Improving Comprehensibility of Medical Information
Proof-of-Concept for Patient Information Leaflets

Till Blesik
University of Cologne, Faculty of Management, Economics and Social Sciences, D-50923 Cologne,
E-Mail: till.blesik@smail.uni-koeln.de

Tobias Dehling
University of Cologne, Faculty of Management, Economics and Social Sciences, D-50969 Cologne,
E-Mail: dehling@wiso.uni-koeln.de

Ali Sunyaev
University of Cologne, Faculty of Management, Economics and Social Sciences, D-50969 Cologne,
E-Mail: sunyaev@wiso.uni-koeln.de

Abstract
Many patients have difficulties with comprehending written and spoken health information presented
to them. Additionally, detailed explanation of medication use and further implications are too complex
and expensive to be carried out by medical professionals whenever they are prescribing medication.
An alternative approach to transform medical information in an easier understandable form could save
cost for medical service providers and help to increase patients’ adherence. We present a requirements
framework for medical information translation systems. Furthermore, concept, architecture, and actual
implementation of a web application leveraging crowdsourcing are illustrated. To demonstrate that the
crowdsourcing approach is suitable to improve comprehensibility of medical information, a proof-of-
concept experiment is conducted.

1 Introduction
Without a medical background, patients are neither able to fully comprehend given written and spoken
health information nor able to interpret it correctly [16, 27]. Transformation of such information into
an easier understandable form by the prescribing medical professional is a complex and therefore
expensive operation. Transformation of the information into a more comprehensible format would
save cost for medical service providers and allow for creation of innovative applications for presenting
medical information. Furthermore, lack of understanding is one barrier to adherence [30] – “the extent
to which a person’s behavior […] corresponds with agreed recommendations from a health care
provider” [11]. By providing easy understandable medical information, it could be possible to enhance
patients’ adherence and increase the quality of their treatment. A better understanding of their illness
could additionally motivate patients to be proactive, empower them to take a more active role in managing their own health, and enable them to reap more benefits from personalized health care.

Transformation of medical information consists mainly of two tasks: translation and enrichment. There are several approaches to perform these tasks with strongly varying degree of computer use ranging from live-transformation during the patient-doctor dialogue to completely automated transformation. While live-transformation provides high quality and customization to patients’ needs, automated transformation delivers fast results with low cost at any time [1]. To combine advantages of both approaches, computer-aided transformation by experts or a crowdsourcing-approach could be used.

Crowdsourcing describes the “act of taking a job traditionally performed by a designated agent and outsourcing it to an undefined, generally large group of people” [19]. In this paper, we apply crowdsourcing for the translation of medical information because crowdsourcing represents a good way to obtain high-quality translations from non-professional translators for relatively low cost [29]. To the best of our knowledge, no concept currently exists that describes problems and requirements for implementation of a crowdsourcing translation system for medical information. We test the applicability of crowdsourcing for transformation of medical information by implementing a web application prototype and conducting a proof-of-concept experiment.

2 Crowdsourcing

Crowdsourcing is derived from outsourcing and was first used by Jeff Howe in an article for ‘Wired’ magazine [18]. In the context of translations, crowdsourcing can be defined as “the process by means of which organizations can tap into the wisdom of their dedicated external community and use the wisdom for their benefit, i.e. with low cost, for more languages, and within the specified time frame” [1]. Besides low cost, a dedicated community represents an additional aspect that makes the definition suitable for the scope of our research. When comparing different translation approaches, crowdsourcing emerges as promising solution. Comparison of human and machine translation leads to the selection of human translation due to significantly better performance in the crucial quality dimension [7]. In contrast to professional translation, the costs are significantly lower when employing a crowdsourcing approach while quality of results remains similar [1, 29]. Since disadvantages of machine and professional translation are avoided, crowdsourcing translation seems to be a promising approach for the developed web application, named ePill crowd (epill.uni-koeln.de/crowd, (in German)).

