INTERMEDIATE SEARCH ELEMENTS AND METHOD COMBINATION IN LEAD-USER SEARCHES

SAMPSA HYYSALO

Department of Design & Department of Management Studies
Aalto University, P. O. Box 31000, FI-00076 Aalto, Finland
sampsa.hyysalo@aalto.fi

PIA HELMINEN[*]

Department of Industrial Engineering and Management
Aalto University, P. O. Box 17700, FI-00076 Aalto, Finland
pia.helminen@aalto.fi

SAMULI MÄKINEN

Department of Engineering Design and Production
Aalto University, P. O. Box 17700, FI-00076 Aalto, Finland
samuli.makinen@aalto.fi

MIKAEL JOHNSON

Helsinki Institute for Information Technology HIIT
Aalto University, P. O. Box 15600, FI-00076 Aalto, Finland
mikael.johnson@aalto.fi

JOUNI K. JUNTUNEN[†] and STEPHANIE FREEMAN[‡]

Department of Management Studies, Aalto University
P. O. Box 21210, FI-00076 Aalto, Finland
[†]*jouni.juntunen@aalto.fi*
[‡]*camisteph@gmail.com*

Published 30 September 2014

Users play an increasingly important role in product and service innovation. Finding the right users can require substantial search effort. Network searches are increasingly popular in searching for rare lead users. In these searches, implicit and inexact referrals have been

---

[*]Corresponding author.

found to comprise a substantial number of network referrals; numbers as high as 70% of the most important referrals to sought people have been reported. To aid handling such referrals during network searches, we explicate their status as intermediate referral types, and how these referral types relate to known search methods. The constraints set by intermediate referrals could potentially be overcome and their potential be capitalized through more extensive method combination in network searches than has been trialed to date. We proceed to offer a proof of concept for such searches through documenting how we ran them in four realworld searches and chart future research avenues.

Keywords: Lead users; user innovation; network search; pyramiding; intermediate referrals; implicit referrals; inexact referrals; combinatory search; rare research subjects.

Introduction

Users play an increasingly important role in product and service innovation. Finding the right users, often those people who know the most about the solution area, can require substantial search effort. To aid these search processes, we clarify the referral types involved in these searches, and the requirements they set for methods used in searching for lead users and other similarly "hard to find research subjects."[1] This also opens opportunities for new types of searches, which we demonstrate to invoke further research on method combinations in such hard to find searches.

Numerous studies have concentrated on identification of user or customer needs and how these needs can be incorporated into products and services (e.g., Hauser and Clausing, 1988; Kaulio, 1998; Pals *et al.*, 2008). Users can also be a source of new product ideas (e.g., von Hippel, 1986, 2005; Hannukainen and Hölttä-Otto, 2006; Jeppesen and Frederiksen, 2006; Raasch *et al.*, 2008; Hyysalo *et al.*, 2013b) and cooperation with lead users has been shown to be an effective means to gain insight into the trends and solutions available in the user domain and to further transform this knowledge into product and service concepts (e.g., Herstatt and von Hippel, 1992; Lilien *et al.*, 2002; Churchill *et al.*, 2009).

However, one of the main questions remains how to find the right people (e.g., Olson and Bakke, 2001; Churchill *et al.*, 2009). Bilgram *et al.* (2008, p 421) conclude that "research in this area indicates that companies are still facing considerable problems in efficiently identifying suitable users." Finding subjects with rare attributes within poorly mapped search spaces also remains a more general problem in social sciences writ large (Sudman and Kalton, 1986; Atkinson and Flint, 2001; von Hippel *et al.*, 2009). User innovation research has contributed to

---

[1]Lead users are users who face needs before the majority of the market and benefit significantly from obtaining solutions to those needs (von Hippel, 1986, 1988).

the range of available search methods (e.g., von Hippel *et al.*, 2009; Bilgram *et al.*, 2008), and has sought to rationalize these search processes in terms of presenting process depictions (e.g., Churchill *et al.*, 2009), formalized some of the strategies developed in doing these searches (von Hippel *et al.*, 2009), as well as pursued comparisons and simulations for establishing the efficiency and efficacy between different lead-user identification methods (Poetz and Prügl, 2010; Stockstrom *et al.*, 2012; von Hippel *et al.*, 2009).

Much of this work has centred around networking strategies in rare subject searches, and to date proceeded by examining how individuals are linked to one another, in affinity to social network analysis (Newman, 2003). The rare subject networking searches, however, include frequent and important episodes where individuals are not linked directly to other people. Poetz and Prügl (2010, p. 906) report only 30.7% of referrals from an initial search domain to another domain as being linked to concrete people, 28.2% pointing to organizations or institutions, and 41.1% pointing to events, professions, products, literature, or technologies. Thus, the referrals to other entities than people amounted to 69.3% of the referrals to the knowledge that was in their line of argumentation potentially most vital. The issue is not limited to analogous fields: referrals to organisations, events, indexes, mass media, and computer-mediated communication (CMC), rather than to concrete people, are common in network searches also within a search domain. To the best of our knowledge, research to date has not addressed such referrals apart from naming them in varying conventions as "implicit" or "less detailed," even though, as the above suggests, these referrals may have considerable importance in rare subject searches. Poetz and Prügl (2010, p. 904) recognize the different amounts of work required to get to differently specific referrals, and we will below explain why and how this is.

A key characteristic of referrals in network search is that they take place in a directed dynamic network (such as pyramid or snowball sampling) where a node (or vertex) belongs to an in-component and out-component (Newman, 2003), where the information on the out-component emerges step by step. An interviewed person can provide the information on the neighboring out-component nodes. Other types of referrals point to nodes — referents — that may be conductive for finding the next person, but cannot provide the subsequent referral information.[2]

---

[2]Note on terminology. We use referral(s) in accordance with all pyramid search literature to denote the information on neighboring nodes of the out-component of directed network or link-tracing search at a given search step node (von Hippel *et al.*, 2009; Stockstrom *et al.*, 2012; Poetz and Prügl, 2010). In marketing, referrals are often called leads, which may better communicate to some readers than the more formal terms. "Referent" is the term we reserve for the next node from a referral provided by a previous network node. The search "step" that connects these two (von Hippel *et al.*, 2009), i.e., the "edge" (Newman, 2003) we call "application of search method" to denote that the step needs to be accomplished by some means.

A referral to an organisation points to an entity, which cannot provide the referrals to the next node like a concrete person can, for instance, through a lead-user survey. At least not before some person within the organisation is reached to give the referral required for continuing the search. Similarly, a referral to a user solution on the Internet, an increasingly common episode in lead-user searches today, points to a node, which lacks the ability to provide information on the out-component needed to carry on with the search. In closer scrutiny the "implicit" or "less detailed" referrals are better understood as forming a larger family of *intermediate referrals* to referents, some of which do not provide information on the out-component at all, some may provide it through an additional step, and some provide it differently than a person would in a traditional respondent-assisted network search. These intermediate referrals can and are used in rare subject searches, but they can pose delay, uncertainty and shifts in the sampling logic that are better explicated than left to practitioners to implicitly grapple with.

This explication can be done by distinguishing between referral types, method types, and their characteristics, which we together define as *intermediate search elements*. The intermediate search elements do not only present hindrances to networking searches but also open possibilities in purposefully combining rare subject search methods. This is an area where trials have emerged using two search methods in parallel (Hienerth *et al.*, 2007) and sequentially (Keinz and Prügl, 2010, p. 280), and where the better understanding of intermediate search elements can aide further combinatory method development.

Our contribution in the current paper is twofold:

(1) We conceptually clarify what are the intermediate elements in networking searches and the effects they have on known rare subject search methods.
(2) We present a proof of concept for purposefully combining multiple search methods to overcome search method requirement constraints, by elaborating an approach that uses multiple methods both in parallel and sequentially, and its application in four real world cases.

We proceed by next recounting the established methods for the finding of lead users. We then clarify the intermediate search elements and the restrictions and opportunities these open for different lead-user search methods. Thereafter, we introduce the idea of purposefully combining different search methods and one way of pursuing it, which we call "mountaineering." We then move to describe the results and the conduct of four principal and two supportive case searches in the fields of online teaching and learning and in renewable energy technologies. The graphical over time search process depictions are available as animations at http://sn.im/mountaineering. We find that these mountaineering searches did achieve their

targets, hence offering a proof of concept that lead-user searches can also be conducted by combining the search methods. We end by discussing three avenues of further research on intermediate search elements and combinatory searches.

Approaches to Rare Subject Identification

Literature to date has suggested several different methods and directions as a means to finding rare subjects for research and R&D. These need to be first recounted before we can turn to discussing how they relate to intermediate search elements and what difference these may make to these methods.

Review of the literature on lead-user identification

Screening is a common approach for finding lead users (von Hippel *et al.*, 2009; Belz and Baumbach, 2010). It is based on collecting information from every member of a population in order to identify the members with desired attributes. However, the rare nature of the sought lead-user attributes can make screening inefficient (Sudman, 1985). For example, Lüthje (2000) reports screening 2,043 persons to identify 22 lead users — a sampling efficiency of only 1.1%.

Snowball sampling (Welch, 1975; Goodman, 1961) or the "*telephone networking approach*" as labelled in the Lead User Project Handbook (Churchill *et al.*, 2009), means that individuals are asked to identify people who have a desired characteristic, or who can provide important information. The Lead User Project Handbook also suggests *site visits* when initial telephone interviews have revealed interesting user-developed innovations (Churchill *et al.*, 2009).

Pyramid sampling (i.e., *pyramiding*) is a variant of snowball sampling; asking for nominations of individuals who know more or have more of the sought attribute (von Hippel *et al.*, 2009; Lilien *et al.*, 2002; von Hippel *et al.*, 1999). It has been found to be more efficient than snowball sampling. von Hippel *et al.* (2009) have tested the efficiency[3] of pyramiding compared to screening, and in their study of 663 pyramiding search chains found the effort of pyramiding search to be only 28.4% of the effort of screening. Stockstrom *et al.* (2012) analysed simulations of a total of 13,188 search chains and found pyramiding to require, on average, 31% of the effort[4] of screening.

---

[3]"Efficiency" here means the chain length, i.e., "number of nodes from start to end point" (von Hippel *et al.*, 2009, p 1401).

[4]"Effort" is used in parallel with 'efficiency', i.e. "number of chain links" (Stockstrom *et al.*, 2012, p. 21).

