data retrieval operations, to be made available to the research community [Prifti 2023]. Web data retrieval, however, is complicated by the technical complexity of dynamic and heterogeneous websites, those that continuously amass large quantities of semi-structured, often partially accessible, user-created data.

1.2 Models of Trust Emerging from Data

In the literature on web interactions and trust, we find two significant trends that essentially define the whole field. The first is about modelling user behaviour and applying models to typical Data Science tasks such as customer profiling, link prediction and shop-basket prediction. The second trend is about creating trust in web users, primarily by creating architectures that guarantee it.

Creating trust often is about money transactions, but also importantly about the veracity of personal interactions, such as verifying user profiles and fact-checking news. An exciting example of the 'creating trust' approach is the recent 'Eunomia' project,[1] which focuses on creating inherently trustable social web architectures. To do so, it leverages technology from peer-to-peer and blockchain to create users' confidence in the contents shared on their platform. We can summarise the wide range of trust 'degrees' that coexist on the web today by highlighting its extremes as follows.

- **Online personal services:** They can be booked online but are delivered in presence. Childcare is perhaps the scenario where stakes are highest, yet families trust the site to bring them qualified childminders (which, of course, may be further vetted in person, once contact is established).
- Somewhere in the middle sits the **Eunomia** approach, and in general approaches that strive to insulate users from the platform. Eunomia's own summary puts it as follows: "EUNOMIA is developing the first social media environment designed to prioritise trust over likes. (...) Its open source tools help the user quickly and confidently assess the trustworthiness of information shared through EUNOMIA without relying on any third-party expert or social media platform to do it for them."

[1]Please see https://cordis.europa.eu/project/id/825171

— On the opposite end of the spectrum sit **anonymous discussion platforms.** The 4Chan anonymous discussion portal is perhaps the best example: no one knows anyone. The stakes are generally low, but we notice that the user base, often considered a fringe 'anarco-libertarian' social segment, actually does trust the platform and its Japan-based localisation to guarantee their privacy and anonymity. This happens even though even though 4Chan users are in fact exposed to profiling and tracking by the long-reaching arm of investigative agencies.

In this article we report on our work on the childcare.co.uk portal and present our data-driven trust model from the data we collected there. The choice of analysing a personal care domain has strong practical and scientific motivations. On the one hand, in fact, childcare.co.uk exposes a wealth of personal and commercial data that, in a different context, could be easily seen as *over-sharing*. On the other hand, it is a commercial site with a clear identification of the roles involved in the childcare process. Mainly, there are *parents* who are looking for a baby-sitter, *carers/service providers* (baby-sitters) and, finally, the *jobs* advertised through the website. Each job is identified by the reviews that the service provider received and, consequently, we can determine how parents are connected to the service providers by inspecting the job entities. The abundance and completeness of the anonymised data we collected is crucial to the success of our study and ensures the correctness of our findings.

From a scientific point of view, there are multiple theoretical models of trust in the literature; as with most empirical approaches, our work is based on the monitoring of social interactions on a web platform. In many platforms, however, interactions are often only weakly related to trust, and in some cases, e.g., posting on a forum, they cannot be seen as indicating trust per se. Conversely, online interactions concerning personal care can be seen as *concrete evidence* of trust, which is what parents/guardians seek.

Thanks to our massive web data extraction campaign, scientific questions that were largely neglected in the past literature could now be addressed. To wit, is it possible to build a data-driven model of trust from (observed) user interactions along with information produced and disseminated by users? Is there a theoretical model of trust that can explain observed behaviour more effectively than others?

1.3 Main Contributions

The main contributions of this article can be summarized as follows:

(1) A new software platform for extracting data from **Online Social Networks of Needs (OSNNs)** is presented. Our system relies on two main components, namely OXPATH [Sellers et al. 2011], a superset of XPath that has also been used in commercial systems to extract data from the web, and *Dr. Web,* a Python package that we explicitly designed to overcome some limitations of OXPATH (which, in its open source version, has limited compatibility with modern browsers). Our architecture easily integrates with Dockers or Kubernetes; this makes our data collection procedure scalable. Furthermore, collected data are mostly semi-structured and thus naturally stored and processed by databases such as MongoDB. We believe that our architecture has the effectiveness and maturity to be adopted in scientific and commercial applications beyond those explored here.

(2) An in-depth analysis of OSNNs, probably the first of this type, is provided. Our analysis makes use of two specific kinds of information: trust measures and geo-referenced data. We believe that social science and network science researchers will be able to use our empirical findings in the design of new social network models, models where the notion of trust is crucial to, e.g., attract new subscribers and ensure the long-time survival of the social network itself.

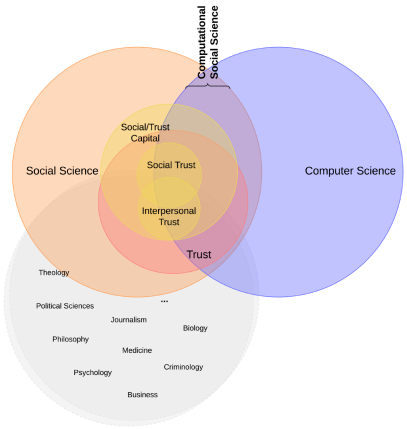


Fig. 1. A Venn diagram summarising the literature on trust as reviewed by Castelfranchi et al. and by Govier.

(3) Centrality analysis, a well-known technique from Network Science, is applied to OSNN data. In our study, we tested some of the mainstream centrality metrics on the collected datasets. As discussed in the rest of the article, centralities, while useful, seem to be only partially able to reveal trust. Clearly, real web datasets seldom come with a ground truth that enables one to assess trust measures. Yet, the results prompt us to conclude that standard centrality measures would need a (rather non-trival) generalisation to networks of heterogeneous entities in order to become effective in trust modelling.

(4) A combined 'centrality plus CF-T' assessment is presented. Unlike centrality, the CF-T model of Castelfranchi and Falcone [2000, 2010] defines trust as a combination of many non-topological, variables e.g., willingness and ability. We consider those surrogate variables in generating an accurate estimation of trust that is well aligned with notions of trust developed in psychology and cognitive science. Reviews and photos posted by users were collected and analysed by means of sentiment and image analysis techniques to calculate variables such as willingness and ability. Thus, our approach provides an alternative, data-driven measure of trust that is in good agreement with psychological theories.

2 Literature Review

Trust is one of those fundamental cognitive concepts that is embedded in and around our everyday life, affects and models our behaviours, yet we rarely have to consciously stop and think about *what is trust*. Within the scope of this research, which focuses on models of interpersonal and social trust emerging from data collected from online communities and social networks, we are proposing the Venn diagram in Figure 1 to help put this literature review in context.

The non-triviality of the Trust Modelling is underscored by the amount of literature dedicated to it, the interlink between the different disciplines, and the challenge to frame the discussion to a number of concepts and disciplines relevant to the research. We clearly separate the concepts of 'social trust' and 'interpersonal trust' seen as highly overlapping but clearly distinct concepts. Whilst the concepts are often used interchangeably in the literature, Govier [1997] summarises (p. 31) the clear difference between the two:

> *Social trust and interpersonal trust are different in significant ways. Interpersonal trust is based on experience, sometimes deep and intimate experience, with another individual. In some cases of social trust, we may have limited experience with the other persons involved; in others, we have none at all. (...) [D]espite these differences (...), it is neither an accident*

> *nor a logical mistake that the word 'trust' should be used across personal and broadly social contexts.*

There are clear traits of distinction, interchangeability and interlinking (i.e., how one affects the other) between 'social trust' and 'interpersonal trust' that are not widely discussed in the literature but emerged naturally during the analysis of the data collected as part of this research. So, we have deployed some of these characteristics to define a new group of online social communities where the traits of *core trust* [Castelfranchi et al. 2009, p. 102] are clearly and explicitly expressed. They see *Core Trust* as follows:

> *A set of mental states* $(MS - CTX, Y)$, *called Core Trust, with these components:*
> *— a set of X's goals and, in particular, one specific of them* (gX) *in order to trust Y;*
> *— a set of X's competence beliefs* $(B - ComX, Y)$ *on Y about* τ*;*
> *— a set of X's disposition beliefs* $(B - DisX, Y)$ *on Y about* τ*; and*
> *— a set of X's practical opportunity beliefs* $(B - PrOpX, Y)$ *on Y about* τ *at that given moment (time) and site (space).*

In Figure 1, we indicate that social trust and interpersonal trust are contributors to another concept: *social capital* (often referred to as Trust Capital). Whilst social capital is not the main focus of this research, we believe it is important to highlight that social capital is most commonly defined, e.g., by Portes [1998], as

> *the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalised relationships of mutual acquaintance or recognition.*[2]

The concept of social capital is important both in the sense that *"When a society has social capital, just about everything is easier because people can turn to others for information and assistance"* [Govier 1997, p. 152] but also *"(...) [social capital] must be built from [interpersonal trust]. (...), trust capital is a macro, emerging phenomenon; it must also be understood in terms of its micro-foundations"* [Castelfranchi et al. 2009, p. 29].
The two statements can be explicitly re-phrased as follows:
(1) An abundance of social capital will influence and promote more trust among individuals and hence higher social trust, and (2) interpersonal trust and social trust are fundamental components of social capital, and hence societies with high interpersonal and social trust will inevitably have higher social capital.