One of the most famous crowdsourcing systems is Wikipedia. Wikipedia provides various medical information. Rajagopalan et al. [23] compared the quality of cancer information found on Wikipedia with a professionally maintained database. Articles of Wikipedia and the professional database have similar depth and accuracy, but the professionally edited articles appeared to be more readable. On the one hand, this confirms other studies saying that crowdsourcing can create high quality information, but on the other hand it imposes the challenge of generating well readable results. Wikipedia’s concept is a promising approach and is adapted for the development of ePill crowd. ePill crowd was developed as an ancillary web application to ePill, which makes patient information leaflet information available in a web application and offers supplementary services [5, 6].
3 ePill crowd Requirements

Medications are created to fight any attack on a person’s health [2]. Unfortunately, tampering with nature eliminates not only threats but creates also risks [15]. Many laws, regulations, and administrations handle this tightrope walk, affecting all drug related areas. Elements covered in this section are not selected by importance for drugs in general, but by relevance for ePill crowd. For example, the European Medicines Agency (EMA, www.ema.europa.eu/ema) is in charge of making recommendations for drug approvals, following strict regulations, and ensures drug quality, safety, and efficiency. Beyond doubt, this approval process is of high importance. Nevertheless, most laws focusing on the approval process do not fit the scope of our research.

3.1 Content accordance check

Directive 2001/83/EC [9], which addresses medicinal products for human use, is highly relevant for ePill crowd because it explains which indications must appear on the exterior packaging and leaflet of medications. Article 54 of the directive contains a list of 14 particulars, numbered (a) to (n), that must appear on the outer or immediate packaging, for example: Product name – denominations depending on dosage and pharmaceutical form (a). A list of excipients known to have a recognized action or effect (d). Instructions how to use the medicine (n). Although some particulars are only relevant for the physical product, like expiry date (h), most are generally applicable. To be in line with current European norms, inclusion of EU-norm particulars is a requirement for ePill crowd.

Article 59 stipulates that “the package leaflet shall be drawn up in accordance with the summary of the product characteristics”. Summary of products characteristics (SPC) provide information on how to safely and effectively use medical products. SPC content must be approved in an official assessment. SPC provide “the most impartial and complete information about prescription-only medicines” [26]. Albeit SPC are not well suited to convey medication information [2], SPC are the most complete and objective information source a patient can access. Hence, provided information should be in accordance with SPC, but easier to understand. Consequently, accordance with the summary of product characteristics is another requirement.

3.2 Quality Assurance

Transformation is supposed to make information easier to understand. Yet, making medical information available holds substantial risks. In the best case, ePill crowd users’ adherence will increase, which will also lead to improvement of users’ quality of life [11]. Erroneous information, misinterpretation or wrongly derived treatment can have fatal consequences [20]; for instance, a misfortune stuck key that results in a typographic error for Insulin or Valium dose rate – e.g., making it ‘11’ units instead of ‘1’ unit – would most likely cause the death of any user relying on it. To avoid negative consequences or harm to users’ health, best possible precautions have to be taken. Therefore, quality assurance of transformed content is an important requirement.

Since transformations require input data, quality assurance for input data is necessary. The underlying database is derived from ePill [5], which provides basic leaflet information that was obtained from an older version of Gelbe Liste (www.gelbe-liste.de). Hence, for the prototype, we assume that the basic dataset is error-free. Necessity of quality assurance for input data depends on the type of medical information processed. In the context of this paper, it can be solved organizationally through design – by only accepting data from published leaflets or using official data sources for the summary of product characteristics. For other systems that work with personal medical records, for example, personal health records where patients enter their own data, more extensive quality assurances have to
be implemented. Directive 95/46/EC requires that “every reasonable step must be taken to ensure that data which are inaccurate or incomplete […] are erased or rectified” [10].

3.3 Overview

<table>
<thead>
<tr>
<th>Requirement</th>
<th>EU-Norm particulars</th>
<th>The system must provide the particulars required by EU-Norm 2001/83/EC – Article 54.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirement 2</td>
<td>Summary of product characteristics</td>
<td>Provided information must conform to information contained in summary of product characteristics.</td>
</tr>
</tbody>
</table>

**Quality assurance**

| Requirement 3 | Quality assurance of transformed content | Quality of transformed information must be ensured. |

Table 1: Requirements framework

Table 1 summarizes the law- and regulation-based core requirements for ePill crowd. If these core requirements are not fulfilled, it will be dangerous and unethical to make the resulting transformed information publicly accessible and they will most likely fail authorization processes for publishing.