Investigation of analogous fields is a name lead-user researchers have given to exploring fields in which similar challenges are present as in the search field under consideration (Lüthje and Herstatt, 2004). Lead users identified in the "advanced analog" fields are found to develop innovations that are most radical relative to conventional thinking (von Hippel, 2005). A well-known example of this is the case of 3M trying to develop surgical drapes (the material that prevents infections from spreading during surgery). The most valuable users were found in veterinary hospitals and among make-up artists in Hollywood (von Hippel *et al.*, 1999). The cross-industry innovation case of the anti-lock breaking system (ABS) transferring from the field of aerospace to standard cars is a further example (von Hippel *et al.*, 1999; von Hippel, 2005). Poetz and Prügl (2010) addressed the potential of pyramiding for crossing domain-specific boundaries by analysing 1,147 interviews conducted in the course of pyramiding search processes in eight lead-user studies. In their study more than one-third of those interviewees who were able to provide a valid referral in their interview, could refer to one or more analogous domains previously unknown to the searcher.

Domain experts (sometimes called lead-user experts) are people who are highly knowledgeable of the user domain area but not necessarily the lead users (or other rare subjects) sought for. Domain experts can be asked to point out lead users (Churchill *et al.*, 2009).

User communities have in some studies been used for finding prominent lead users; e.g., in mountain biking and kayaking (Lüthje *et al.*, 2005; Hienerth, 2006).

Broadcasting means advertising the need for a solution or expertise in hope that relevant people self-select to respond (Lakhani, 2006; Jeppesen and Lakhani, 2010). A common form of broadcasting is to post a problem on an Internet discussion forum or a mailing list of a special interest group. Broadcasting has been combined with pyramiding in several lead-user projects (e.g., Hienerth *et al.*, 2007).

Idea competitions follow the idea of broadcasting. Submissions to the contest are evaluated by an expert panel and users whose submissions score highest receive an award from the manufacturer (which is often granted the right to exploit the solution in its domain). Piller and Walcher (2006) claim that idea competitions are often faster and less laborious (and expensive) compared with screening and pyramiding.

A *virtual stock market* (*VSM*) means bringing a group of participants together via the Internet and allowing them to trade shares of virtual stocks. Spann *et al.* (2009) explored the use of VSMs in identifying lead users in the product category "movies." They concluded that VSMs are an effective means to attract and filter large numbers of anonymous customers for the identification of lead users on the Internet.

Seeking out innovative solutions to reveal innovating users behind them is another way to find lead users as they are likely to be more invested in such

development than other users (Bilgram *et al.*, 2008). Many lead users have developed prototypes, modifications, or other iterations of existing products to meet their needs, which the products on the market do not yet satisfy (von Hippel, 1976, 1988; Baldwin *et al.*, 2006).

Doing "*netnography*" in user forums is a recently established method for analysing online communities (Kozinets, 1998, 2010). It was applied by Belz and Baumbach (2010) to identify lead users, who might be actively blogging or reading and commenting on blogs, such as those relating to new technology (Droge *et al.*, 2010). Bilgram *et al.* (2008), similarly hypothesised that in web 2.0 "leading edge users are likely to be already committed to communities as active members."

Intermediate search elements in lead-user identification methods

In the lead-user literature all of the above ways to approach the sought-after lead users have been denoted as "methods" (albeit researchers undoubtedly debate their status as such). This conflates differences between referral types, search methods, and their characteristics.

Let us begin with clarifying the character of referrals in the process of snowball or pyramid sampling. The literature on lead-user searches has operated on individuals giving referrals only to other individuals, but in real-world searches individuals often refer to companies, solutions, or different fields of expertise, for example. Whilst undoubtedly very handy ways to get to lead users, analogous fields and user communities are strictly speaking not methods, but referrals to intermediate referral types that can aid in the networking search.

To clarify the matter further, let us examine it analytically. When a lead-user networking search deals with only people, each referral points in fact to a union of a person and their "attribute," usually "lead-user characteristic" pertaining to solution information. Furthermore, as the search is a dynamic directed network (Newman, 2003) each node is arrived at on the basis of the information of the neighboring in-component nodes and must provide information on its neighboring out-component nodes for the search to continue. A person is an extraordinary node in that the person, attribute, and the information on the in-neighbors and out-neighbors can all be fused into one, and hence a search that can establish a contact to a person has high hopes of attaining both the attribute information, by, for instance, a self-assessment survey or even a display of a solution a user has made, as well as a set of further referrals to more knowledgeable people (Fig. 1). Because of this union, we call referrals to people *immediate referrals* in that all of the sought information is immediately at hand at successful contact.

The family of intermediate referrals enters the picture when the referral is not a concrete person but an entity that is associated with him. What we call
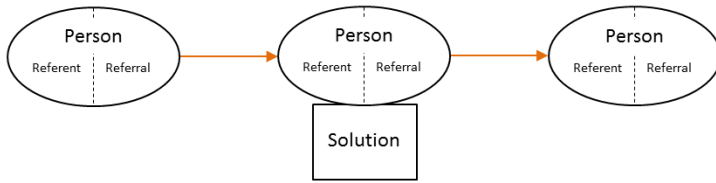
Fig. 1. A search chain, where a referral to a person is made. In the middle, the union of the person and their "attribute," for instance a solution they have made and knowledge they have of it. Each person is a referent and can provide the referrals needed to carry on with the networking search.
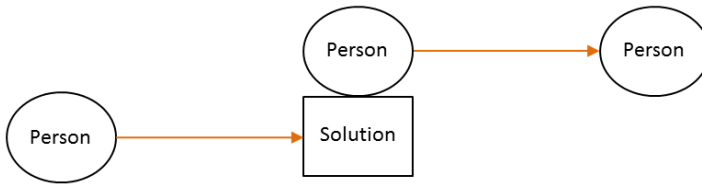

Fig. 2. A search chain, where a semi-immediate referral to a solution is made. To continue the search, the person behind the solution needs to be contacted.

semi-immediate referrals point to the "other side" of person–solution knowledge dyad, the *solution* (Fig. 2). This can give great evidence of the status of the person with regard to the sought attribute, by, for instance, showing how innovative is the solution. The solution, however, does not have a similar out-component as a person does and it cannot nominate further referrals. Certainly content analysis can reveal out-neighbors to other solutions, or even other people, and many solutions have inbuilt information about their maker, but neither is necessarily the case.[5] Contact to a person is needed to gain the same kind of out-neighbor information as in the person-to-person networking. An example is seeing an innovative solar panel system on the roof of a house, whose owners are not in, and having to gain contact with them by some other means before the search can be continued.

A common variant of semi-immediate referrals exists in *CMC*. For instance, in Internet forums both the solution and its maker are often displayed in the same post (Hyysalo *et al.*, 2013b), but in fact what is displayed is a solution description and the user's pseudonym (nickname). Whilst the interactivity of CMC can be used to address the person behind, they may not respond or respond with lag if they visit the CMC site infrequently (Fig. 3).

---

[5]Equally important is that gaining information on out-neighbors through content analysis is different with regard to sampling logic, which we discuss further.
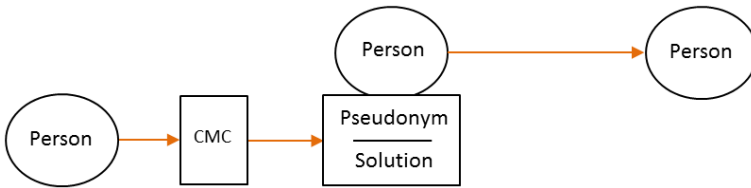
Fig. 3. A search chain where an on-line discussion forum reveals a user solution posted by a pseudonym. The forum can be used to send messages to person behind it but only asynchronously.
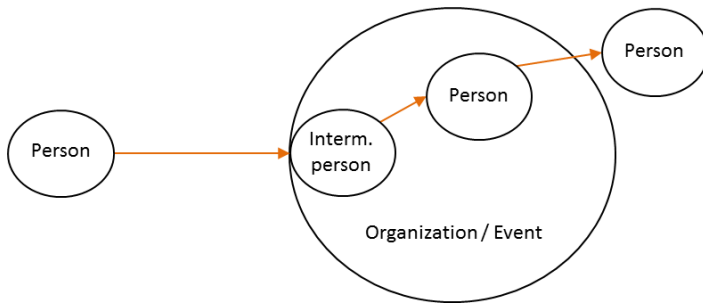

Fig. 4. A search chain, where a semi-intermediate referral to an area, i.e., an organisation or event is made. Further networking within the area, via e.g., a host, is required to be able to continue the search.

Semi-intermediate referrals no longer refer to the referent, but to an "area" in which the referent is to be found, and the area itself is of the type, which provides an association point that can be used to continue networking towards the referent. An *organisation* and *event* are common referrals that tend to have contact information or hosts that can be contacted with regard to members or participants respectively (Fig. 4).

Pure intermediate referrals do not point to the referent, nor do they provide a contact point that can give a referral. *A location*, *a profession*, *a technology*, or *an analogous field* is a referral that can be very useful, but must be sampled by other means than respondent-assisted in order to get to the next node. For instance "the aviation field" can be searched through indexes or by broadcasts or by conducting purposive sampling, but there is no direct or indirect node that could be pursued (Fig. 5).

Mass media as one-way (and often second-party-edited) communication and *indexes* such as phonebook or search engine searches, offer referrals whose character can vary from semi-immediate, such as an interview story on how "A fireman Pasi Sillanpää invents a pellet burning cradle," to pure intermediate, such as a search engine created list of veterinary doctors. They can never, however, be
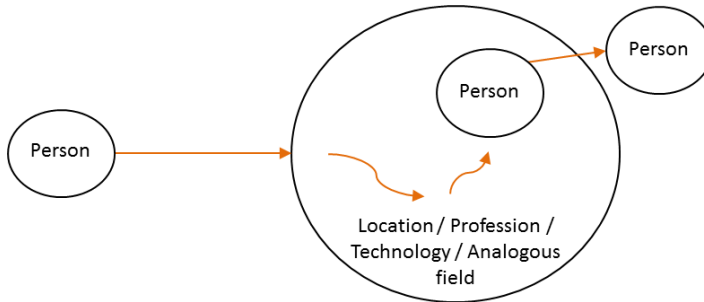
Fig. 5. A search chain, where a person cannot provide a referent or a contact point, but a pure intermediate referral, e.g., an analogous field. To continue the search, the field requires sampling for identifying the next contact point(s).

immediate referrals that can be asked respondent-assisted questions. We have placed them under pure intermediate referrals for the reason that both indexes and mass media typically require the researcher to draw on further sampling rather than provide a clear referral which to pursue.

In Table 1 we summarise the referral types thus far known to us having surfaced in lead-user searchers and the distinct ways of engagement both for researchers and potential lead users. A formal organisation has a contact person, one can participate in an event, go to a location, read media, register on an online discussion forum, demonstrate or try out a user solution, grab a reference book for a professional field, search an index, etc. Interactive computer media is distinguished from mass media because of its different communicational form; many-to-many versus one-to-many communication (McQuail, 1987), degree of editing involved prior to publication, and different content search mechanisms. In addition, the researcher sometimes can draw networking starting points from their own membership groups, i.e., personal network, but since these are only starting points, they are not included in this table. Furthermore, some empirically encountered referrals feature a combination of referral types, particularly in on-line services. For instance, a referral to a Facebook group would get one to a site that features properties of interactive computer media (peer discussions), mass media (edited pages with mostly one-way communication) and indexes (searchable parts). In the same fashion, popular services that previously could be characterised as mass media have extended their services with indexes and two-way communication sections (CMCs). We will also consider the case of network searches that do not proceed via people, but directly from solution to another, but an excursion to sampling methods is needed to fathom out what is involved.