This concept is clearly stated by Putnam [Leonardi et al. 2001, p. 177], who suggests that, because of the nature of social capital and trust, societies will converge to an equilibrium of either high social capital and high trust or the opposite spectrum:

> *Stocks of social capital, such as trust, norms, and networks, tend to be self-reinforcing and cumulative. Virtuous circles result in social equilibria with high levels of cooperation, trust, reciprocity, civic engagement, and collective well-being. (...) Conversely, the absence of these traits in the uncivic community is also self-reinforcing. Defection, distrust, shirking, exploitation, isolation, disorder, and stagnation intensify one another in a suffocating miasma of vicious circles.*

The result of our work is to bridge the abstract model above (and Bordieu's vision of *society as a spectacle*) and Online Social Networks with actual users and data. By aggregating the trust values computed for each agent, we compute the aggregated social capital values and study their

[2]Whilst this definition of social capital is already in Bourdieu [1980, p. 248], we are referring to Portes [1998], who analyses the origins and various definitions of the term.

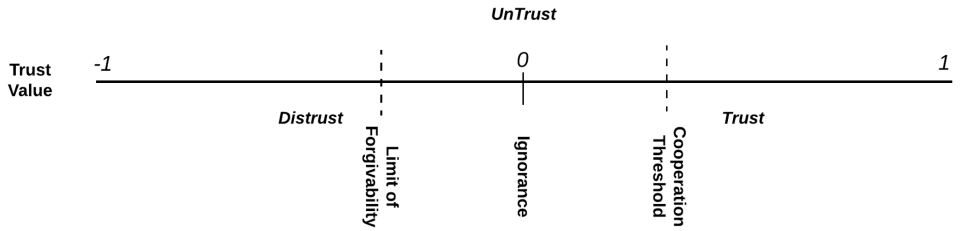


Fig. 2. S. March's trust continuum: The limit of forgivability.

correlation with other geographically distributed social measures, namely social capital and wealth generation.

2.1 Literature on Computational Definitions of Trust

Most of the research efforts on quantifying and computing trust are part of developments in information security. This concept is usually described as Computational Trust. Quoting from Wikipedia:[3]

Computational Trust applies the human notion of trust to the digital world.

Some of the models of trust discussed in the later sections of this section are developed in the context of computational trust. In our work we take a different approach, looking closer at the traits of trust emerging between cognitive agents interacting via digital means.

Marsh [1994] pioneered computational trust, in particular by introducing the concept of the *trust continuum,* which he further refined in Marsh and Briggs [2009]. He discussed trust as a continuum between −1, a situation of complete distrust, and +1, complete trust. Also, he clarified the difference between *Untrust* or *Lack of Trust* and *Distrust*. The former is a situation of ignorance, where there is not enough information to either trust or distrust. On the other hand, this is different from *Distrust*, where the Trustor firmly believes that the Trustee will perform an action that goes against his interest (intentionally or unintentionally).

In March's trust continuum spectrum, represented in Figure 2, we also find the 'cooperation threshold' put in perspective. For example, whilst looking for a babysitter on online sitter communities, we review many babysitters and trust them to be good with children, yet we look further until we find the few we trust beyond the cooperation threshold, and these are the ones we will contact to offer the job. Similarly, March discussed a 'limit of forgivability,' beyond which it is unlikely that the trustor would trust again and might have decided about the intentionality of the trustee's action going against the desired outcome.

Marsh's work is important not only because he discussed trust in its composing components (delegation, cooperation, forgiveness, regret, etc.) but also because he proposed a first formal model for trust. This consisted of concepts of *situations* $(\alpha, \beta, \ldots)$, *agent* $(a, b, c, \ldots)$, *knowledge* (e.g., x knows y $K_x(y)$), *importance* (e.g., of α to x, $I_x(\alpha)$), *utility* (e.g., of α to x, $U_x(\alpha)$), *basic trust* (i.e., the propensity an agent x has to trust someone in general, T_x), *general trust* (i.e., similar to basic trust but in relation to another agent y, $T_x(y)$), and lastly *situation trust* (i.e., the trust an agent x has on agent y to perform action α, $T_x(y, \alpha)$, with the situational trust expressed as

$$T_x(y, \alpha) = U_x(\alpha) \times I_x(\alpha) \times \widehat{T_x(y)}. \tag{1}$$

With the proposed formalisation in Equation (1) Marsh extends the concept of trust as discussed in Game Theory (e.g., as the quantifiable amount in the Prisoners' Dilemma) and Artificial

[3]Please see the Wikipedia entry for computational trust.

Intelligence (where trust is discussed in terms of situations and agents with cognition in terms of utility).

While Marsh approached trust in an algorithmic fashion, Castelfranchi et al. [2009] set up to build a non-reductionist model that also formalises the concepts of *Trust*, *Lack of Trust* and *Distrust* but goes further than Marsh's in separating *Distrust* from *Mistrust* based on the trustee's intentionality. They defined mistrust as follows.

Definition 1. **Mistrust:** Agent i mistrusts j to ensure $\neg\phi$ by performing action α if and only if:

(1) i wants to achieve ϕ;
(2) i expects that
- j has the opportunity to ensure $\neg\phi$ by performing action α AND
- j intends to perform action α AND
- the internal preconditions for the execution of action α by agent j hold AND
- the external preconditions for the execution of action α by agent j hold.

There is a clear distinction in this definition between *Distrust* and *Mistrust*. The former is the belief of the Trustor that the Trustee won't perform action α and hence the Trustor won't achieve the goal ϕ. In the latter, the Trustee is actively pursuing an action α that will achieve the goal $\neg\phi$. Mistrust is the active pursuit of actions for the Trustor *not* to achieve his goal. Castelfranchi breaks down trust into a number of components on another dimension to those discussed by Marsh. These are the components of *'basic trust'* broken down as (i) *competence,* (ii) *disposition,* (iii) *dependence,* (iv) *fulfilment,* (v) *willingness,* (vi) *persistence,* (vii) *self-confidences* and (viii) *motivation.*

In Falcone and Castelfranchi [2001] we see a clear distinction between *Trust* and *Social Trust*, with the latter requiring the trustee (or agent y) to be a cognitive entity. This is discussed in terms of delegation (delegation being the decision to trust), distinguishing between weak delegation (i.e., no presupposition in terms of agreement, deals or promises) and strong delegation (i.e., the presupposition of goal adoption by the trustee y). As they put it:

> *(...) social trust in the strong delegation, which is its typical and strict sense in the social sciences.*

Another contribution of their work is the discussion of the quantification of trust and the definition of the degrees of trust, as in the following definition.

Definition 2. **Degrees of Trust**: The degree of trust the cognitive agent X has on the cognitive agent Y about the situation τ is

$$DoT_{XY\tau} = DoC_X[Opp_Y(\alpha, g)] * DoC_X[Ability_Y(\alpha)] * DoC_X(WillDo_Y(\alpha, g)], \quad (2)$$

where

— $DoC_X[Opp_Y(\alpha, g)]$ is the degree of credibility of X's beliefs about Y's opportunity of performing α to realize g;
— $DoC_X[Ability_Y(\alpha)]$ is the degree of credibility of X's beliefs about Y's ability/competence to perform α;
— $DoC_X(WillDo_Y(\alpha, g)]$ is the degree of credibility of X's beliefs about (cognitive agent) Y's actual performance and can be broken down as

$$DoC_X(WillDo_Y(\alpha, g)] = DoC_X[Intend_Y(\alpha, g)] * DoC_X(Persist_Y(\alpha, g)].$$

In other words, the degree of trust can be expressed as a composition of beliefs and credibility of agent X on agent Y's opportunity, ability and competence, intentions and persistence to perform the action required to achieve X's goal. In conclusion, the 'degrees of trust' definition of Equation (2) is the model we adopt.

2.2 Trust Modelling in Social Networks

Trust has been widely employed in Social Network Analysis, and in particular, some authors suggest using the network of social relations between users in a community to propagate trust scores: in other words, starting from a relatively small set of *seed* actors in a social network, we can make use of the social connections between social network members to predict the users that a target user u can trust (i.e., the users that are not directly connected to u in the social graph).

In general, trust is categorised into *explicit trust* (i.e., we assume that in some social networks users can declare which members of the network they trust) and *implicit trust* (i.e., the level of trust between two actors in the social network is calculated by the similarities between user behaviors).

A number of approaches use knowledge of the topology of the trust network to infer trust values, e.g., Agreste et al. [2015], Golbeck [2005] and Ziegler and Golbeck [2015]. The work of Golbeck and Hendler [2006] is one of the most relevant as it implements a revised version of the **Breadth-First Search (BFS)** algorithm on the trust network. Each BFS step is performed twice: first, to find paths that lead from a source node i to a target node j, and then to backtrack trust values to i. Finally, the node i collects trust values and averages them to compute the trust score it would assign to j. Ziegler and Golbeck [2015] propose a procedure for spreading trust in a social network inspired by the *spreading activation* mechanism, which has been extensively studied in neuroscience [Quillian 1966]. In spreading activation, a starting node (called *seed*) is endowed with a certain amount of energy that is equally distributed among its neighbours; if the amount of energy that can be transferred from the seed node to each of its neighbours is greater than a threshold, then the seed node's neighbours are activated, and the energy recursively propagates through the trust network until all nodes are activated. The trust propagation process can die out prematurely if new nodes are not activated.