4 System Design

4.1 Concept

ePill crowd is supposed to support the translation process by leveraging crowdsourcing. A translation process has three main elements; a source language, a translation process, and a target language [24]. The ePill crowd process flow is derived from the elements of the translation process and adapted to the identified requirements (Figure 1). First, information (source language) has to be made available. It has to be considered how the information should be stored, how to best present information, and how to reach high accessibility. Second, tools have to be provided to make it as easy as possible for the crowd to transform the information (translation). Third, it has to be assured that transformed content (target language) meets the quality requirements.

*Provide an easily usable platform to display patient information leaflets.*

*Use crowdsourcing to transform and enrich patient information leaflets.*

*Quality assurance of transformation results.*

Figure 1: Process flow of ePill crowd
4.2 Architecture

An initial idea of ePill crowd is already insinuated by the process flow (cf. Figure 1). The rise of crowdsourcing is closely connected with advancing internet technology [8]. The internet perfectly supports crowdsourcing mechanisms. Therefore, ePill crowd will be developed as web application. A web application has many advantages: it can be “easily accessed and consulted at any time, can ease locating of desired information, it can more comfortably point users to further information, it can be updated easily, and it offers a customizable structure as well as interactive content” [5].

![Figure 2: Simplified web application architecture](image)

As shown in Figure 2, basic web application architectures typically consist of clients connected via the internet to a server that has access to a database. Clients send requests via the internet to the application server. The application server extracts necessary data from the database and delivers response information back to the clients via the internet. Large and complex web applications need an appropriate architecture. For instance, load balancing could be necessary to handle the workload of highly frequented applications. Redundant databases could help to assure data security and consistency. Replication of the elements of the host architecture helps to increase applications’ availability and scalability [5].

ePill crowd is built upon the Model-View-Controller (MVC) architectural pattern [3]. An architectural pattern [3] divides software into predefined subsystems, defines responsibilities for each subsystem, and provides guidelines to organize subsystem relationships. For systems that feature human-computer interaction – so called interactive systems, like ePill crowd – the MVC pattern is well-suited [3]. Applications following the MVC pattern are divided into three components as shown in Figure 3. The model contains data and database logic. Views are responsible for displaying information. Program logic and event handling is implemented in the controller. In combination, views and controllers comprise the user interface (UI). The controller acts as intermediary between model and view. For example, if a user opens the web application, the controller queries data from the model, which extracts it from the database, and transfers it to the view where the information is graphically represented and displayed to the user. The MVC pattern makes it easy to develop and maintain an application [5, 21]: The application’s look is independent of the model and controller components, and can be changed without changing underlying data structures or logic. Multiple interfaces can simultaneously be maintained to adapt to specific purposes. Furthermore, the MVC pattern works well with complex system architectures. An application with high workload, multiple interfaces, and large data volume could be spread about many server clusters. Application logic and load balancing can be implemented on servers with high processing performance. Models can be handled on database servers with large storage capacities. Different types of views can be handled on specialized servers offering, for example, data-stream compression for mobile clients or high graphics performance for thin-clients.
5 Implementation

ePill crowd employs the Wikipedia approach because it represents a promising approach for creation of high quality medical information [23]. In addition to the constructs used by Wikipedia, like editable content pages, references, and message boards, two additional constructs are used: Activity and question. Activity refers to actions for changing medication texts, changing text statuses, changing medication discussion text, posting or answering questions, and changing question statuses. Questions can be posted for parts of medication information or the whole medication text. Type or content of questions is not specified, users can ask for general advice or specific advice regarding incomprehensible text parts.