The referral types hold implications for rare subject search methods. What we currently have amidst "rare subject search methods" is a mix of different contact

Table 1. Different referral types and key requirements and possibilities associated to them.

Referral category	Referral types
Immediate referrals	**Person**
	An individual with a name
Semi-immediate referrals	**CMC**
These referrals point to referents directly, but respondent-assistance is not immediate.	Interactive computer media (e.g., blog, forum, wiki, mailing list, social networking site and online community). Two-way communication.
	Solution
	User innovations, prototypes, etc. displayed without their maker.
Semi-intermediate referrals	**Organisation**
These referrals have a responsible gatekeeper that can assist the researcher.	Formal organisation (e.g., company, agency, non-profit, school).
	Event
	Conference, seminar, fair, etc.
Pure intermediate referrals	**Mass media**
These referrals cannot be used directly: A researcher-driven sample must be obtained next.	Mass-broadcasted one-way communication (newspaper, TV, radio and company website).
	Index
	Searchable index of things, people and their personal information[a] (e.g., census, health care, and tax records, databases, search engines).
	Field
	Professional field or domain[b] (e.g., superconductors, banking, public health care).
	Location
	A meeting place where people hang out, a subway station, a gallery, etc.

[a]Poetz and Prügl (2010) mention also "literature" that we would place under Index in network search terms.
[b]In Poetz and Prügl (2010) also "profession" and "technology."

and sampling methods. Broadcasting, idea competitions, and virtual stock markets are foremost methods of sampling in that they can be used for gathering a self-nominated sample that can later be investigated by another method; for instance by screening or by networking within or onwards from the sample by, for example, pyramiding.

Apart from lead-user identification, rare subject search methods have been discussed elsewhere in the social sciences, and this larger body of work provides important cues for how to bring clarity to the above classification problems. We

discuss this through moving from different logics of sampling to different referral types and requirements they set for search methods.

A common way to cluster methods is around whether their sampling logic is *respondent-assisted* or *researcher-driven* (e.g., Daniel, 2012). Let us first take a look at methods that rely on respondent-assisted sampling. We distinguish between two cases: (1) interviewed persons nominate other individuals (or other type of referrals) who could be asked to give more information, e.g., snowball and pyramid sampling, and (2) self-nomination to an action started by the researcher, that is, people answering a post, sign, or other broadcast of a message, e.g., broadcasting. We contrast these respondent-assisted sampling logics with (3) researcher-driven sampling: purposive, quota, and probabilistic sampling, including hybrids such as netnography, and (4) saturation sampling.

Saturation sampling, where all members of a population or a sample are surveyed, has been used in lead-user searches under the term "screening" (von Hippel *et al.*, 2009) Screening has been carried out in two distinct ways. First, which we call screening by survey, has meant going through a survey with every member in a community or a sample, or every person known to have innovated within an area (e.g., von Hippel *et al.*, 2009; Lüthje *et al.*, 2005; Franke and Shah, 2003). The other type of screening, which we call screening by content analysis, is going through every member of a population (or a sample) by examining the content of their innovations. This is found, for instance, in studies where all innovations within a field have been first collected and then their proportions of user-made and manufacturer-made have been decided based on the information available (von Hippel, 1976; Hienerth *et al.*, 2013).

In addition to these, we should keep in mind the case where a researcher is not intentionally looking for referrals, but stumbles upon a referral while doing something else. This is one form of availability sampling that we here call unintentional.

Table 2 summarises the methods and their sampling logic, the characteristics of these methods, and provides references to literature.

The excursion to different sampling logics in different search methods is needed to understand how they differ with regard to intermediate referral types, the topic to which we move now. Table 3 outlines how different referral types relate to key methods used in rare-subject searches. The table indicates a few things. First, intermediate referral types open possibilities for nominated sampling, i.e., snowball and pyramid sampling, but also pose additional steps and may cause waste of time and factually require the use of researcher-driven sampling steps when it comes to pure intermediate referrals. Self-nominated sampling methods, i.e., broadcasting, can be used in conjunction with most referral types apart from the point that solutions cannot self-nominate themselves. Self-nominated sample

Table 2. Methods and their sampling logic.

Method	Sampling logic	Characteristics	Reference literature
Screening	No sampling logic, census or saturation sampling.	Not limited to individuals but could mean inspecting all members of any group, be it cars on a parking lot, posts on an Internet forum, or participants of an event. The screening effort varies substantially depending on the size and nature of population. Often wise to first sample.	Also known as "complete enumeration" (Gobo, 2004), "saturation sampling" (Spreen, 1992).
Snowball and Pyramid Sampling	Respondent-assisted: Nominated sampling, link-tracing sampling.	Most applicable in searches where characteristics are of a sensitive nature or when there are no available statistics — that is, the size of the population is unknown (Atkinson and Flint, 2001). Allows refining the questions during the sequential search, as the searchers learn more of the search domain and population at hand. Requires a person as a point of contact.	Can be placed within the wider set of link-tracing methodologies[a] (Spreen, 1992, p 42). Has been used in studies of "hard-to-reach" or "hidden" populations such as prostitutes (McNamara, 1994; Faugier, 1995), drug users (Avico *et al.*, 1988; Griffiths *et al.*, 1993; Kaplan *et al.*, 1987), pickpockets (Inciardi, 1977), and AIDS sufferers (Pollak and Schiltz, 1988).
Broadcasting, idea competitions, virtual stock markets	Respondent-assisted: Self-nomination to research action.	Relies on the active role of participants, who self-nominate to participate. Since the researcher is not in control of the response time, broadcasting early in the search process is advised. A common form of broadcasting is to post a problem on an Internet discussion forum or	Rooted in the view that innovations or drivers of change and progress often come from outside or from the margins of established research communities (Chubin, 1976; Crane, 1969; Edge and Mulkay, 1974) so in order to increase the probability of a

Table 2. (*Continued*)

Method	Sampling logic	Characteristics	Reference literature
		a mailing list of a special interest group, although creativity in choosing channels may point to e.g., technology blogs (Droge et al., 2010), or a paid ad in a relevant magazine or social media, such as Facebook.	successful response, problems need to be broadcasted to a heterogeneous set of solvers not necessarily associated with the problem holders or their scientific and technical domains (Lakhani, 2006).
Purposive, quota, and probabilistic sampling	Researcher-driven: Researcher decides sampling logic.	Instead of screening a discussion forum of perhaps thousands of posts, one can sample out, for instance, members with the most posts, the longest discussion chains, or the most popular posts, or use random sampling for yielding a lower and screenable amount of posts. Often a practical starting point for familiarising oneself with the field where rare subjects are to be searched involves systematic and/or opportunistic sampling of various media sources such as journals or newspapers. It can, for example, take a form of an extensive literature review, utilising online search engines, or following feeds from social media applications. Retrieving information from databases and using online search engines such as Google also belong under the umbrella of sampling.	Numerous sampling strategies available in the literature: A sample of extreme situations in order to maximise variation, quota sampling for objects that contain a wide range of statuses, emblematic sampling (Gobo, 2004), location sampling (e.g., Kalton, 2009; Sudman, 1980), random sampling (e.g., Lohr, 2009), and samples targeting typical or critical cases (e.g., Patton, 2002), or anomalies. A combination of sampling methods has been used in for example the "netnographic" study of Belz and Baumbach (2010).

Table 2. (*Continued*)

Method	Sampling logic	Characteristics	Reference literature
Miscellaneous Encounters	Unintentional availability sampling: Referrals come to the researcher by chance.	Miscellaneous and informal encounters, such as casual coffee table discussions, can provide referrals for more systematic searches.	
		These encounters refer to situations where the researcher has found a referral by being in the right place in the right time, but the main activity has been something other than systematic searching.	
		One can also spot a relevant referral for a rare subject while browsing through an online newspaper, for example.	
		It is important to stay alert for possible new referrals throughout the search process.	

[a]Link-tracing methodologies presume the existence of some kind of "linkage" or "bond" with other people in the sample population (Spreen, 1992, p. 35).

Table 3. Referral types and applicable methods.

Referral type		Respondent-assisted sampling methods nominated sampling		Complete enumeration methods		Researcher-driven sampling methods	Respondent-assisted sampling methods self-nomination to research action
		Pyramid search	Snowball search	Screening by survey	Screening by solution content analysis	Purposive, quota, and probabilistic sampling	Broadcasting, idea competitions, virtual stock markets
Immediate	Person	Yes		Yes	No	Yes	Yes
Semi-immediate	CMC	Yes, delay and uncertainty in response[a]		No, response rate is unlikely to produce screening	Yes	Yes	Yes
	Solution	Yes, delay and uncertainty in response		No	Yes	Yes	No, requires link to person.
Semi-intermediate	Organisation	Additional step		Yes	No	Yes	Yes
	Event	Additional step		Yes	Yes, if solutions on display	Yes	Yes
Pure intermediate	Mass media	May require research-driven sampling step(s)		No	Yes, if solutions on display	Yes	Yes
	Index	May require research-driven sampling step(s)		No	Yes, if solutions on display	Yes	Yes
	Field	Requires research-driven sampling step(s)		Yes	Yes, if solutions on display	Yes	Yes
	Location	Requires research-driven sampling step(s)		Yes	Yes, if solutions on display	Yes	Yes

[a]As discussed in Intermediate search elements in lead-user identification methods, solution and CMC can have an immediate out-neighbor in the person involved, but in some cases this connection remains unaccomplished or accomplished only with considerable delay.

as a search strategy, however, is not applicable or sufficient in all rare subject searches even if it may provide additional referrals in most searches. Complete enumeration, i.e., screening, either by survey or solution content analysis, can be targeted to referents "within an area" (intermediate referrals). However, without prior sampling such approach can be extremely time consuming and costly. For instance, the full coverage of 220,000 messages in CMC in one renewable energy Internet forum we discuss below would have presented an unfeasible workload for most searches. It took us three person months to screen by document analysis just the selected key sections. Finally, even as intermediate referrals cannot be directly used for nominated sampling, they are amenable for researcher-driven sampling; for instance, using a solution as information of what kind of solutions one could search for in indexes. But it should be kept in mind that this is, logically speaking, no longer a respondent-assisted or in cases of pure intermediate leads even pure link-trace sampling strategy but researcher-driven sampling. It further appears that closer scrutiny of referral types and search methods reveals complementarities in where and to what they are capable of working.

We have now reviewed the literature on rare subject identification and organised the field by clustering the main search methods according to their sampling logic and by presenting a classification of referral types together, and their implications for search methods. We next move to considering the opportunities that may emerge from combining different search methods to exploit their complementarities in the face of requirements and constraints imposed by different referral types.