Trust in a social network has often been used to support *recommender systems*. It is in fact well known that recommender systems suffer from the *data sparsity* problem since the users of a platform consume (and rate) only a small portion of the items/services that the platform offers; an effective strategy to fight against data sparsity is to combine user ratings with trust values to generate more accurate user profiles. In turn, more precise user profiles are useful to predict with greater confidence whether a specific item is of interest to a specific user.

One of the first approaches that combines trust relationships and user ratings is SocialMF [Jamali and Ester 2010]; in detail, it integrates trust propagation with a probabilistic matrix factorisation method to generate high-quality recommendations. A further approach to mention is TrustSVD of Guo et al. [2015]; it applies an extension of the Singular Value Decomposition algorithm to combine ratings and trust values to produce more accurate recommendations. jao Chen and Gao [2018] describe the recommendation process through diffusion mechanisms originally studied in the field of statistical physics (e.g., heat diffusion) and show that the effectiveness of the diffusion mechanism in selecting objects relevant to user needs can be significantly improved by incorporating user trust values. Khaledian and Mardukhi [2022] provide a matrix factorization method that integrates rating information and trust relationships to overcome both data sparsity and the cold start problem (i.e., the lack of information for users who only recently joined the platform). In a subsequent work, Khaledian et al. [2023] introduce a new dictionary learning technique based on trust information (called TrustDL). TrustDL combines multiple sources of information (e.g., trust statements and ratings) to alleviate both cold start and data sparsity.

3 Finding Instances of Trust and Evaluating Models

Trust evaluation is, in general, concerned with trust metrics and trust models. Ziegler [2009] provides a categorisation for trust metrics based on their features, shown in Figure 3.

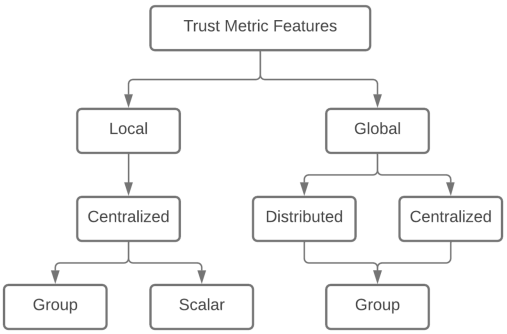


Fig. 3. A classification of trust metric features proposed in Ziegler [2009].

Summarising our findings with the literature review, out of 18 trust models reviewed, 12 used graph-based evaluations. Network Centrality measures are at the core of social network analysis and often the basis for trust metrics [Barbian 2011]. Next, PageRank, Trust Centrality (see again [Barbian 2011]) and EigenTrust are all trust metrics derived from network centrality measures. We now have social trust data from multiple sources. In this section, we are going to

— use data from childcare.co.uk to build graphs and run social network analysis;
— use the collected data to create multiplex networks;
— calculate the value of the trust (see equations) for agents that have received reviews, use factorisation machines to predict trust for agents without reviews and evaluate its predictive power over time; and
— discuss the trust relations between different online social networks of needs, the total trust and its variations over time.

As mentioned above, it is common in the literature to analyse reviews and derive trust metrics. There are well-known limits to this approach since reviews and recommendations are known to suffer from positive bias of opinions [Askay 2015].[4]

We have defined and tested a novel approach to trust that is attribute based, does not rely on reviews and is more comprehensive in the network evaluation. In the later sections of this section, we compare the results between our attribute-based trust models and models derived by analysing the network characteristics of the reviews.

3.1 Overview of Our Web Data Extraction System

We designed an ad hoc software solution for collecting data from the web; it is specifically tailored for OSNNs are web platforms where interactions between members are mostly face to face. We first turned to OXPATH [Sellers et al. 2011], a superset of the XPath query language, which has been used commercially in many web data extraction projects. OXPATH is an excellent choice in a number of scenarios, such as automatic full website extraction. However, we encountered several situations where OXPATH was not the most appropriate choice (e.g., OXPATH has limited support with modern web browsers).

To get around OXPATH's limitations, we developed an open source Python package.[5] Our web scraping system can be easily integrated with Kubernetes and Docker, and thus, it is possible to use commodity hardware to scale our web scraping system. It is worth noting that the technologies mentioned above allow new nodes to be added to existing clusters (and they also allow for creating

[4]The low number of reviews and how it negatively correlates with trust demand is discussed in Prifti [2023].
[5]The entire code base is available from https://github.com/yllipifti/dr-web-engine

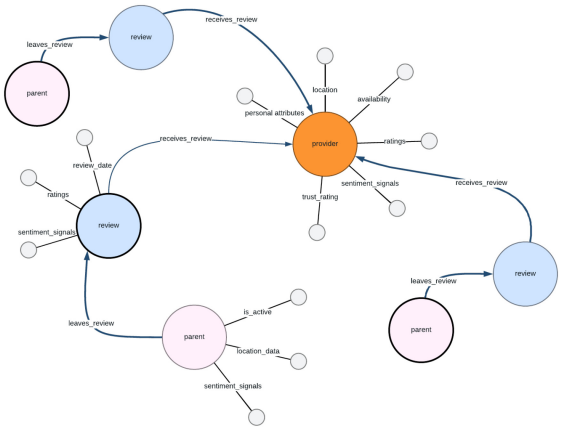


Fig. 4. A graph-based representation of data extracted from the *childcare.co.uk* platform. The figure highlights three types of nodes (parent, review and provider) along with their links. For each node we also display some of its attributes.

new clusters); such an increased computational power can join the already running data retrieval process, and it can take over the tasks in sync with the existing nodes.

Finally, the output of our system is semi-structured and can be easily integrated with natively semi-structured database systems such as MongoDB. More details on the implementation of the system can be found in Prifti [2023].

3.2 OSNNs of Needs

Right from the landing page of childcare.co.uk, we are presented with the following text:

> *Childcare.co.uk is an award-winning online social networking platform for parents, childcare providers and private tutors with over 2.5 million members. Start your childcare search or childcare job search today.*[6]

Apart from efficiently describing the utility of the social network, this sentence also highlights the nodes and relations of the graph we built. In fact, we have: • **providers**, • **parents** and • **jobs** as nodes and • **awards job** and • **receives service** as links. The jobs are represented by the reviews received by the service providers, and they link parents to providers.[7] In fact, this is a common pattern among all OSNNs. Figure 4 shows a visual representation of nodes and links in childcare.co.uk.[8]

We build the graph using a distributed graph database, specifically Neo4j [Miller 2013]. The result is a disconnected directed bipartite multi-graph made of 80,325 provider nodes, 22,885 nodes representing parents, 24,087 nodes representing reviews, 2,512 nodes representing UK locations, 9 region nodes, 3 country nodes and a total of 196,220 links. The graph is directed, because there is clear direction from parents leaving reviews to providers receiving reviews. The graph is also bipartite, since we can separate parents and provider nodes from review nodes. Furthermore, there

[6]Visited on 25 November 2024.
[7]From now on we will use *review* instead of *job* and [*leaves review*, *receives review*] instead of [*awards job*, *receives service*].
[8]The Jupyter notebook at https://github.com/yprift01/4chan-data-project/blob/main/colab/CreateChildcareGraph.ipynb implements network creation. It includes read-only access to source data but requires a new Neo4j server. Neo4j offers a free tier for its cloud-native Distributed Graph Database that can be used to replicate the research in this section.

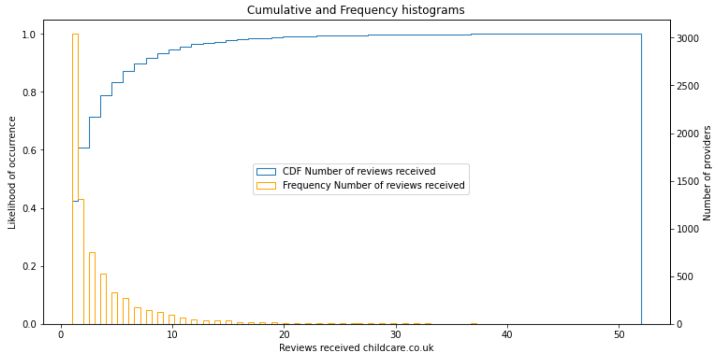


Fig. 5. Childcare provider received review distribution.

are no cycles since only parents can initiate and leave a review, and providers can only receive reviews once they have delivered their services.

In Listing 1 we provide Cypher code to extract the data associated with the parents, the reviews that the parents leave and the providers who receive those reviews.

```
match (p:parent) - [lr:leaves_review] - (r:review) - [rr:receives_review
    ] - (pr:provider) return p, lr, r, rr, pr
```

Listing 1. The Cypher code we employed to extract relationships between parents, reviews and suppliers.