5.1 Provision of medication information

With respect to page layout, it is recommended to use recognizable environments, conform to standard design paradigms [17], and imitate existing components [13]. Since ePill crowd is similar to Wikipedia, ePill crowd offers a similar graphical user interface (cf. Figure 4). A search bar is placed on the center top (A). After typing a couple of letters, possible autocomplete matches are shown. If the entered search string matches a medication exactly, the medication will directly open after submission. For ambiguous search terms, a list of possible matches pops up and medications can be opened by clicking on them. There is a menu on the left side (B), which is grouped by functionality. The first item leads to the start page.

Elements on the start page (Figure 5) are displayed as tiles. This eases possible customization of the start page by removing or adding content. At the top (Start-A), a welcome text is displayed including links to system functions and content, for example, links to manuals, which help to get started. Furthermore, unsolved questions (Start-B), newest questions (Start-C), and newest activities (Start-D) are shown. Users can directly see open questions and proof-read recent changes. Thus, this design helps to foster participation.
The second menu group of (B) gives access to several lists that provide possibilities to get a detailed overview of specific content. In the first list, all medications are alphabetically listed. The second list shows medications for which questions exist. All activities are displayed in the third list. The third menu group of (B) holds pages offering information on ePill crowd itself and assistance with ePill crowd use. First steps are described on the page linked first. Second, a more detailed manual is provided. The last page provides general information about the system, the authors, and the motivation for ePill crowd. In the fourth and last menu group of (B), a chat area can be accessed and feedback can be given. Users can register and log in on the upper right side (C). Registered users can use some additional features. If a text they edited is edited again or if the status of one of their questions is changed, they will receive notifications. Extra menu points are also available. They can view their own questions and activities. If desired, logged in users will get recommendations for medications they could improve through adaptation of SuggestBot’s [4] functionality. To find recommendations, ePill crowd combines the knowledge of which texts a user edited and searches for similar medications using contained substances. If a leaflet is opened, functions and views related to the opened leaflet are reachable over the tabs on top of the content area (D). Leaflet information is provided in the first tab. The second tab holds questions regarding the respective medication and a corresponding discussion section. In the third tab, specific information is provided, that is, SPC can be posted. Activities are listed in the fourth tab, representing the change history of the medication page. The last tab generates a list of medications which are similar with respect to medication group and substances. To navigate leaflet contents, a dropdown menu is always shown on the upper right of the content area (E).

5.2 Transformation of Medication Information

The goal of the transformation is to enrich leaflet information and improve comprehensibility. To foster user contribution, crowd contribution systems must allow for higher degrees of freedom [12, 14]. ePill crowd must be able to handle heterogeneous input and support consolidation of input into a joint artifact. The input must be text. Existing text can be edited and new text can be added. It is, for example, possible to format text by making parts of it bold or structure it using lists. Users can transform text in every possible way within each editable part. To avoid deletion of required particulars, it is not possible to edit the highest structural level of leaflets. If text for required particulars is missing, a message requesting users to contribute missing information will be displayed. Figure 6 shows leaflet text in normal mode (left) and in edit mode (right). To edit text, users click the pen icon (.AppendFormatPen) directly above the text. They can then use the disk icon (SaveChanges), which replaced the pen icon, to save their changes. Users can post questions to each text part or about the medication as a whole by clicking the question mark (AskQuestion). Posted questions are displayed and discussed in the discussion tab. Question status can be switched between answered and unanswered. The question and discussion
feature solves the dilemma that people who could improve text usually already understand the text and do not see the need for improvement. Questions can be seen as an indication for bad comprehensibility and provide a function to solve complicated or very special cases without messing up leaflet information. The activity tab shows all activities for the respective medication, can be used to see previous text versions, and to undo them. Activities can be filtered for changes applied to leaflet information and can even be reduced to one text category. Users who want to contribute are supported by the ‘similar products’ tab and the ‘Give me work!’ menu point. With these features, users can find further products that could spark their interest and lie within their area of expertise.

Figure 6: Leaflet text in normal mode and in edit mode

5.3 Quality Assurance

It is crucial that high quality of provided information is assured. Misinterpretation or wrongly derived treatment can have fatal consequences [20]. In ePill crowd, two functions are implemented to support quality control even further than the Wikipedia approach of having multiple control readers and editors. The control elements for both quality assurance functions are shown in Figure 7.