Purposefully Combining Referral Types and Search Methods: Lead-User Mountaineering

The basic metaphor of pyramiding is finding one's steps up a pyramid to reach the top lead user(s). One of the earliest illustrations of such a lead-user search (see Fig. 6) was the networking approach (von Hippel *et al.*, 1999). To be precise, the figure presents an approach with search methods (snowball and pyramid sampling) and two referral types (person, field).

To date, pyramiding and broadcasting have been combined both in parallel (Hienerth *et al.*, 2007) and sequence (Keinz and Prügl, 2010), and multiple starting points have been used for pyramiding (Poetz and Prügl, 2010, p. 910). The above explication of intermediate referral types adds interest to method combinations in rare subject searches as they can potentially overcome hindrances and delays that a single search method can face. To take these experiments further we introduce more encompassing combinatory search that uses several referral types and several
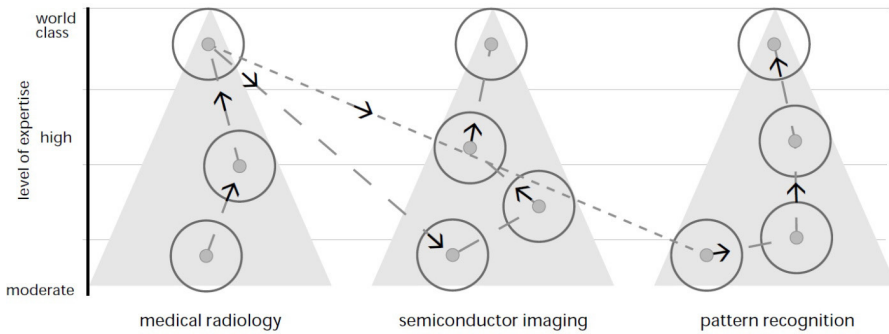
Fig. 6. Project teams network their way up "pyramids of expertise" to identify lead users and experts, first in the target market and then in other key fields (von Hippel *et al.*, 1999).

search methods to overcome constraints in referral types that become available during a search. We call this search approach *mountaineering*, because it is foremost multiple method hill-climbing, a way of "traversing upwards" towards those people, who have the sought-after characteristics (von Hippel *et al.*, 2009), but not limited to respondent-assisted pyramiding. The basic idea is thus to purposefully combine the referral types and search methods listed in Tables 1 and 2 in order to get at the sought after rare research subjects step by step. This can take place via multiple routes in parallel but emphasising those referrals and search methods that are most promising in a given moment. Such concurrent integrative search approach can be started with many given starting points and methods and can keep several search chains alive simultaneously, in so much that they do not jeopardise each other, for instance through same people being contacted repeatedly or by several means (see Fig. 7). As mountaineering is, in principle, not applicable to identifying only users with expertise, but any rare subjects, our *Y*-axis measures the amount of sought characteristic.

The sequence that such a search can take varies considerably with regard to the referral types encountered. As we shall document in the next section, searches in fields where Internet presence is widespread can rely on CMC and solution referrals displayed therein, and begin searching the people behind solutions and in those people's knowledge networks as a secondary step. In contrast, searches without publicly available Internet repository may be better served by taking multiple starting points and a respondent-assisted strategy concentrating on pyramiding that can be complemented by broadcasts, media scanning and more limited CMC analysis. Regardless of such shape of search process, this search approach means continuous consideration as to which search chains to pursue first and which ones are secondary, and which chains should be left to die off or if new search chains could or should be started. It thus, in principle, would
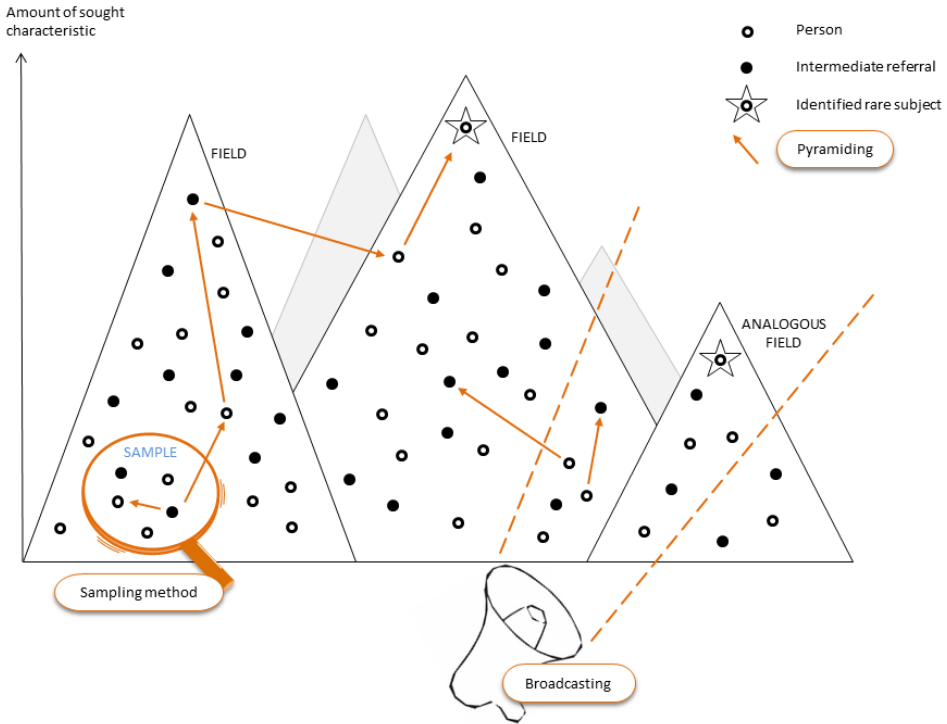
Fig. 7. The idea of mountaineering: "Traversing upwards" towards those persons or intermediate referrals who have the sought characteristics, with a suite of means from which one can select the most appropriate to the situation at hand.

capitalise on the learning effect during the search in advancing that search experiment, which is most sensible at a given step on the way (von Hippel *et al.*, 2009).

The progress of the rare subject search can be followed by monitoring the amount of sought characteristic of each referral. In lead-user searches it means monitoring the level of lead userness that can be measured with self-assessment questions (see below "Data and Methods"). In cases where such respondent-assisted monitoring is unattainable, expert evaluation or researcher's heuristic evaluation of a found solution can be used as proxy.

Elaborating Mountaineering through Real Life Cases

Data and methods

Below we report four principal and two supportive lead-user and user invention searches conducted in Finland during the years 2009–2012 by our six-person team

using the mountaineering search strategy in basic and applied research projects. All of the interviews conducted have been transcribed and lead userness has been assessed with self-assessment questions drawing on Franke *et al.* (2006), using a similar operationalisation of lead-user characteristics. Lead userness was measured by four seven-point Likert-scale questions; the scores were totalled without weighting, leading to a maximum rating of 28. The form of the questions was retained, while the content reference was changed from the original kite surfing context to the searches at hand, e.g., heating equipment (Appendix A). Table 4 presents the cases and their characteristics. The four principal cases are presented in detail in the following section; the first two relying more on parallel search strategies among different networks and the latter two based more on investigating user forums. The supportive cases resemble closely the wood pellet searches, and were here left without detailed description to save space in the face of little additional information gained.

In the diagrams,[6] which graphically document the actual search processes, the horizontal axis represents time from left to right in relative terms, not as an absolute scale. The vertical axis represents lead userness, that is, the sum of the self-assessment score. The lead userness of those users whose inventions were identified in forums but who did not respond to our contact requests, were rated with the aid of three domain expert evaluators who also rated the innovativeness of the user-developed concept. A referral that is not yet investigated is represented by a smaller circle, positioned close to the actor providing the referral, and is transformed into a bigger circle at the time of contact. Starting points for searches are placed close to the bottom of the vertical axes, since the starting points do not have a lead userness score. With regard to positioning of the circles in the graphs, while persons have been given a lead userness rating (i.e., clear *y*-axis position), the positioning of other circle types depend on nearby circles, as well as default values of the force-directed graph library. Where lines, labels, or circles would have become otherwise unreadable due to clutter, minor manual adjustments were made.

In the diagrams, we distinguish between different referral types with different coloured circles. Similarly, different search methods are represented with different coloured lines. Figures 8 and 9 explain how the referral types and search methods map to different colours, which is needed to interpret the diagrams in the following sections. We have hence condensed the different referral types (see "Intermediate search elements in lead-user identification methods") and search methods to nodes and edges (circles and lines) to give an overview of how the searches proceeded.

---

[6]The full search process depictions can be found in http://sn.im/mountaineering as animations.

Table 4. Case comparison.

		Principal cases				Supportive cases	
		Web service for teachers	Solar panels	Wood pellets	Solar thermal collectors	Ground heat pumps	Air heat pumps
Facts	Consumer good	no	yes	yes	yes	yes	yes
	High tech	no	yes	no	no	yes	yes
	User population[a]	65,000	50,000	25,000	10,000	70,000	440,000
	Well-defined domain	no	yes	yes	yes	yes	yes
	Public visibility of user activities	high	med	med	med	med	med
	Large online forum	no	no	yes	no	yes	yes
Project	Commissioned	yes	no	no	no	no	no
	Aim was to identify	5 workshop participants	5 workshop participants	as many lead users and user inventions as possible	as many lead users and user inventions as possible	as many lead users and user inventions as possible	as many lead users and user inventions as possible
	Multiple researchers	yes	no	no	no	yes	yes
	Tight schedule	yes	no	no	no	no	no

[a]Current user population of the product or service class in question. NB: Lead-user search is not limited to this population.

Circle Types

○ person ● organization ● index ● res.p.network

● cmc ● event ● field ○ referral

● solution ● mass media ● location

Fig. 8. The mapping between circle colours and referral types.

Line Types

—— screening

—— snowball and pyramid sampling

—— broadcasting

—— purposive, quota, and probabilistic sampling

—— solution content analysis

—— misc. encounters
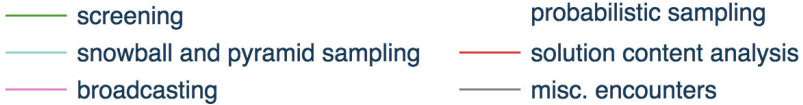
Fig. 9. The mapping between line colours and search methods.

Case web service for teachers

Project context

Finland's national public service broadcasting company (Yle) needed to redesign its web service for teachers, Opettaja.tv ("Teachers' TV"). The search goal was to find five lead users representing different trends relevant to the service to take part in a workshop. The trends included "richer web content and activities relevant to it, such as finding, managing, using, or producing the content," "easier feedback management from users to company and from users to users," "new ways of learning," "greater visibility," and "technologies for online learning." The service was open to everyone, although clearly aimed for teachers of elementary and secondary education who number approximately 65,000 in Finland.