Unusually for today's OSNs, most providers have not received any reviews. There are 80,325 active providers in the network, but only 7,143 (i.e., 8.9%) have received a review. There are 24,087 reviews in the network.

The chart in Figure 5 shows the frequency and cumulative distribution of the number of reviews received by each provider. About 95% of those who have received reviews have in fact only received 10 or less. Relations in high-trust-demanding networks are harder to form and supposedly long-lasting. The graph density is only 3.29×10^{-5} with average degree of 1.78.[9]

A visual representation (see Figure 6) of a random walk of 3,000 provider nodes and their review and parent links shows a highly clustered and disconnected graph. Unsurprisingly, the graph is provider centric and shows characteristics of an egocentric network. Providers create loyalties for long-term relations, and parents will only work with a few trusted providers. The code snippet in Listing 1 shows the Neo4j cypher query used to generate the graph in Figure 6.

We can expand the network analysis to include the locations where the services are provided and where job exchanges (i.e., reviews) have happened. Locations are represented by the outcode of a UK postcode[10] and their links into region nodes, districts and countries.

Whilst the network of parents and reviews is disconnected (i.e., not enough parents leave reviews to distinct providers to form a connected network), the location-centric graph is connected since most providers offer services in multiple locations. In fact, the results of community detection, as defined, e.g., in Clauset et al. [2004], on the two different views of the graph show the different natures of the two. There are 6,403 communities detected in a provider-centric network but only 456 in a network that additionally features links to locations.

The Betweenness centrality on the connected network that includes the locations is greatly dominated by the geolocation structure and association of postcode locations and regions, featured

[9]The graph density is slightly higher, 4.13×10^{-5}, when you include locations, regions and respective links.
[10]Please see https://en.wikipedia.org/wiki/Postcodes_in_the_United_Kingdom

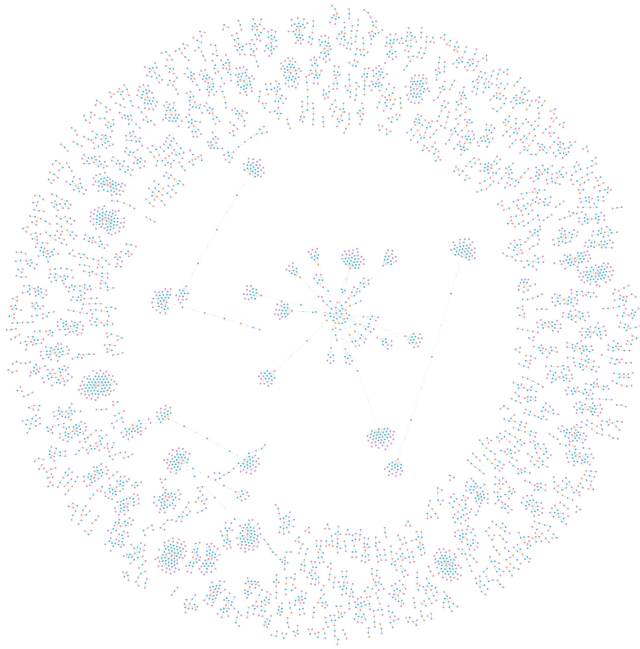


Fig. 6. A graph-based representation of the relationships between parents, reviews and providers drawn from the *childacare.co.uk* platform. The data consists of a sample of 3,000 items (parents, reviews and providers) randomly extracted from the data collected by our scraping system.

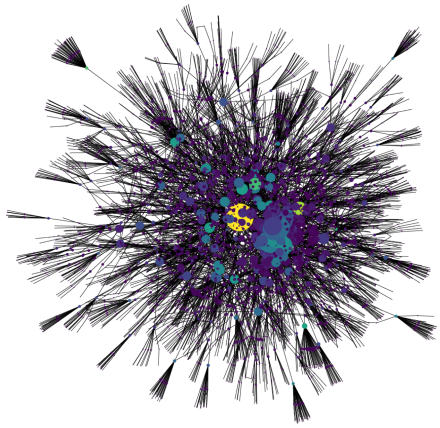


Fig. 7. Eigenvector Centrality from a sample of 5K nodes from the childcare.co.uk network in London. The size of each circle is proportional to its centrality, and, in particular, the larger the radius of a circle, the greater its centrality.

in the new graph. The Eigenvector centrality, shown in Figure 7, instead weights the in-links and weights of the connected nodes. The providers with the highest PageRank are the ones that are connected to reviews from parents who have worked with other providers and connected to locations where many providers with high PageRank offer their services.[11]

[11]Here *PageRank* and *Eigenvector centrality* are used interchangeably.

(a) Betweeness Centrality (b) Closeness Centrality

Fig. 8. Betweenness and Closeness Centrality measures from a sample of 5K nodes from the childcare.co.uk network in London. The size of each circle is proportional to its centrality, and, in particular, the larger the radius of a circle, the greater its centrality.

Figure 8 shows the differences between the two centrality measures. Betweenness centrality is, in general, dominated by locations. There is a direct correlation between providers with the highest betweenness centrality and providers with the largest numbers of reviews received. The same is not true for the providers with the highest PageRank score.

Interestingly, none of the providers with the highest betweenness centrality appears in the list of the providers with the highest PageRank centrality. The top 3 providers with the highest closeness centrality feature in the top 10 providers with the highest PageRank centrality. This pattern did not change when we ran the same calculation on other UK regions or when running on the full network rather than using sampling. The Betweenness centrality is more closely associated with a higher number of reviews, whilst the PageRank centrality seems not to feature any of the providers with the highest number of reviews. We found Closeness centrality values about an order of magnitude higher than the other two centrality measures, as shown in Figure 8.

Another pattern emerging from the regional networks is that the eccentricity measure is always 8. This measure is a network representation artefact, which is the total number of links between the two most distant nodes (by definition). In fact, 8 is the number of links in the following path: *parent* $\rightarrow_1$ *review* $\rightarrow_2$ *provider* $\rightarrow_3$ *location* $\rightarrow_4$ *region* $\rightarrow_5$ *location* $\rightarrow_6$ *provider* $\rightarrow_7$ *review* $\rightarrow_8$ *parent*.

These characteristics can be seen in Table 1. Closeness centrality shows a much smaller standard deviation from the mean than the others; its normalised standard deviation is 0.00132. By comparison, the standard normal deviation for the Betweenness centrality and PageRank centrality are larger by a factor of 8, at 0.009613 and 0.00974, respectively.

The childcare network is disconnected in its triad of nodes, parents, reviews and providers. It is, however, connected by locations where services are provided, and this allows us to derive some conclusions from the centrality measures. Yet, most of the providers that haven't received any reviews are left out of these calculations. A trust measure that weights the features of the individual providers would provide more insight into the trustworthiness of each provider.

4 The Attribute-based CF Trust Model

Most of the trust measures in the literature, and those analysed in the previous section, are based on networks, centrality measures and an evaluation of probability and likelihood based on

Table 1. Top 10 Providers by Centrality Measures

Betweenness Centrality				Eigenvector Centrality				Closeness Centrality			
Prov	Cent	R	L	Prov	Cent	R	L	Prov	Cent	R	L
73,286	.0.016	41	1	2,449,523	0.011	3	4	1,093,188	0.194	14	2
109,854	0.016	44	1	848,517	0.009	10	3	1,252,990	0.194	2	2
70,763	0.014	38	1	340,251	0.009	2	3	1,095,304	0.194	1	2
1,073,016	0.011	28	1	2,313,178	0.009	2	4	970,779	0.194	1	2
14,005	0.010	25	1	2,255,913	0.008	1	3	837,197	0.193	18	2
212,922	0.009	22	1	1,402,304	0.008	4	3	109,854	0.193	44	1
339,021	0.008	21	1	214,141	0.008	8	3	1,904,057	0.193	11	2
105,020	0.007	18	1	1,983,931	0.008	5	3	130,340	0.193	9	2
300,229	0.007	18	1	1,093,188	0.006	14	2	468,559	0.193	9	1
123,798	0.007	16	2	1,252,990	0.006	2	2	2,623,136	0.193	9	2

R: number of reviews received and L: number of locations where they provide their services.

previous experiences expressed as reviews and recommendations. However, trust is much more than subjective probability [Castelfranchi and Falcone 2000]:

> *...trust is about somebody: it mainly consists of beliefs, evaluations, and expectations about the other actor, his capabilities, self-confidence, willingness, persistence, morality (and in general motivations), goals and beliefs, etc. Trust in somebody basically is (or better at least includes and is based on) a rich and complex theory of him and of his mind. Conversely, distrust or mistrust is not simply a pessimistic esteem of probability: it is diffidence, suspect, negative evaluations relative to somebody.*

As we have seen, many approaches make use of a graph-based representation to model trust networks in which humans correspond with nodes of the network; edges instead specify the trust between a pair of members of the trust network, and in general, edges are directed and weighted. Network-based approaches display several strengths, and one of their strongest points is that the importance of a node in a network (which can be measured by one of the many measures of centrality available in the literature) measures the level of trust of the individual that that node represents. Network-based approaches have several limitations: in almost all real cases, in fact, we lack a ground truth to guide us in choosing the best centrality measure. A second limitation is that real social networks arise from the interaction of heterogeneous entities that are often described by some explanatory attributes. In case of OSNNs we highlighted three types of entities (parents, reviews and providers), each of which is endowed with specific attributes. The generalisation of standard centrality measures (such as PageRank or closeness) to trust networks made up of heterogeneous nodes with attributes is not a trivial task. Furthermore, the network topology plays a key role in the calculation of the centrality measures, with the undesirable consequence that some centrality metrics may be more negatively affected by topology than others.