Figure 7: Elements to operate quality assurance functions

If text is edited, the new version does not replace the original text but is saved separately. Further changes will apply to the separately saved text. Texts can have two states: either verified or edited. Edited text can be changed by everyone and is only subject to quality control of the crowd. Verified versions are former edited versions that were approved by authorized users. If a verified version and an edited version exist, the edited version is displayed. The verified version can always be accessed using the ‘sheet with check mark’ (☑) icon, which also indicates the existence of two versions by being displayed or not displayed. Rights to approve edited versions, which changes the status to verified and triggers replacement of former verified versions, is only given to users with a sophisticated medical background. Approval is the sole function not available for every user and should be used carefully. Systems like ePill crowd should have clear guidelines stating who is allowed to approve edited versions. Such guidelines depend on the context the system is used in. A possible guideline would be to allow approvals only if the edited text was verified by an official agency. In the ePill crowd proof-of-concept experiment only medically educated users were allowed to approve edited versions. The second quality assurance function enables users to annotate text with references (add citation: ☐, remove citation: ☐) to more specific information, which are displayed as tooltip via mouse-over effect (e.g., terms like ‘seldom’ or ‘often’ can be linked to a table that lists actual percentages). In edit mode, marked text can be linked to a corresponding text part of a reference.
References can only be made to text that is provided in the ‘specific information’ tab. Medically educated users can discuss whether referenced text parts are sufficient to prove leaflet information and, if necessary, edit it. Next to each text part, a small colored circle is displayed. The circle changes its color, depending on the percentage of referenced text in the information. Switching between red, yellow and green, the circle uses a traffic light metaphor to signal users the reliability of provided information. This approach could motivate users to apply references in order to reach a green light level. Functionality to differentiate between verified and normal information is the only function not fitting the Wikipedia approach. Given the special system focus, this consistency break is justifiable by the need to always be able to offer verified information.

6 Fulfillment of Requirements

<table>
<thead>
<tr>
<th>Content accordance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Requirement 1 (EU-Norm particulars)</td>
<td>The leaflet’s top-level structure cannot be edited and shows always the required particulars. If no information is available, the particular is shown and the user is asked to contribute the respective information.</td>
</tr>
<tr>
<td>✓ Requirement 2 (Summary of product characteristics)</td>
<td>The reference function supports the process of making leaflet information conform to summary of product characteristics (SPC). SPCs can be contributed for each product. Users are motivated to reference leaflet information to SPC information.</td>
</tr>
</tbody>
</table>

Table 2: Achievement of requirements

As shown in Table 2, the implementation meets all requirements. Requirement 2, conformance to summary of product characteristics, is the only one not completely fulfilled. Even though the system supports conformance to SPC and motivates users to apply it, it is not enforced. This issue could be addressed by allowing information to get approved only if at least 50% of it is referenced. Utility of such restrictions should be carefully addressed in further research in order to avoid implementation of misdirecting incentives. Depending on the operator’s intentions, some implementation decisions could be adjusted, but overall the system is fit for use.

7 Proof-of-Concept

As proof-of-concept, a scenario-based experiment where users had different tasks depending on their role was conducted. System usage data was collected and combined with responses to a post-experiment questionnaire. The experiment was focused on eliciting findings regarding comprehensibility improvements and implications for similar systems. Multiple measurements were employed: usability was measured using the system usability scale (SUS) [25] and motivation was assessed using methods adapted from Nov [22] and Yang and Lai [28]. At the end of the experiment, 16 users were registered in ePill crowd and an email was sent to all users, asking them to fill out an online questionnaire. Finally, 14 completely filled out questionnaires were received (response rate = 87.5%). Except for one questionnaire, all could be mapped to registered users: 5 medically educated users and 9 normal users responded. 6 participants were female (42.8%) and 8 participants were male.
Till Blesik, Tobias Dehling, Ali Sunyaev

The average age was 30; 12 participants were between 23 and 28, and two participants were 55 and 56 years old respectively. All participants were either students with a finished universal degree or trained nurses. Given the small sample size (n=14), findings can only be seen as tendencies.