Process

We used a range of sources to familiarise ourselves with the service and domain, and then ran two workshops with users and designers to refine the search goals, i.e., the trends relevant to the service. The initial starting points were acquired by various sampling strategies such as media scanning, Internet search, and miscellaneous means, and included personal contacts, workshop participants, project partners at Yle, and newspapers (Fig. 10). Miscellaneous sources for new starting points were not neglected at any point of the process.

After the first stage, the search was characterised by our pyramiding efforts beginning to hit users that had created their own inventions or were otherwise very knowledgeable of the teaching media domain (lead-use experts). Some of these
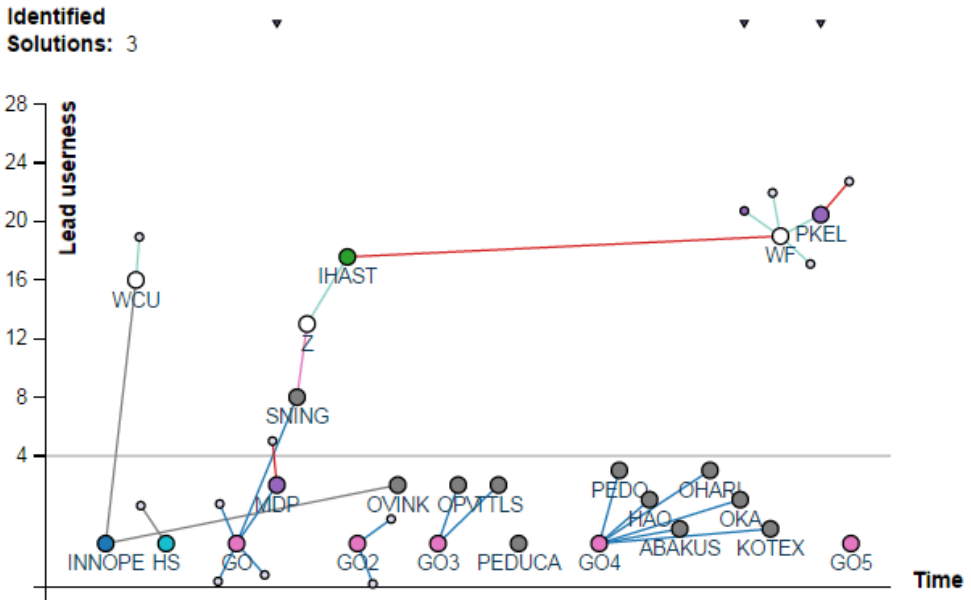
Fig. 10. First stage of the web service lead-user search and the longest chain taking shape: *Researcher-driven Sampling on index* (on Google GO), *broadcasting* (in an online community SNING), *pyramiding (*from a person Z), *solution content analysis* (inside an organisation IHAST), *pyramiding* (from a person WF), and finally *solution content analysis* (person behind a solution PKEL). The person at the end of the chain was later confirmed to be a lead user. The lower section shows the multitude of referrals from early familiarising of which several would later be used as a place for broadcasting.

people provided extensive lists of further referrals and the team had to concentrate on choosing which referrals to follow. In Fig. 11 this is visible in the great number of referrals that still remained unfollowed at the end of the process. The selection of the referrals to follow was based on how promising the referrals appeared to be for the researchers and whether they would help in covering all the relevant development areas of the service. Pyramiding continued until the final stages of the process, broadcasting still running at the background, and late in the process we found one more starting point that eventually led to a lead user.

In all, the search was comprised of 33 referrals to persons of which 19 were contacted and 82 other referrals. Eventually five lead users representing different areas of the service were identified and selected for a lead-user workshop. Many of the referrals were deemed unnecessary to follow up because they were deemed to be irrelevant. Several referrals were either commercial or open established solutions and many were also bypassed due budget and time constraints as the required five "lead user enough" persons were found in relevant areas of the web service. As von Hippel *et al.* (2009, p. 1403) underscore: "In real life searches it is almost
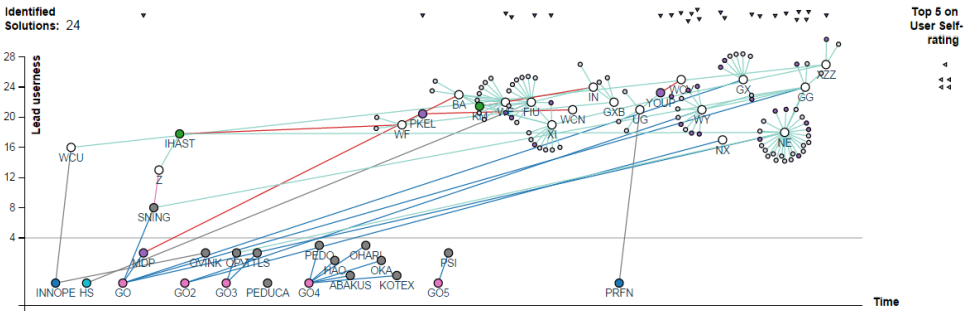
Fig. 11. Final stage of the web service lead-user search: Pyramiding further with broadcasting still on the background. A late new starting point (PRFN) led to a lead user.

always the case that one does not necessarily need to reach the 'top of the pyramid' in order to get an appropriate solution." The workshop results and insights gained during the process were condensed into a new concept of Opettaja.tv web service, which was considered thrilling and warmly welcomed at Yle. Its uptake and implementation is ongoing.

Case highlights

This case highlights the use of multiple parallel search strategies, adapting to the situation, and the use of several starting points. Lead users were found through referrals to solutions and organisations. The longest chain, taking shape in Fig. 10 and continuing in Fig. 11, contains five different referral types and four different search methods. New starting points after the initial phase revealed a user innovator behind his superior solution. Some of the advanced social media solutions were found from analogous fields.

Lead-user example

A high school mathematics teacher who has, for years, uploaded short self-made "how-to" video clips on YouTube to help his own pupils do their homework; at the time a rare and new activity.

Case solar panels

Project context

The solar panels (photovoltaics, PV) search was designed to become comparable with the web service case through having a specific business target, five relevant

trends, and a technology domain that does not feature a large national Internet user forum (see cases below). The search was then continued in the second stage to find as many user innovations as possible, but here we report this comparable first stage only. Solar PV technology is a high-tech domain unlikely to be accessible to user innovators, yet a working installation is much more than solar panels: mounting angle, location, and structure, as well as electrical power infrastructure.[7] The trends were "solar tracking," "humidity and snow mitigation techniques," "design (aesthetically pleasing installations)," "installation mounting structures," and "hybrid systems." Most solar panel installations in Finland are off-grid, and a cautious estimate would be that there are around 50,000 installations, mostly in summer cottages.

Process

The search process started with referrals from knowledgeable colleagues and the solar electricity discussion section of ilmaisenergia.info Internet forum (identified in the supporting solar thermal case, see below). 32 discussion threads with 636 messages and sporadic discussions concerning a handful of DIY-projects were found. The most leading edge user had already been identified and interviewed in the solar thermal case. The researcher continued gathering starting points, which included participating in two professional development events on solar energy. Figure 12 illustrates the early phase of this case.

The next steps included interviews with the two most referred referrals (ZJ and PA), but they had no further contacts to relevant solar PV lead users, nor did the two other interviewed solar business referrals (CR and EE, see Fig. 13). However, one of the professional development events pointed to a lead-use expert, who then referred to a lead user.

Discussions with engineer friends in the personal networks of one of us, revealed two other lead users. Some time into the search, a user, who was in the middle of figuring out solar tracking, posted new messages on the forum. As that user turned out to be a lead user, the goal of finding five lead users or lead-use experts was met. Although there were no large solar panel online forums available, even the small number of posts led to identifying two lead users. In all, the solar panels search was comprised of 43 referrals to persons of which 15 were contacted and 36 other referrals. Five lead users representing relevant trends were identified.

---

[7]Current regulators, DC/AC current inverter, cables, and possibly batteries to store electricity, as well as reserve power generators.
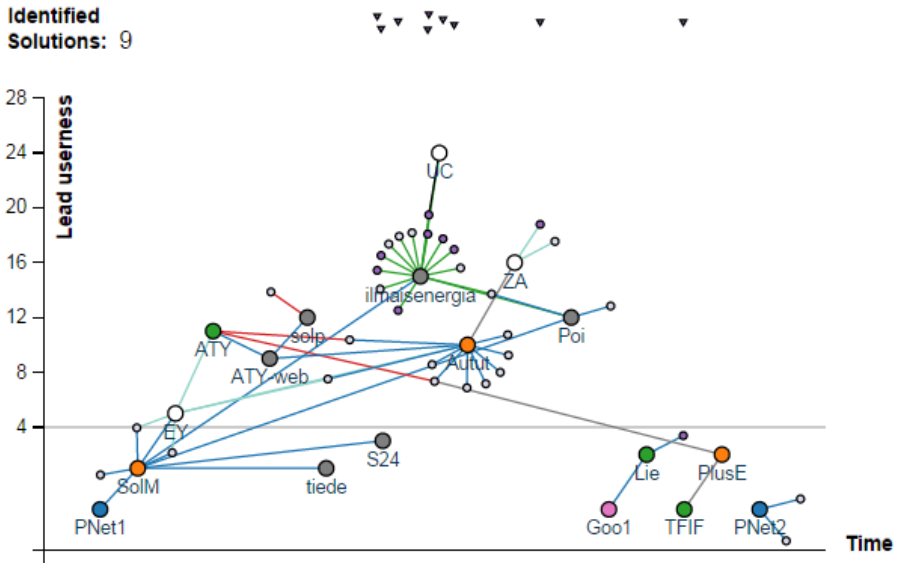
Fig. 12. First stage of the solar panel mountaineering: As forums did not yield sufficient referrals, new starting points were sought.
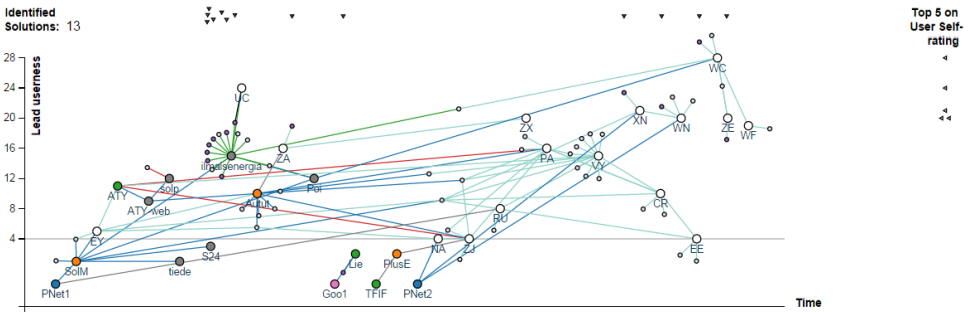

Fig. 13. Final stage of the solar panel mountaineering: With the help of new starting points and new messages on the previously followed forum, the search accelerated and lead users were found.