The trust model proposed by Castelfranchi and Falcone (CF-T or CF Trust) is, in our opinion, able to address the criticisms above. The CF-T, in fact, is a socio-cognitive model that relies on several types of information, such as opportunity, ability and willingness, which we reported above. Rather than considering trust between two individuals as a variable to be measured or predicted, the CF model relies on a number of surrogate parameters that, taken together, provide a quantitative measure of the degree of trust of two individuals that is well aligned with the notion of trust defined in cognitive and behavioural science. This is a significant step forward because it alleviates the lack of ground truth: in fact, instead of using missing information about pairwise trust scores (which is often difficult or impossible to obtain), we use surrogate variables

to quantify the degree of trust of two individuals. Moreover, high-quality surrogate variables can be obtained at a reduced cost: in our application scenario we collected abundant data on the actors involved (parents, providers and reviewers), which were subsequently analysed through sentiment and image analysis, and this enabled us to measure, with a high degree of accuracy, the value of opportunity, ability and willingness. Furthermore, the CF model is independent of the relationships between the involved actors and is based only on attributes. A further advantage of the CF model is that its calculation is independent of the number of reviews received by a provider and, thus, it does not suffer from the problem of missing data that might affect centrality-based approaches operating on sparse trust networks.

At the top level, the CF model of trust is the sum of the three elements of the CF analysis: opportunity, ability and willingness (or in the extended form 'opportunity,' 'ability and competence,' 'intent and persistence'). In this section, we evaluate trust scores for childcare providers over the three 'axes' described by the CF model. In doing so, we intentionally won't use reviews in the calculation because *reviews already represent an expression of trust* and are a way to validate the results. The number of reviews received is used as the independent variable to validate the results. We expect a positive correlation between the computed value of Attribute-based Trust (henceforth, CF Trust) and the number of reviews received. Further, when using a non-linear solver to calibrate the belief parameters, we use the number of reviews received as the (solver) objective.

Trust is situational and subjective. We formalise it as a function $DoT_{x,y,\tau}$ that represents the degree of trust for the cognitive agent x (or the *Trustor*, or in our case *parents*) toward the cognitive agent y (or the *Trustee*, or in our case *provider*) for the situation τ (in our case the *care service*). We are computing the degree of trust for care providers as (hypothetically) perceived by parents looking for care services. The reviews in the platform are seen as an expression of trust for a situation and are subjective since both a parent and a provider are involved. Our computation is objective since we aim to generalise the perceived trust for all parents. The parameters $\alpha_{i,x,y}$, $\beta_{i,x,y}$ and $\gamma_{i,x,y}$ are the subjectivity parameters. Each parent would have different beliefs, hence different values of how much, for example, the number of qualifications (i.e., $\beta_{\text{qualifications},x,y}$) contributes in the evaluation of the ability and competence a provider has to offer their services, or when evaluating its willingness and persistence (i.e., $\gamma_{\text{qualifications},x,y}$). The value of 'number of qualifications' for provider y is the $T_{\text{qualifications},y}$ of the CF-T model.

Starting from a position of no knowledge, we can only decide whether an attribute contributes to the value of opportunity, ability and willingness based on common sense. Whilst this decision can be subjective, it does not undermine the trust evaluation since we calibrate the parameters as the next step of this evaluation. Table 2 shows the attribute contributions to the degrees of trust calculation.

Attributes are retrieved from the provider pages as exposed in the childcare platform. Derived attributes are used in the calculations. We have intentionally left out *Rating Count* and *Rating*, since these will be used for model validation. *Positive Sentiment* is derived from analysing the sentiment of the *About Me* text and using the positive score.

Next, we deployed the off-the-shelf Vader [Loper and Bird 2002] sentiment analyser[12] to evaluate the texts. The final step is the usage of DeepFace [Serengil and Ozpinar 2021] to analyse the profile image and derived several attributes including *Happy Emotion*, *Age*, *Gender*, *Race,* etc. We are only using Happy Emotion for trust evaluation purposes; however, the other attributes can be useful for trust segmentation and bias analysis.

We have intentionally included attributes that positively contribute to trust. The degree of trust is a number in [0,1]. The calculation can be expanded to include negative trust in a range between

[12]Please see the description on www.nltk.org/_modules/nltk/sentiment/.

Table 2. Attribute Contributions

Attribute	Derived Attribute	Opportunity	Ability	Willingness
About Me	Length	-	-	X
	Positive Sentiment	-	X	X
Profile Image	Happy Emotion	-	X	X
	Age	-	-	-
	Gender	-	-	-
	Race	-	-	-
My Qualifications	Length	-	X	X
My Local Schools	Yes/No	X	-	X
Availability Timetable	Yes Count	X	-	X
Last Updated Timetable	Days Ago	X	-	X
My Documents	Length	-	X	X
Rating Count		-	-	-
Rating		-	-	-

[-1, 1]. Other attributes, such as *Anger Emotion,* can be included with a negative contribution. Missing values from some attributes or out-of-range values can also have a negative contribution rather than 0.

Each attribute is normalised as follows:

$$N_i = \begin{cases} \frac{V_i}{Max(V_i) \times Max(i)} & \text{if } 0 \leq Max(V_1) \leq 1 \\ \\ \frac{log(V_i)}{log(Max(V_i)) \times Max(i)} & \text{if } Max(V_i) > 1 \end{cases}, \tag{3}$$

where

N_i is the normalised value between 0 and 1 for the ith attribute;

V_i is the value of the ith attribute;

$Max(i)$ is the number of attributes, and

$Max(V_i)$ is the highest value for V_i among all providers for the ith attribute; $\log(\cdot)$ is used on values greater than 1 to limit the impact of out-of-range values.

Equation (3) encapsulates attribute value normalisation $V_i \div Max(V_i)$ and coefficient normalisation $1 \div Max(i)$, since the coefficients are equal among all attributes. The core Python code snippet for CF computation is shown in Listing 2.

```
def opportunity(item) -> float:
    return (get_local_schools(item, NR_OPP)
            + get_my_timetable_avail(item, NR_OPP)
            + get_about_me(item, NR_OPP))

def ability(item) -> float:
    return (get_my_qualifications(item, NR_ABB)
            + get_about_me(item, NR_ABB)
            + get_my_docs(item, NR_ABB)
            + get_image_score(item, NR_ABB))

def willingness(item) -> float:
```

Table 3. Top 10 Providers with the Highest CF Trust and the Number of Reviews Received

Prov	CF	R	C	Prov	CF	R	C	Prov	Rand	R	C
143,427	0.436	5	3	2,273,546	0.049	5	53	3,087,320	1	0	0
2,049,616	0.411	5	7	20,688	0.032	5	46	3,215,773	1	0	0
234,392	0.405	0	0	109,854	0.064	4	44	3,092,006	1	0	0
2,875,684	0.402	5	2	2,354,656	0.179	5	41	3,222,468	1	0	0
2,152,564	0.395	4	19	1,141,164	0.085	4	41	1,472,774	1	0	0
189,455	0.372	5	18	73,286	0.152	4	41	1,193,106	1	0	0
2,739,757	0.370	0	0	103,071	0.047	4	39	1,321,398	1	0	0
682,981	0.369	0	0	70,763	0.077	4	38	3,303,579	1	0	0
3,223,827	0.367	5	4	1,862,858	0.301	4	37	2,228,202	1	0	0
2,302,155	0.364	0	0	3,146,887	0.286	5	37	3,359,468	1	0	0

Prov: Childcare Provider, CF: CF-T Trust Score, Rand: Random Trust Score, R: Rating, C: Review Count.

```
    return (get_my_timetable_avail(item, NR_WILL)
            + get_about_me(item, NR_WILL)
            + get_timetable_last_update(item, NR_WILL)
            + get_my_docs(item, NR_WILL)
            + get_image_score(item, NR_WILL))

def cf_trust(provider):
    return ( ability(provider)
            * willingness(provider)
            * opportunity(provider))

cf_ratings = [cf_trust(provider['l']) for provider in provider_list]
```