For usability, a SUS-Score of 74.46% was calculated. Following the categorization of Tullis and Albert [25] issues exist, but it is close to a good system with minor issues. In order to conduct user-based comprehensibility measurement, all edited text parts and their original versions were opposed to each other in the survey. Users were asked about their assessment of comprehensibility improvement. In total, the approval level was rated with 4.57 out of possible 5 points. All users perceived the text’s comprehensibility to be increased. Applied changes, that could explain the increase of perceived comprehensibility, can be grouped in three categories. First, abbreviations were changed to their written-out form. Second, text parts were added to explain medical terms. Third, lists and font-effects were applied to structure text parts. In user-participation based systems, it is important to identify user’s motivation. To measure participation, two indicators were used: (1) number of activities performed by the user and (2) number of logins. Activities and logins were both logged by ePill crowd. Nov [22] used six motivational categories that explain volunteers’ motivation (protective, values, career, social, understanding, and enhancement) and two motivations frequently used in research on open source software development (fun and ideology) to explain users’ motivation for contribution. Nov found fun and ideology to be the top motivations, but only fun correlated significantly with the amount of contributions. In the case of ePill crowd, ‘fun’ was the main motivational factor (Pearson r=0.72, p=0.17) for users with specific medical knowledge and ‘social’ was significantly (Pearson r=0.69, p=0.04) the most important motivational factor for users without medical knowledge. Yet, it is doubtful, if this will be the case in a non-experimental system environment. In addition, computed statistics are only indicative and should be considered with care due to the small sample size. Yang and Lai [28] constructed an integrated motivation model to examine the knowledge sharing behavior of Wikipedians. Their method measures how knowledge sharing behavior is affected by intrinsic motivation, extrinsic motivation, external self-concept and internal self-concept. They found out that internal self-concept based motivation has the highest influence and that, in contrast to Nov’s finding [22], people may initially engage in virtual communities for fun but intrinsic motivation is rarely a dominant motivation for knowledge sharing. In our ePill crowd proof-of-concept experiment, intrinsic motivation factors were the most important for users (Pearson r=0.45, p=0.15). This result contributes to the former finding that ‘fun’ is the most important motivation factor for contribution. Moreover, users felt that sharing knowledge in a system like ePill crowd would improve their professional status. Interestingly, this is the only extrinsic motivation factor with significant and at least moderate correlation (Pearson r=0.58, p=0.03).

8 Conclusion

Requirements for medical information transformation systems can theoretically be fulfilled by a web application. Functions can be implemented to automatically check data or at least support and motivate users to perform checks that cannot be automated. Since all users perceived the text’s comprehensibility to be improved, a crowdsourcing approach is capable of producing easier understandable medical information. We presented a framework of core requirements to create a system that conforms to basic laws and regulations and built a web application based on it. The implemented web application’s concept is derived from the main elements of a translation process, fostering the goal of information transformation. Information system quality goals like availability, scalability or maintainability are satisfied by the used architecture. A web application is accessible
with any browser from any device and synergizes with the crowdsourcing concept. For further research, different aspects of the presented approach can be considered. First of all, usability should be further improved. The identified success factors have to be validated in an evaluation with larger sample size and findings regarding comprehensibility should be assessed with sound methods besides users’ perception. In a broader perspective, the requirement framework’s conformity with local laws and regulations all over the world has to be analyzed. In addition, integration with existing health care infrastructures and pharmaceutical systems can be an interesting subject. Overall, a crowdsourcing-based web application is a suitable approach to transform medical information into an easier understandable form.

9 References


[23] Rajagopalan, MS; Khanna, VK; Leiter, Y; Stott, M; Showalter, TN; Dicker, AP; Lawrence, YR (2011): Patient-Oriented Cancer Information on the Internet: A Comparison of Wikipedia and a Professionally Maintained Database. Journal of Oncology Practice 7, 5, 319–323.