Case highlights

This search featured a mix of personal networks, business networks, and online methods to identify lead users. Since solar panel technology is very high-tech and had low public visibility in Finland, starting points were scarce. The researcher took part in business events to learn about the technology, related trends, and referrals. Most central actors in the business network were domain experts, as expected, but these did not know any lead users.

Lead-user example

A building engineer, who was disappointed with the market offering at the time, decided to import solar panel system components himself, in order to achieve lower costs and improved materials suited for the humidity of the Finnish Archipelago. The successful solving of importing and installation problems for his summer cottage led him to start a business in the domain.

Case wood pellets (supported by heat pump cases)

Project context

This case was part of a research project to identify and analyse user inventions in renewable home heating systems. The goal was to find as many as possible lead users and user inventions in wood burning pellets, ground heat pumps, and air heat pumps in Finland. In these searches we found 192 user inventions and modifications, which three domain experts verified for us. In wood pellet burning systems we found in total 79 inventions (Hyysalo *et al.*, 2013a,b). All of these three technologies and the modifications made to them are actively discussed in large national Internet forums, which affected strongly what kinds of search sequences were sensible. As these three cases resemble each other we describe here only the wood pellet search in detail.

Process

Having conducted two searches on heat pumps that benefited from screening Internet discussion forums (CMC), we again first scanned the web for forums relating to wood pellets and found a large active forum pellettikeskustelu.net, which had 1,897 registered members featuring 46,830 posts in 3,194 threads and an "own inventions/constructions" section comprising 1,635 messages in 123 threads. All messages in the section was screened by content analysis for user inventions and the effort provided us with the potential and confirmed lead users visible in Fig. 14. Alongside, we carried out simultaneous pyramiding by starting with 10 interview requests through the website. The names coming up in the interviews were, however, also found through screening, partly because the forum also revealed *indigenous pyramiding* within, as the forum posts internally referenced leading users to other people.

In the course of the study we sent 20 more interview requests. Thirteen out of the total of 30 responded and agreed to be interviewed. These interviews led to four new suggestions of lead users on the forum, of which two agreed to be interviewed but gave no further referrals. The interviewees pointed us to specific
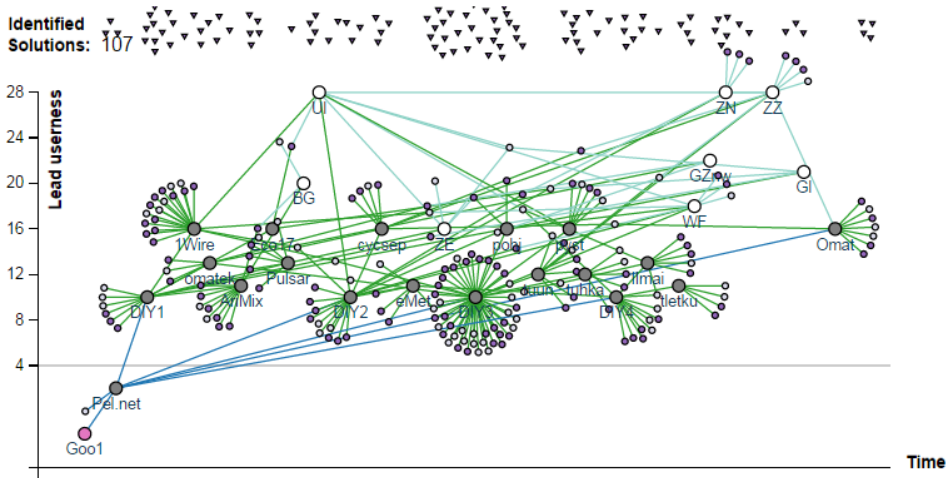
Fig. 14. Early stage of the wood pellet lead-user search: Forum screening reveals more lead users who become confirmed in the course of pyramiding interviews and a new forum section (the rightmost circle, Omat) is screened only after a second referral, from ZZ.

sections in the picture gallery of the forum, which lead to also screening these sections fully (see Fig. 14). This led to one already identified lead user and three further potentials that did not respond.

In addition we used five other reference points external to the forum. Our Google searches conducted in the very late stage still revealed one lead user (WYn). The leading national newspaper "Helsingin Sanomat" (HS) featured an article about a user innovators' pellet burner project, but the maker did not respond to our interview request. Another research project "eco-forerunners" featured a user-manufacturer, who then pointed to two further subjects, of which one turned out to be a domain expert and the other deceased (Fig. 15).
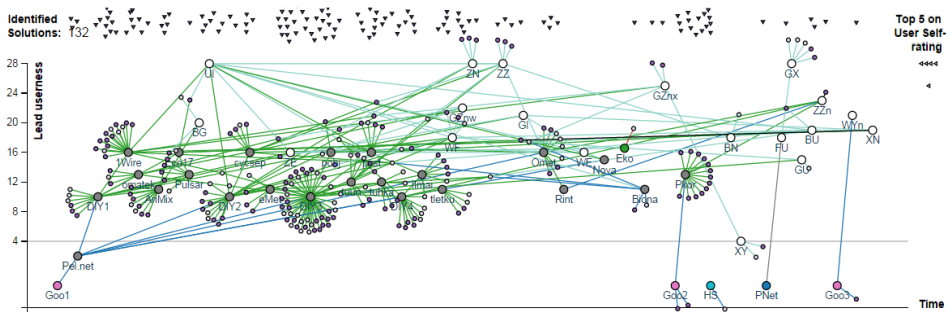

Fig. 15. Final stages of wood pellet lead-user search: Forum screening reveals still more lead users, new forum section, and five more starting points (forum referrals reduced in the illustration).

In all, the wood pellet search was comprised of 84 referrals to persons of which 18 were contacted and 161 other referrals. The 18 contacted persons included 10 people who had the score of 20 or more in the lead-user self-assessment. In total these searches revealed 67 user innovations or modifications.

Case highlights

Wood pellet and heat pump user innovation searches exceeded our expectations in that we were able to identify a greater number and variety of user inventiveness than we expected. User forums appear to present a promising environment for rare subject searches, particularly if the aim is to cover the area. However, content analysis of posts takes time and may not be feasible in commercial projects. Forum activity and other "netnography" methods can lead to lead users, but not necessarily. For instance, in ground source heat pumps the inventive users had only a moderate level of postings, and details revealing them required reading into the threads in detail. Pyramiding in CMC is greatly hampered by slow response times and receiving no responses — we experienced lags of weeks and months. This calls to question attempts at optimising pyramiding by just opting for the most promising subject and discarding other routes. Second referrals led to decisions to screen more forum sections and a Google search in the very late stage still revealed one lead user. Thus, a combination of search methods and use of multiple starting points also appears as a feasible strategy in the presence of large Internet forums, particularly if one needs to perform the search quickly.

Lead-user example

A metal technician-plumber-IT person developed a novel pellet transfer system because there was no commercial product available. It was also cheaper to do-it-himself from what comprised of mostly recycled materials. His pellet transfer system draws pellets from a larger area than a spiral conveyer, and digs down to the bottom of the pellet silo thus preventing pellets from arching and ash from accumulating. He produces and sells the system "Pellet Elephant" to other forum members (so far 100 pieces) for a minimal mail delivery fee.

Case solar thermal collectors

Project context

This case is also part of the renewable home heating systems project. The goal was to find as many as possible lead users and user inventions in Finland in solar thermal collectors, which features approximately 10,000 installations.

Process

Our first starting point was an offshoot from pyramiding in the earlier heat pump study: A lead user pointed us to a smaller renewable energy forum (ilmaisenergia. info) (CMC) that had a section on solar thermal collectors. The energy section in Tiede magazine's Internet forum (a Finnish science magazine) was another small forum starting point for us. The discussion scope was modest in both forums, and full screening resulted in 13 and 3 potential contact points respectively in these two forums (Fig. 16). We also already knew one lead-user contact given by a lead-use expert and our earlier renewables searches had already pointed to another lead user in solar collectors. We also carried out Internet searches to find users with blogs, resulting in three names.

Altogether we had 21 potential lead-user names, which already included two persons who were interviewed in earlier renewable energy studies. Interview invitations went to 19 persons. Eleven gave a positive response and were interviewed. We used further pyramiding and snowball sampling in these 11 interviews, leading to 15 new referrals three previously recognised names and three companies. However, of these 15 new pyramiding referrals, only two users accepted our interview request. The new referrals obtained in these interviews led
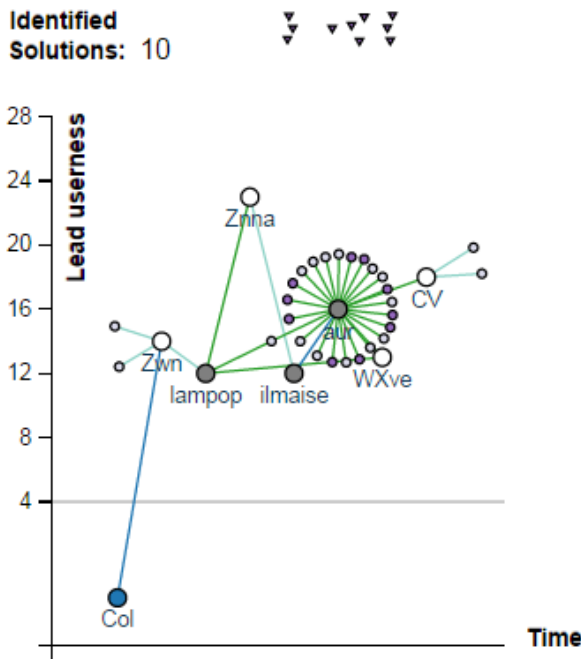

Fig. 16. First stage of solar thermal lead-user search: Screening on forums reveled few potential lead users.
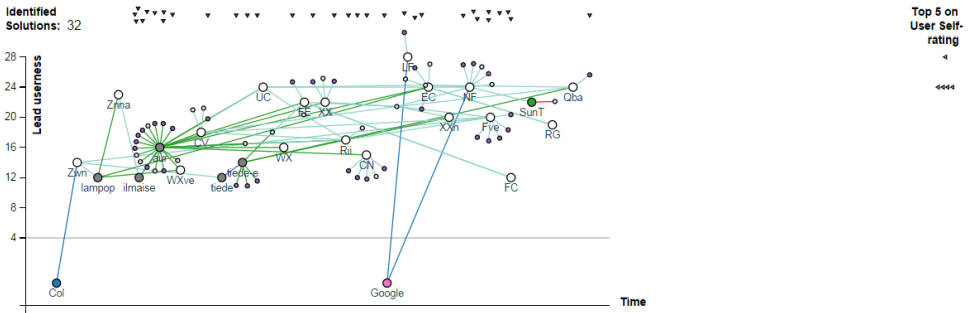
Fig. 17. Final stage of solar thermal lead-user search: Internet search, pyramiding and snowball sampling have saturated.

only to already known contact points (Fig. 17). At this point we were "at a peak," but could not tell if it was a lower peak or not. Alternative sampling such as broadcastings could have been used to get us beyond the networks we had traced through, but were not pursued due to lack of time in the project. In all, the solar thermal collector search was comprised of 35 referrals to persons of which 19 were contacted and 41 other referrals. The 19 contacted persons included 9 who had the score of 20 or more in the lead-user self-assessment.