Listing 2. Degrees of Trust Calculation

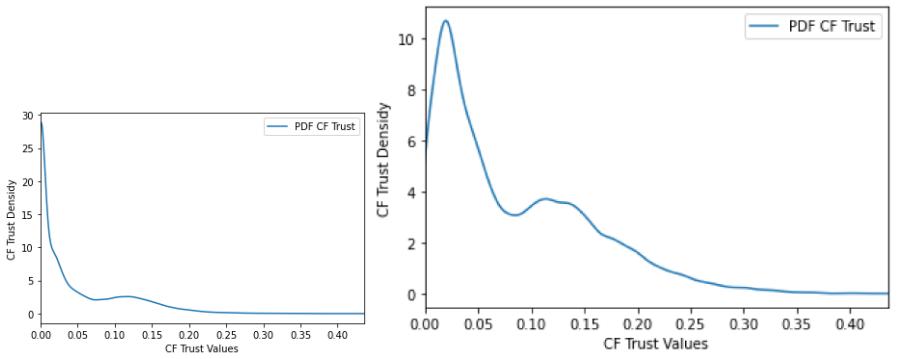
The inner functions are the normalisation calculations for each of the trust components. For example, the *get_about_me* normalisation function is shown in Listing 3, where *get_pos* returns the positive sentiment of the *about_me* text, *MAX_ABOUT_ME* is the maximum value the positive sentiment has in all the population and *contr* is the number of attributes for the current contribution.[13]

```
def get_about_me(item, contr) -> float:
    return (get_pos(item['about_me']))/(MAX_ABOUT_ME * contr)
```

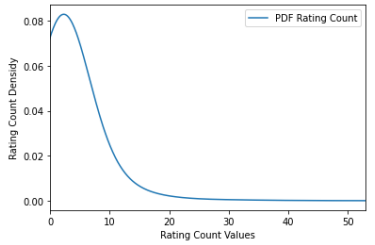
Listing 3. Normalisation function on about_me text sentiment

The left columns of Table 3 show the top 10 providers with the highest CF Trust (column 'CF') and the corresponding aggregated rating and number of reviews. The middle columns show the top 10 providers with the highest number of reviews, their ratings and the corresponding Castelfranchi–Falcone's Trust. The four columns on the right show the top 10 providers with the highest random trust (i.e., a random number between 0 and 1 assigned to each provider) and their corresponding rating and review count.

[13]The Jupyter Notebook with the complete calculations of the Castelfranchi Degrees of Trust model applied to the childcare network can be found at http://srv-us.prifti.us:28888/lab/tree/CastelfranchiTrustCalculation.ipynb



(a) PDF CF Trust (b) PDF CF Trust of providers with reviews



(c) PDF Rating Count

Fig. 9. Probability distribution functions for CF Trust, Rating and Number of reviews received.

The probability distribution functions for the three measures, CF Trust, Number of reviews and Rating, are shown in Figure 9. Along with calculating the CF Trust, we generate a random number (referred to as 'Random Trust'), which is assigned uniformly at random by using *random.uniform*[14] from the NumPy package in Python. It generates random numbers between [0, 1) with a probability density function of $p(x) = \frac{1}{b-a}$. We compare CF Trust to random trust as a performance baseline measure between the calculated trust and randomly generated numbers. As a first result, we can easily see CF Trust performs better than random.

There are over 80K childcare providers in the network, and the 10 with the highest random trust score all have a score with five 9s after the decimal point; hence, all are rounded up to 1. However, it is interesting that they have not received any reviews (i.e., they have not been trusted by parents yet to provide childcare services). The fact that by a random selection there is a higher probability of selecting childcare providers that parents have not yet trusted is supported by the fact that there are 73,182 providers that have not yet received any reviews, and only 7,143 providers have received reviews. Less than 10% of the childcare providers in the network have received any reviews. The number of reviews received is an important measure, independent of the rating the parent leaves afterwards.

The CF Trust calculation is independent of the number of ratings received or the rating values. These were intentionally not included in the degree of trust calculation. This differs from the centrality measures that mainly depend on the number of reviews received since reviews are the incoming links of the care providers' network. The CF Trust performs substantially better than a random measure (i.e., random trust). For example, the top 100 higher CF-Trust-rated providers

[14]Please see https://numpy.org/doc/stable/reference/random/generated/numpy.random.uniform.html

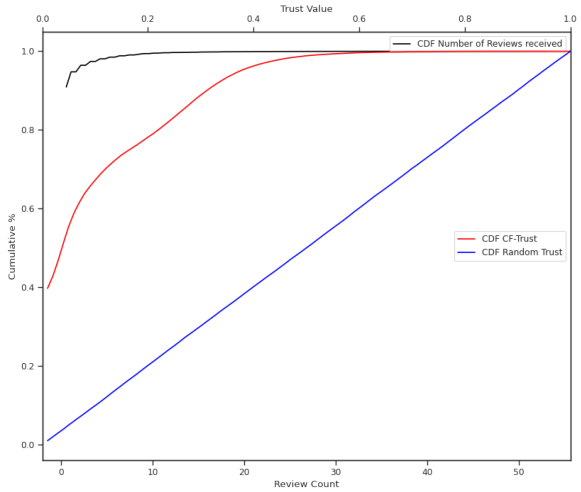


Fig. 10. CDF comparison of Random Trust, CF Trust and reviews received.

account for 330 reviews received (i.e., they have been trusted 330 times), whilst the 100 higher-rated providers with random trust account only for about 23 reviews. CF Trust performs as well as PageRank. The number of reviews received (i.e., the number of times providers have been trusted) for the highest-ranked providers is comparable between CF Trust and PageRank. Yet, the CF Trust measure is independent of the number of ratings, whilst the number of ratings received is embedded and positively influences the PageRank calculation.

More importantly, the centrality measures do not offer any information about providers that have not received any reviews since the network for calculating the centrality measure focuses on the triad *parent* $\rightarrow_1$ *review* $\rightarrow_2$ *provider*. On the other hand, CF Trust depends solely on provider attributes, and every provider participates in the trust scoring. There are four providers in the top 10 highest CF Trust providers that have not received any reviews.

4.1 Validating the Predictive Abilities of the CF Model

We have trust data represented by the number of reviews each provider has received (the more reviews a provider receives, the more times they have been trusted with care work). We also have computed CF Trust, attributed-based trust calculations that are independent of the reviews received. We want to validate if attribute-based trust is a good predictor of reviews received (potentially a good predictor for a trusted provider). KS was introduced as a test for underlying exponential distributions. To validate this correlation, we start by comparing the distribution properties of PDF and CDF. The CDF in Figure 10 shows the cumulative distribution of values between CF Trust and the Number of Reviews received. About 90% of providers have not received any reviews, whilst about 90% of providers have a CF Trust value of 0.4 or less. About 99% of providers have 30 or fewer reviews and have a CF Trust value of 0.6 or less.

As the next step, we compute Spearman's correlation coefficient to see if the two distributions are similar and correlated.

Spearman's correlation coefficient between the two distributions *'Number of Reviews Received'* and *'CF Trust'* is 0.238 with a p-value of 0. These results indicate that the two distributions are positively correlated with high confidence (the p-value is very small). In Figure 12, results along the 'y=x' diagonal are the highly correlated results, indicating providers who perform similarly with both CF Trust and the number of reviews received. Points on the lower right-hand side are

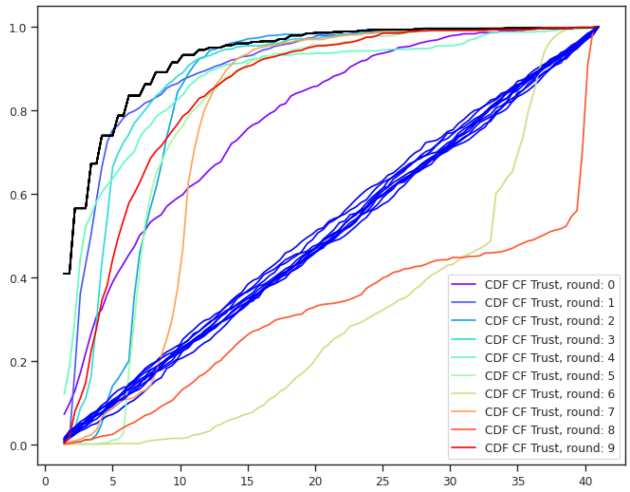


Fig. 11. CDF convergence over 10 iterations of belief calibration.

the providers where CF Trust calculation indicated low trust yet they have received many reviews and are highly trusted in the platform. These providers indicate an opportunity to improve and better calibrate the belief parameters.

Results in the scatterplot where the number of reviews received is 0 but the CF Trust is in $[0, 1)$, by far the majority of providers, indicate results where, even though we have a lack of information when trying to rank them based on number of reviews received, we can assign trust values based on attribute-based trust computation.

We have an opportunity to calibrate the belief parameters $\alpha_{i,x,y}$, $\beta_{i,x,y}$ and $\gamma_{i,x,y}$ based on observations. Ordinary Least-squared Optimisation for linear regression in Python [Seabold and Perktold 2010] was deployed to search and optimise the parameters. We used a 10% providers sample from the entire network to run the optimisation search. Next, 50% of the providers with ratings from the 10% sample were used to create the training set. We normalised the rating count between 0 and 1 and set $\mathbf{y}_{n\times 1}$ as the optimisation target. For each provider in the training set, we created an array of 13 parameters that participate in the CF Trust calculation. These were used as the training matrix of dimension $\mathbf{X}_{n\times 13}$, where n is the number of sampled providers with ratings.

After each iteration, the belief parameters (i.e., $\alpha_{i,x,y}$, $\beta_{i,x,y}$ and $\gamma_{i,x,y}$) are derived from the Ordinary Least Squares model and used in the calculation of CF Trust. Even with only 10 iterations over a 10% sample, we can visually see the convergence and get a new set of belief parameters that perform better than the belief parameters result of the normalisation process. Figure 11 shows how convergence, essentially, is embedded in the training model via the target array $\mathbf{y}_{n\times 1}$, while the rating distribution pattern isn't.