Case highlights

The solar collector search indicates complementarities between search methods and referral types when one or the other of the methods is not clearly superior for the purpose at hand. The parallel searches helped to make a fast and efficient study despite the facts that none of the forums or user communities concentrated user innovation, and that there were no means to delineate effectively a population to which, for example, a lead-user survey could have been sent.

Lead-user example

A user created a home-sized concentrated solar collector with an advanced solar tracker that was optimised for Finnish sun conditions all year round. The concentrator heated fluid in a small tank and the fluid circulation to the heat exchanger was also controlled with sensors and algorithms to retain optimal working temperature and maximise yield.

Summary of cases

Table 5 summarises the key statistics of the search process for the cases, including the number of referrals suggested during the whole search process, how many

Table 5. Summary of cases and characteristics of searches for finding 5 lead users in each case.

	Web service for teachers	Solar panels	Wood pellets	Solar thermal collectors	Percentage of total	Range of case totals
Use of Methods						
Screening	0	1	20	3	16.0%	0.0–44.4%
Snowball and pyramid sampling	16	18	16	19	46.0%	35.6–70.4%
Broadcasting	12	0	0	0	8.0%	0.0–29.3%
Purposive, quota, and probabilistic sampling	5	12	7	4	18.7%	12.2–32.4%
Solution content analysis	5	2	1	1	6.0%	2.2–12.2%
Miscellaneous encounters	3	4	1	0	5.3%	0.0–10.8%
Total	41	37	45	27	100.0%	
Starting points						
Personal network	2	2	1	1	30.0%	20.0–50.0%
CMC	1	0	0	0	5.0%	0.0–11.1%
Organisation	0	1	0	0	5.0%	0.0–25.0%
Media	1	0	1	0	10.0%	0.0–20.0%
Index	5	1	3	1	50.0%	25.0–60.0%
Total	9	4	5	2	100.0%	
Referrals						
Person	33	43	84	35	37.9%	28.7–54.4%
CMC	22	7	24	5	11.3%	6.6–19.1%
Solution	24	13	132	32	39.0%	16.5–53.9%

Table 5. (*Continued*)

	Web service for teachers	Solar panels	Wood pellets	Solar thermal collectors	Percentage of total	Range of case totals
Organisation	33	13	5	4	10.7%	2.0–28.7%
Event	0	3	0	0	0.6%	0.0–3.8%
Mass media	1	0	0	0	0.2%	0.0–0.9%
Index	0	0	0	0	0.0%	0.0%
Field	2	0	0	0	0.4%	0.0–1.7%
Location	0	0	0	0	0.0%	0.0%
Total	115	79	245	76	100.0%	
Top 5						
Total steps (= chain length)	3/1/3/7/2	3/8/1/1/5	3/3/5/3/2	6/1/1/3/3		
Case highlight	Multiple starting points	Personal network	Internet Forum	Shifts between search methods		

people were contacted, and how many initial starting points were used. The variance in search method usage is visible under "Use of Methods." The term "Top 5" refers to the parts of the search process leading to the top-5 lead users, which we chose as an additional comparison frame between the cases. It documents the total number of steps from the starting point to the finding of each of the top-5 lead users.

Discussion and Conclusions

Network searches are increasingly common in lead-user identification processes. To date, these processes have been investigated as chains where individuals refer to other individuals (e.g., Churchill *et al.*, 2009; Hienerth *et al.*, 2007; Lüthje and Herstatt, 2004; Stockstrom *et al.*, 2012), however, real-world network searches also include a substantial number of implicit and inexact referrals (Poetz and Prügl, 2010). These intermediate referrals, while useful if addressed appropriately, may complicate the search process with delay and uncertainties. To clarify matters we distinguished between referral types, method types, and their characteristics that we defined as intermediate search elements.

In the course of this paper we have explicated the range of intermediate referral types that commonly arise in real-world rare subject searches, and elaborated the different constraints they set for different search methods in terms of contacting, sampling, and investigating these referrals. In particular, we drew attention to necessary shifts in the sampling logic when it comes to pure intermediate referrals: in effect resulting in a mix of respondent-assisted and researcher-driven sampling. With semi-intermediate and semi-immediate referrals delay and extra effort may ensue, sometimes also uncertainty over whether such contact will ever respond. This clarification is our primary contribution.

The intermediate referral types and sampling logic changes invite considering the possibilities that search method combinations may have in overcoming referral type limitations and capitalising on different strengths of main rare subject search methods (cf. Hienerth *et al.*, 2007; Keinz and Prügl, 2010). We thus introduced a more encompassing combinatory search approach, a concurrent integrative search we called mountaineering for clarity (as well as, to pay homage to the hill-climbing idea and pyramiding search that inspired us in the work by von Hippel *et al.*, 2009).

As a first proof of concept we ran six searches with this search approach and achieved the search outcomes to the satisfaction of our research goals and those of our client. This proof of concept of purposeful combinatory searches is our secondary contribution.

Literature to date has suggested a range of methods and directions as a means to finding lead users. Effort has been made to compare and choose between methods (e.g., von Hippel *et al.*, 2009; Stockstrom *et al.*, 2012), but possibilities for systematically combining different methods have remained under-explored.[8] Our mountaineering searches covered the whole gamut of different methods, both *respondent-assisted* and *researcher-driven* by sampling logic, as well as the full range of referral types identified thus far. Looking at our cases, the most productive search strategy appears to vary significantly from one case to another as well as with regard to the phase of the particular case. The mix of a domain expert–broadcasting–pyramiding search strategy used in the web service case would have been less productive in the heat pump and wood pellet searches where we could opportunistically use large Internet forums by first drawing researcher-driven samples and then screening these by content, followed by pyramiding. Screening by content Internet or other communities of interest can be effective, particularly if there is a self-nominated subsection of the population active regarding the sought search attribute. Otherwise, a researcher-driven sampling strategy would be needed targeting the sweet spots for the information and people searched for.[9] Similarly, the known downside of link-tracing strategies in missing isolates (Atkinson and Flint, 2001; van Meter, 1990) can be compensated by using multiple starting points as we did in all of our searches: Figure 11 relating to the web service search and Fig. 15 to the wood pellet search show a lead user being found very late in the process and not connected to the initial chains.

To explore further the effects and possibilities that result from multiple referral types and their requirements for search methods the following further research avenues appear to be the most salient. The first is to extend pyramiding simulations so that multiple referral types are included within the population examined instead of only persons to see what effects this would have on pyramiding searches. This would in all likelihood mean treating all referral types as information

---

[8]Discussion similar to this has recently emerged in the field of human–computer interaction — a field where the debate on choosing the "best" method for some specified context has prevailed for the past decades. Woolrych *et al.* (2011) argue that only very few comparative research studies investigate methods as they are mostly used in practice: as combinations of methods and their components to best fit the task at hand.

[9]Our experience has made us skeptical of relying on mechanistic netnography such as forum statistics, search functions, or crawlers in uncovering user innovation and lead users. We tried them and they delivered a small subsection of the innovators, which surfaced in the less mechanistic method combinations we pursued. The mechanistic means did not result in any such solutions or people we did not discover otherwise.

containers that can be investigated with different costs and which point to the next source, albeit in a different manner. Further work is here needed, however, to develop heuristics for the degree of sought attribute in referrals that are not respondent-assisted, as well as to estimate sensible values for the time and other costs involved in the periods of researcher-driven sampling necessarily involved in following through pure intermediate referrals.

A second avenue would be set-ups to measure search effectiveness and efficiency in multiple method searches. Our empirical studies indicate that chain length used in social network analyses and pyramiding simulations may not be the only viable measure in combinatory searches. Long chains can unravel fast and with relatively little effort, whilst short chains can take a long time and tedious effort to accomplish if respondents are not forthcoming or much investment needs to be made to attract their response. Search time, both invested effort time as well as calendar time, could be explored as alternative measures.

A third avenue that we encourage is also the main management implication from the present research. In many problem areas that are both practical and theoretical, it has in the course of history been wise to pursue both. Thermodynamics emerged 50 years after steam engines had been in productive use (de Solla Price, 1984). As recently as in 1980s the psychological principles for human–computer interaction were sidestepped by practical usability guidelines that did a better job at improving user interfaces. It took several years for psychologists to work out the underlying principles and why the guidelines worked, all the while designers and users were happy using and developing them, assisting the psychologists efforts in doing so (Kuutti, 2001). In both cases the eventual theoretical explanations required a stock of well-documented empirical material to identify the underlying principles. In rare subject search method combinations it may also be wise not to wait for years for the ultimate truths to emerge as to when and where different combinatory searches may be most effective or which may, in principle, be the most effective combinations, but also to conduct and document such searches alongside. A cautious "something for nothing" trial strategy would be to first use parallel or multiple method search sequences in cases where a single method search is stuck, tedious or slow. A more daring trial would be to experiment with limited parallel searches and method combinations to speed up the search calendar-wise for lead-user projects. Our "all out" mountaineering type search, using several parallel search sequences, all referral types and multiple search method combinations, is also a possibility. It is too early to talk of its efficiency, but we have shown that it can be used in both lead-user project searches as well as in user innovation research projects in different settings to the satisfaction of clients and for meeting the search needs in our research projects.

Appendix A. Self-Assessment Questions Used; Example from the Solar Thermal Collector Case

These questions follow the operationalisation of lead-user characteristics by Franke *et al.* (2006) and Stockstrom *et al.* (2012). Franke *et al.* developed a set of questions for each lead-user characteristic, from which Stockstrom *et al.* selected the question with the highest Item-to-Total correlation. While Franke *et al.* measured the "Ahead of trend" construct with a Thurstone scale, Stockstrom *et al.* used a Likert-scale question, and we followed the latter (Table A.1).

Table A.1. Lead userness questions, measured on a seven-point Likert scale.

Lead-user characteristics	Question
Ahead of a Trend	I have improved heating equipment on my own.
Technical Expertise	I can make technical changes to my heating equipment on my own.
High Benefit Expected	I have already had problems with my heating equipment that could not be solved with the manufacturer's conventional offerings.
Community-Based Resources	I know many other people who optimise heating equipment and have a thorough knowledge of heating equipment.

Acknowledgments

References

Atkinson, R and J Flint (2001). Accessing hidden and hard-to-reach populations: Snowball Research Strategies. *Social Research Update*, 28(1), 93–108.

Avico, U, C Kaplan, D Korczak and K van Meter (1988). *Cocaine Epidemiology in Three European Community Cities: A Pilot Study Using a Snowball Sampling Methodology*. Brussels, Belgium: Health Directorate, Commission of the European Communities.

Baldwin, C, C Hienerth and E von Hippel (2006). How user innovations become commercial products: A theoretical investigation and case study. *Research Policy*, 35(9), 1291–1313.

Belz, F-M and W Baumbach (2010). Netnography as a method of lead user identification. *Creativity and Innovation Management*, 19(3), 304–313.

Bilgram, V, A Brem and K-I Voigt (2008). User-centric innovations in new product development — Systematic identification of lead users harnessing interactive and collaborative online-tools. *International Journal of Innovation Management*, 12(3), 419–458.

Chubin, DE (1976). The conceptualization of scientific specialties. *The Sociological Quarterly*, 17(4), 448–476.

Churchill, J, E von Hippel and M Sonnack (2009). *Lead User Project Handbook: A Practical Guide for Lead User Project Teams*. Mass.: MIT. available at: http://web.mit.edu/evhippel/www/Lead User Project Handbook %28Full Version%29.pdf.

Crane, D (1969). Social structure in a group of scientists: A test of the "invisible college" hypothesis. *American Sociological Review*, 34(3), 335–352.

Daniel, J (2012). *Sampling Essentials: Practical Guidelines for Making Sampling Choices*. Thousand Oaks, CA: SAGE Publications.

de Solla Price, D (1984). The science/technology relationship, the craft of experimental science, and policy for the improvement of high technology innovation. *Research Policy*, 13(1), 3–20.

Droge, C, MA Stanko and WA Pollitte (2010). Lead users and early adopters on the web: The role of new technology product blogs. *Journal of Product Innovation Management*, 27(1), 66–82.

Edge, DO and MJ Mulkay (1974). Case studies of scientific specialties. University of Edinburgh, Science Studies Unit.

Faugier, J (1995). Looking for business: A descriptive study of drug using female prostitutes, their clients and their health care needs. PhD thesis, University of Manchester, School of Nursing Studies, Manchester, UK.

Franke, N and S Shah (2003). How communities support innovative activities: An exploration of assistance and sharing among end-users. *Research Policy*, 32(1), 157–178.

Franke, N, E von Hippel and M Schreier (2006). Finding commercially attractive user innovations: A test of lead-user theory. *Journal of Product Innovation Management*, 23(4), 301–315.

Gobo, G (2004). Sampling, representativeness and generalizability. In *Qualitative Research Practice*, C Seale, G Gobo, JF Gubrium and D Silverman (Eds.), pp. 435–456. London: Sage Publications.

Goodman, LA (1961). Snowball sampling. *Annals of Mathematical Statistics*, 32(1), 148–170.

Griffiths, P, M Gossop, B Powis and J Strang (1993). Reaching hidden populations of drug users by privileged access interviewers: Methodological and practical issues. *Addiction*, 88(12), 1617–1626.

Hannukainen, P and K Hölttä-Otto (2006). Identifying customer needs — Disabled persons as lead users. In *Proceedings of IDETC/CIE 2006*, September 10–13. Philadelphia, Pennsylvania, USA: ASME 2006 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.

Hauser, JR and D Clausing (1988). The house of quality. *Harvard Business Review*, 3(May–June), 63–73.

Herstatt, C and E von Hippel (1992). From experience: Developing new product concepts via the lead user method: A case study in a "low-tech" field. *Journal of Product Innovation Management*, 9(3), 213–221.

Hienerth, C (2006). The commercialization of user innovations: The development of the rodeo kayak industry. *R&D Management*, 36(3), 273–294.

Hienerth, C, M Pötz and E von Hippel (2007). Exploring key characteristics of lead user workshop participants: Who contributes best to the generation of truly novel solutions? In *Proceedings of DRUID 2007*, June 18–20, Copenhagen, Denmark.

Hienerth, C, E von Hippel and M Berg Jensen (2013). User community vs. producer innovation development efficiency: A first empirical study. *Research Policy*, 43(1), 190–201.

Hyysalo, S, JK Juntunen and S Freeman (2013a). Internet forums and the rise of the inventive energy user. *Science and Technology Studies*, 26(1), 25–51.

Hyysalo, S, JK Juntunen and S Freeman (2013b). User innovation in sustainable home energy technologies. *Energy Policy*, 55(Special issue), 490–500.

Inciardi, JA (1977). In search of the class cannon: A field study of professional pickpockets. In *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings*, RS Weppner (Ed.), pp. 55–77. Beverly Hills, CA, USA: Sage.

Jeppesen, LB and L Frederiksen (2006). Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. *Organization Science*, 17(1), 45–63.

Jeppesen, LB and KR Lakhani (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization Science*, 21(5), 1016–1033.

Kalton, G (2009). Methods for oversampling rare subpopulations in social surveys. *Survey Methdology*, 35(2), 125–141.

Kaplan, CD, D Korf and C Sterk (1987). Temporal and social contexts of heroin-using populations: An illustration of the snowball sampling technique. *The Journal of Nervous and Mental Disease*, 175(9), 566–574.

Kaulio, MA (1998). Customer, consumer and user involvement in product development: A framework and a review of selected methods. *Total Quality Management*, 9(1), 141–149.

Keinz, P and R Prügl (2010). A user community-based approach to leveraging technological competences: An exploratory case study of a technology start-up from MIT. *Creativity and Innovation Management*, 19(3), 269–289.

Kozinets, RV (1998). On netnography: Initial reflections on consumer research investigations of cybetrculture. *Advances in Consumer Research*, 25(1), 366–371.

Kozinets, RV (2010). *Netnography: Doing Ethnographic Research Online*. London: Sage.

Kuutti, K (2001). Hunting for the lost user: From sources of errors to active actors — and beyond. In *Cultural Usability Seminar at Media Lab, University of Art and Design Helsinki*. Available at http://mlab.uiah.fi/culturalusability/papers/Kuuttipaper.html.

Lakhani, KR (2006). Broadcast search in problem solving: Attracting solutions from the periphery. MIT Sloan School of Management working paper.

Lilien, GL, PD Morrison, K Searls, M Sonnack and Eric von Hippel (2002). Performance assessment of the lead user idea-generation process for new product development. *Management Science*, 48(8), 1042–1059.

Lohr, SL (2009). *Sampling: Design and Analysis*, 2nd Edition. Boston, MA: Brooks/Cole.

Lüthje, C (2000). *Kundenorientierung im Innovationsprozess: Eine Untersuchung der Kunden-Hersteller-Interaktion in Konsumgütermärkten*. Wiesbaden: Gabler.

Lüthje, C and C Herstatt (2004). The lead user method: An outline of empirical findings and issues for future research. *R&D Management*, 34(5), 553–568.

Lüthje, C, C Herstatt and E von Hippel (2005). User-innovators and "local" information: The case of mountain biking. *Research Policy*, 34(6), 951–965.

McNamara, RP (1994). *The Times Square Hustler: Male Prostitution in New York City*. Westport, CT, USA: Praeger.

McQuail, D (1987). *Mass Communication Theory: An Introduction*, 2nd Edition. Thousand Oaks, CA: Sage Publications, Inc.

Newman, MEJ (2003). The structure and function of complex networks. *SIAM Review*, 45(2), 167–256.

Olson, EL and G Bakke (2001). Implementing the lead user method in a high technology firm: A longitudinal study of intentions versus actions. *Journal of Product Innovation Management*, 18(6), 388–395.

Pals, N, MGD Steen, DJ Langley and J Kort (2008). Three approaches to take the user perspective into account during new product design. *International Journal of Innovation Management*, 12(3), 275–294.

Patton, MQ (2002). *Qualitative Research & Evaluation Methods*, 3rd Edition. Thousand Oaks, CA: Sage Publications, Inc.

Piller, FT and D Walcher (2006). Toolkits for idea competitions: A novel method to integrate users in new product development. *R&D Management*, 36(3), 307–318.

Poetz, MK and R Prügl (2010). Crossing domain-specific boundaries in search of innovation: Exploring the potential of pyramiding. *Journal of Product Innovation Management*, 27(6), 897–914.

Pollak, M and MA Schiltz (1988). Does voluntary testing matter? How it influences homosexual safer sex. In *Proceedings of Fourth International Conference on AIDS*, Stockholm, Sweden.

Raasch, C, C Herstatt and P Lock (2008). The dynamics of user innovation: Drivers and impediments of innovation activities. *International Journal of Innovation Management*, 12(3), 377–398.

Spann, M, H Ernst, B Skiera and JH Soll (2009). Identification of lead users for consumer products via Virtual Stock Markets. *Journal of Product Innovation Management*, 26(3), 322–335.

Spreen, M (1992). Rare populations, hidden populations, and link-tracing designs: What and why? *Bulletin de Méthodologie Sociologique*, 36, 34–58.

Stockstrom, CS, RC Goduschiet, JH Jørgensen and C Lüthje (2012). Identification of individuals with special qualities — Assessing the performance of pyramiding search. In *Proceedings of DRUID 2012*, June 19–21, Copenhagen, Denmark.

Sudman, S (1980). Improving the quality of shopping center sampling. *Journal of Marketing Research*, 17(4), 423–431.

Sudman, S (1985). Efficient screening methods for the sampling of geographically clustered special populations. *Journal of Marketing Research*, 22(1), 20–29.

Sudman, S and G Kalton (1986). New developments in the sampling of special populations. *Annual Review of Sociology*, 12, 401–429.

van Meter, KM (1990). Methodological and design issues: Techniques for assessing the representatives of snowball samples. In *NIDA Research Monograph 98 "The Collection and Interpretation of Data from Hidden Populations."*, EY Lambert (Ed.), pp. 31–43. Washington, DC: NIDA.

von Hippel, E (1976). The dominant role of users in the scientific instrument innovation process. *Research Policy*, 5(3), 212–239.

von Hippel, E (1986). Lead users: A source of novel product concepts. *Management Science*, 32(7), 791–805.

von Hippel, E (1988). *The Sources of Innovation*. New York, NY, USA: Oxford University Press.

von Hippel, E (2005). *Democratizing Innovation*. Cambridge, Mass., USA: MIT Press.

von Hippel, E, N Franke and R Prügl (2009). Pyramiding: Efficient search for rare subjects. *Research Policy*, 38(9), 1397–1406.

von Hippel, E, S Thomke and M Sonnack (1999). Creating breakthroughs at 3M. *Harvard Business Review*, 77(5), 47–57.

Welch, S (1975). Sampling by referral in a dispersed population. *The Public Opinion Quarterly*, 39(2), 237–245.

Woolrych, A, K Hornbæk, E Frøkjær and G Cockton (2011). Ingredients and meals rather than recipes: A proposal for research that does not treat usability evaluation methods as indivisible wholes. *International Journal of Human-Computer Interaction*, 27(10), 940–970.