Observing Figure 11, we notice some interesting facts about the belief calibration:

— In *round 1* the CF Trust looks the closest to the observed expressed trust on the childcare.co.uk platform.
— The belief coefficients are not bound; however, the best performing results show coefficients within $[-1, 1]$.
— R^2 shows that the model only explains about 6% of variability. Independently, the adjusted belief parameters yield results that converge closer to observations.
— We started from a position of no knowledge with normalised belief coefficients and now have a set of coefficients that improve the CF Trust match observations.

Table 4. Rolling Parameter Optimisation with Increasing Sample Sizes; Each Stage Uses the Output Parameters from the Previous Step

Experiment	1	2	3	4	5
Execution time (minutes)	10	18	13	36	41
Sample rate (%)	0.2	0.4	0.6	0.8	1
Sample size (number of records)	14	28	42	57	71
Objective function value	0.936	2.893	6.010	6.802	7.748
Total drift with input parameters	45.03	12.62	18.87	24.44	30.109
Total drift with optimised parameters	5.55	12.62	18.87	24.44	30.109

Table 5. Optimising a Random Walk: The Same Sample Is Explored Five Times, and Each Time a Different, Randomised, Setup Is Used

Experiment	1	2	3	4	5
Execution time (minutes)	10	11	11	11	11
Sample rate (%)	0.2	0.2	0.2	0.2	0.2
Sample size (number of records)	14	14	14	14	14
Objective function value	1.144	0.472	0.847	0.534	1.051
Total drift with input parameters	77.94	79.70	53.36	63.07	37.49
Total drift with optimised parameters	6.612	5.845	5.93	6.512	5.831

The linear regression used is bound to have strong limits in optimising the coefficients in a non-linear distribution. For better results for coefficient optimisation, we used non-linear optimisation methods, more specifically, differential evolution[15] and its Python implementation.[16]

We ran two different experiments:

— Progressive optimisation with different sample size and
— Random walk with the same sample size.

The purpose of the first experiment is to evaluate how the parameter optimisation progresses as we run multiple experiments with increased sample size, where each stage uses the output parameters from the previous step. The results are presented in Table 4.

The results show that the first optimisation, even with a minimal sample size, improves the outcome by about 10 times. The total drift of the CF-T, calculated using random belief parameters in input, is 10 times bigger than the total drift when using the differential evolution optimised parameters. After that, however, all successive executions that use the parameters optimised in the previous step do not improve the total drift further. This is despite the increased sample size, which results in a significant increase in execution time.

In the second experiment, we ran the optimisation five times with a sample size rate of 0.2%, starting from a random starting position (i.e., using random belief parameters). We compared the results shown in Table 5. In all five experiments, across both methods, even though both the parameters and the sample data were different, the resulting optimised belief parameters converge to the same values. These results were confirmed over hundreds of executions of the experiments. Whilst the optimisation method does not guarantee that the parameters are a global minima of the

[15]Please see https://en.wikipedia.org/wiki/Differential_evolution
[16]Please see the Scipy documentation.

Table 6. Optimised Belief Parameters: We Measure How Each Component of the Profile Impacts the Three CF Dimensions

	Opportunity Belief	Ability Belief	Willingness Belief
About Me Sentiment	−3.47e-06	0.2411	−0.0083
Profile Picture Happiness Score	−1.01e-07	0.0103	−0.0031
Services Local Schools	−1.23e-07	0.0292	0.0009
Has Qualifications	7.16e-06	−0.2678	−0.0118
High Timetable Availability	−2.31e-06	0.0185	0.0218
Recent Updates Timetable Availability	9.61e-07	0.0024	0.0024
Has Uploaded Personal Documents	0.1906	−0.0358	−0.0358

Uploading personal documents greatly impacts the Opportunity Belief parameter, with a marginal degradation of the other two. No component could improve all three at once.

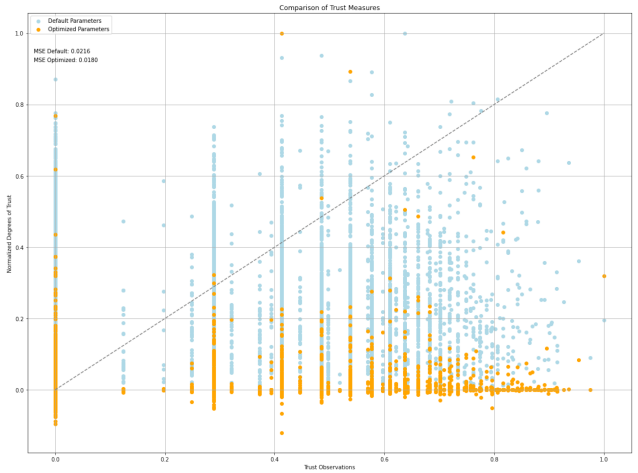


Fig. 12. Comparison of CF Trust with random and optimised parameters: optimisation (orange circles) moves data points closer to the diagonal (where correlation between CF measures and reviews is perfect). We observe a marked decrease of the overall MSE.

drift, there is high confidence they represent a solid solution. The resulting belief parameters are shown in Table 6.

Looking at it, we notice that, in general, the chosen attributes have little influence on the Opportunity belief. Providers who have uploaded personal documents are likely to be judged as having higher opportunities to do the job by a factor of 19%. The other parameters influence the opportunity belief close to 0. On the other hand, providers with a positive sentiment on their 'about me' section are likely to be judged high on their ability by a factor of 24% (i.e., the 'about me' sentiment contributes 24% to the overall judgment of ability in the trust evaluation).

The belief parameters, shown in Table 6, are the result of observation based on how the attributes contributed to trust for providers that the community has already trusted. Whilst we only have less than 10% of providers having received any feedback on the platform, we can now calculate the CF Trust values for all providers. Figure 12 plots Trust Observations and CF Trust calculated with random parameters and CF Trust calculated with the optimised parameters.

Table 7. Social Capital Score per 100K Households

Households	Region	Trust Capital	Trust Capital/Household
3,521,483	London	6,158.859684	174.893921
3,934,639	South East	1,402.256877	35.638768
2,652,878	East of England	625.644869	23.583628
2,383,082	South West	531.846623	22.317596
3,103,743	North West	639.411579	20.601306
2,530,062	West Midlands	419.348362	16.574628
2,361,567	Yorkshire and The Humber	382.643381	16.202944
1,142,240	North East	167.124897	14.631329
2,185,457	East Midlands	252.483454	11.552890
598,514	South East Wales	66.912271	11.179734
693,329	Northern Ireland	66.088602	9.532070
315,757	South West Wales	12.123391	3.839469
290,634	North Wales	5.647836	1.943281

Whilst regression models and non-linear optimisation help calibrate belief coefficients, we use factorisation machines to look at similarities among attributes and determine if trust patterns can emerge for providers with similarities.

4.2 Degrees of Trust Predictability

In the previous section, we used the mathematical representation of Castelfranchi–Falcone's Degrees of Trust to calculate its values across the childcare service provider network. We also used the existing knowledge about expressed trust on the network to calibrate the belief parameters. Finally, we used the best-performing belief parameters closely related to the number of received reviews on the network to recalculate and eventually normalise CF Trust scores between 0 and 1. This section explores whether the CF Trust scores are good predictive scores for future jobs.

In Table 7, we can see the ranking of social capital observed for every 100K households grouped by region. An interesting observation in this table is that the ranking closely matches the per capita GDP for each region.[17]

The strong correlation is confirmed by the Pearson's Correlation between CF Trust and per capita GDP, sorted by UK region: we obtained a correlation value of 0.98 and a p-value that is less than 10^{-7}. Such a small p-value is an indication we should reject the null hypothesis that the correlation is due to random noise. This is an important independent quantification and validation of the social science concepts discussed by Rousseau et al. [1998] about Social Capital, its effect on economic welfare and how trust is a composing component of Social Capital.

The experimental results can now induce some consideration. As we have seen, many approaches make use of a graph-based representation to model human communities, and in particular, community members correspond with nodes of the trust network; edges in a trust network instead specify the trust between a pair of members in the trust network, and in general, edges are directed and weighted. Network-based approaches display several strengths, and

[17]The top regions for per capita GDP are (1) London: 55,974; (2) South East: 34,516; (3) East of England: 29,176; (4) North West: 28,257; (5) South West: 28,012; (6) West Midlands: 26,281; (7) East Midlands: 25,956; (8) Yorkshire and The Humber: 25,696. Source: ons.gov.uk.

one of the strongest points of approaches based on Network Science to model trust is that the importance of a node in a network (which can be measured by one of the many measures of centrality available in the literature) measures the level of trust of the individual that that node represents. Network-based approaches, however, have several limitations: in almost all real cases, in fact, we lack a ground truth to guide us in choosing the best centrality measure. A second limitation that emerges from our study is that real social networks arise from the interaction of heterogeneous entities that can interact in various ways; moreover, available entities are often associated with some explanatory attributes. In case of OSNNs we highlighted three types of entities (parents, reviews and providers), each of which is endowed with attributes that characterise it. The generalisation of standard centrality measures such as PageRank or closeness in networks made up of heterogeneous nodes with attributes is not a trivial task. The topology of the network plays a key role in the calculation of the centrality measure, and in our scenario, OSNNs are sparse and comprise groups of nodes that are strongly interconnected. This peculiar topology may benefit some centrality measures to the detriment of others.

The CF-T model is, in our opinion, able to address the criticisms seen above. The CF model, in fact, is a socio-cognitive model that relies on several types of information, such as opportunity, ability and willingness. Rather than considering trust between two individuals as a variable to be measured or predicted, the CF model relies on a number of surrogate parameters that, taken together, provide a quantitative measure of the degree of trust of two individuals that is well aligned with the notion of trust defined in cognitive and behavioural science. This is a significant step forward because it alleviates the lack of ground truth: in fact, instead of using missing information about pairwise trust scores (which is often difficult or impossible to obtain), we use surrogate variables to quantify the degree of trust of two individuals. Moreover, high-quality surrogate variables can be obtained at a reduced cost: in our application scenario we collected abundant data on the actors involved (parents, providers and reviewers), which were subsequently analysed through sentiment and image analysis, and this enabled us to measure, with a high degree of accuracy, the value of opportunity, ability and willingness. Furthermore, the CF model is independent of the relationships between the involved actors and is based only on attributes. A further advantage of the CF-T model is that its calculation is independent of the number of reviews received by a provider and, thus, it does not suffer from the problem of missing data that might affect centrality-based approaches operating on sparse trust networks.

5 Conclusions and Future Work

This work, among others, proved the worthiness of a query-based architecture, system design and implementation for a distributed system capable of collecting large amounts of data for extended periods. This system was consistently used across the work of this research and was one of our early goals when looking to solve one of the dimensions and challenges when working with trust: trust information collection. The ability of user queries to extract data is guaranteed by two query engines: OXPath and DR-Web-Engine. Further, we found in Castelfranchi–Falcone's model a suitable theoretical model to fit to the data we collected.

There are, of course, still some limitations to our work, which we intend to explore and overcome in our future research. A remaining limitation of this study is its focus on the specific domain of childcare, with a focus on the UK. Therefore, the detailed analysis carried out here may still lack generalisability to other domains or countries. It would be interesting to examine other application areas and/or datasets and to compare the results obtained (e.g., employment of carers and nurses for disabled, elderly or dependent persons).

A second, more standard, limitation is the potential bias introduced by studying online interactions only. Families who do not use OSNNs are not covered here, and it is unclear whether they

simply don't need to, their needs being covered locally by informal channels, or because they do not trust the whole concept of seeking personal care online.

A third limitation is technical. In Castelfranchi–Falcone's work, there is often a positive correlation between the calculated CF-T value and the number of ratings received. At times, such correlation may become problematic as we may find a bias towards positive reviews, in the sense that individuals on a social platform are more likely to give positive than negative reviews. Also, the idea of counting the number of positive reviews as an indirect measure of trust still does not fully account for differences in review style; some users may like to post positive reviews, while others may remain silent.

Acknowledgments

For the purposes of open access, the authors have applied a CC-BY public copyright licence to any author accepted manuscript version arising from this submission.

References

Santa Agreste, Pasquale De Meo, Emilio Ferrara, Sebastiano Piccolo, and Alessandro Provetti. 2015. Trust networks: Topology, dynamics, and measurements. *IEEE Internet Computing* 19, 6 (2015), 26–35. https://doi.org/10.1109/MIC.2015.93

David A. Askay. 2015. Silence in the crowd: The spiral of silence contributing to the positive bias of opinions in an online review system. *New Media & Society* 17, 11 (2015), 1811–1829.

Guido Barbian. 2011. Trust centrality in online social networks. In *2011 European Intelligence and Security Informatics Conference*. IEEE, 372–377.

Pierre Bourdieu. 1980. Le capital social: Notes provisoires. *Actes de la Recherche en Sciences Sociales* 31, 1 (1980), 2–3.

Cristiano Castelfranchi and Rino Falcone. 2000. Trust is much more than subjective probability: Mental components and sources of trust. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. IEEE, 10 pp.

Cristiano Castelfranchi and Rino Falcone. 2010. *Trust Theory: A Socio-cognitive and Computational Model*. Vol. 18. J. Wiley & Sons.

Cristiano Castelfranchi, Rino Falcone, and Emiliano Lorini. 2009. A non-reductionist approach to trust. In *Computing with Social Trust*, Jennifer Golbeck (Ed.). Springer London, London, 45–72. https://doi.org/10.1007/978-1-84800-356-9_3

Aaron Clauset, Mark E. J. Newman, and Cristopher Moore. 2004. Finding community structure in very large networks. *Physical Review E* 70, 6 (2004), 066111.

Rino Falcone and Cristiano Castelfranchi. 2001. Social trust: A cognitive approach. In *Trust and Deception in Virtual Societies*. Springer, 55–90.

Jennifer Golbeck. 2005. *Computing and Applying Trust in Web-based Social Networks*. University of Maryland, College Park.

Jennifer Golbeck and Jim Hendler. 2006. Inferring binary trust relationships in Web-based social networks. *ACM Transactions on Internet Technology* 6, 4 (2006), 497–529.

Trudy Govier. 1997. *Social Trust and Human Communities*. McGill-Queen's Press-MQUP. 31 pages.

Guibing Guo, Jie Zhang, and Neil Yorke-Smith. 2015. TrustSVD: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence,*. AAAI Press, 123–129.

Mohsen Jamali and Martin Ester. 2010. A matrix factorization technique with trust propagation for recommendation in social networks. In *Proceedings of the 2010 ACM Conference on Recommender Systems, (RecSys'10)*. ACM, 135–142.

Ling jao Chen and Jian Gao. 2018. A trust-based recommendation method using network diffusion processes. *Physica A: Statistical Mechanics and its Applications* 506 (2018), 679–691.

Navid Khaledian and Farhad Mardukhi. 2022. CFMT: A collaborative filtering approach based on the nonnegative matrix factorization technique and trust relationships. *Journal of Ambient Intelligence and Humanized Computation* 13, 5 (2022), 2667–2683.

Navid Khaledian, Amid Nazari, Keyhan Khamforoosh, Laith Abualigah, and Danial Javaheri. 2023. TrustDL: Use of trust-based dictionary learning to facilitate recommendation in social networks. *Expert Systems with Applications* 228 (2023), 120487.

Robert Leonardi, Raffaella Y. Nanetti, and Robert D. Putnam. 2001. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton, NJ.

Edward Loper and Steven Bird. 2002. NLTK: The natural language toolkit. *arXiv preprint cs/0205028* (2002).

Stephen Marsh and Pamela Briggs. 2009. Examining trust, forgiveness and regret as computational concepts. In *Computing with Social Trust, Human-Computer Interaction Series*. 9. https://doi.org/10.1007/978-1-84800-356-9_2

Stephen Paul Marsh. 1994. *Formalising Trust as a Computational Concept.* Technical Report. University of Stirling.

Justin J. Miller. 2013. Graph database applications and concepts with Neo4j. In *Proceedings of the Southern Association for Information Systems Conference*, Vol. 2324.

Alejandro Portes. 1998. Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology* 24, 1 (1998), 1–24.

Ylli Prifti. 2023. *The Emergence of Interpersonal and Social Trust in Online Interactions.* Ph.D. Dissertation.

M.-Ross Quillian. 1966. *Semantic Memory.* Air Force Cambridge Research Laboratories, Office of Aerospace Research.

Denise M. Rousseau, Sim B. Sitkin, Ronald S. Burt, and Colin Camerer. 1998. Not so different after all: A cross-discipline view of trust. *Academy of Management Review* 23, 3 (1998), 393–404.

Skipper Seabold and Josef Perktold. 2010. Statsmodels: Econometric and statistical modeling with Python. In *9th Python in Science Conference.*

Andrew Jon Sellers, Tim Furche, Georg Gottlob, Giovanni Grasso, and Christian Schallhart. 2011. OXPath. In *Proceedings of the 20th International Conference Companion on World Wide Web (WWW '11).* ACM Press. https://doi.org/10.1145/1963192.1963304

Sefik Ilkin Serengil and Alper Ozpinar. 2021. HyperExtended LightFace: A facial attribute analysis framework. In *2021 International Conference on Engineering and Emerging Technologies (ICEET'21).* IEEE, 1–4. https://doi.org/10.1109/ICEET53442.2021.9659697

Cai-Nicolas Ziegler. 2009. On propagating interpersonal trust in social networks. In *Computing with Social Trust*, Jennifer Golbeck (Ed.). Springer London, London, 133–168. https://doi.org/10.1007/978-1-84800-356-9{_}6

Cai-Nicolas Ziegler and Jennifer Golbeck. 2015. Models for trust inference in social networks. In *Propagation Phenomena in Real World Networks.* Intelligent Systems Reference Library, Vol. 85. Springer, 53–89.

Received 31 January 2024; revised 5 December 2024; accepted 7 January 2025