

# Towards Generating Personalized Hospitalization Summaries

Sabita Acharya, Barbara Di Eugenio, Andrew Boyd, Richard Cameron, Karen Dunn Lopez, Pamela Martyn-Nemeth, Carolyn Dickens, Amer Ardati

University of Illinois at Chicago

Chicago, Illinois

sachar4, bdieugen, boyda, rcameron, kdunn12, pmartyn,  
cdickens, aardati@uic.edu

## Abstract

Most of the health documents, including patient education materials and discharge notes, are usually flooded with medical jargons and contain a lot of generic information about the health issue. In addition, patients are only provided with the doctor’s perspective of what happened to them in the hospital while the care procedure performed by nurses during their entire hospital stay is nowhere included. The main focus of this research is to generate personalized hospital-stay summaries for patients by combining information from physician discharge notes and nursing plan of care. It uses a metric to identify medical concepts that are *Complex*, extracts definitions for the concept from three external knowledge sources, and provides the simplest definition to the patient. It also takes various features of the patient into account, like their concerns and strengths, ability to understand basic health information, level of engagement in taking care of their health, and familiarity with the health issue and personalizes the content of the summaries accordingly. Our evaluation showed that the summaries contain 80% of the medical concepts that are considered as being important by both doctor and nurses. Three patient advisors (i.e individuals who are trained in understanding patient experience extensively) verified the usability of our summaries and mentioned that they would like to get such summaries when they are discharged from hospital.

## 1 Introduction

In the current hospital scenario, when a patient is discharged, s/he is provided with a discharge note along with the patient education materials, which contain more information about the health issue as well as the measures that need to be taken by the patient or the care-taker to continue with the much needed care. However, with statistics showing that

over a third of US adults have difficulty with common health tasks like adhering to medical instructions (Kutner et al., 2006), not many people will be able to understand basic health information and services needed to make appropriate health decisions. More often, patients end up discarding the health documents that are provided to them, either because they get overwhelmed with a lot of information, or because they find it hard to comprehend the medical jargons that such documents are usually flooded with (Choudhry et al., 2016).

Our solution is to generate concise and comprehensible summaries of what happened to a patient in the hospital. Since patients with chronic health conditions such as heart failure (as is our case) need to continue much of the care that is provided by nurses in the hospital even after they are discharged, we integrate the information from both the physician and nursing documents into a summary. We also develop a metric for determining whether a medical concept (a single word or multi-word term) is *Simple* or *Complex* and provide definitions for *Complex* terms. Unlike the “one size fits all” approach that is used for creating health documents, we generate summaries that are personalized according to the patient’s preferences, interests, motivation level, and ability to comprehend health information. In this proposal, we will briefly explain our work on summarizing information and simplifying medical concepts. We will also describe our ongoing efforts on personalizing content and our plans for future evaluations.

## 2 Related Work

Most of the existing approaches on using natural language generation (NLG) for multi-document summarization work only for homogeneous documents (Yang et al., 2015; Banerjee et al., 2015), unlike our case where the physician and nursing

documentation contain different type of content (free text vs concepts). As concerns identifying terms that are *Complex*, some applications assume that all the terms that appear in specific vocabularies or corpora are difficult to understand (Ong et al., 2007; Kandula et al., 2010). These methods are unreliable because none of the currently available vocabularies are exhaustive. For providing explanations to terms that are identified as difficult, Horn et al. (2014) and Biran et al. (2011) use the replacement that was provided to terms from Wikipedia in the Simple Wikipedia parallel corpus. Elhadad (2006) supplements the selected terminologies with definitions obtained from Google “define”. In medical domain, some work has been done in obtaining pairs of medical terms and explanations: Elhadad and Sutaria (2007) prepare pairs of complex medical terms by using a parallel corpus of abstracts of clinical studies and corresponding news stories; Stilo et al. (2013) map medical jargon and everyday language by searching for their occurrence in Wikipedia and Google snippets. Our simplification metric is similar to that of (Shardlow, 2013) but we use five times as many features and a different approach for distinguishing between *Simple* and *Complex* terms. We provide definitions to terms similar to Ramesh et al. (2013), but we are not restricted to single word terms only. Unlike Klavans and Muresan (2000), we refer to multiple knowledge sources for definitions of medical concepts.

There are several existing systems that produce personalized content in biomedical domain (Jimison et al., 1992; DiMarco et al., 1995) as well as in non-medical domains (Paris, 1988; Moraes et al., 2014). However, only a few of the existing biomedical systems generate personalized content for the patients (Buchanan et al., 1995; Williams et al., 2007). PERSIVAL system takes in a natural language query and provides customized summaries of medical literature for patients or doctors (Elhadad et al., 2005). BabyTalk system (Mahmood and Reiter, 2011) generates customized descriptions of patient status for people occupying different roles in Neonatal Intensive Care Unit. However, this system relies on handcrafted ontologies, which are very time intensive to create. Our approach to personalization uses several parameters that determine the content to be included in our summary, similarly to the PERSONAGE system (Mairesse and Walker, 2011), a parameteriz-

Attending: Dr. PHYSICIAN.  
 Admission diagnosis: acute subcortical CVA.  
 Secondary diagnosis: hypertension.  
 Discharge diagnosis: Right sided weakness of unknown etiology.  
 Consultations: Physical Medicine Rehabilitation.  
 Physical Therapy.  
 Occupational Therapy.  
 General Medicine.  
 Nutrition.  
 HPI (per chart): The patient is a AGE y/o AAF with H/o HTN, DM, CVA in past transferred from MRH for evaluation for possible stroke. per patient, she was ne until saturday at around 700 pm, while attempting to go to bathroom, she fell down as her body below chest suddenly gave away.[...]

Figure 1: Portion of the doctor’s discharge note for Patient 149.

Patient Name	: Patient 149	Medical DX	: Acute subcortical cerebrovascular accident
DOB	: 8/8/1940	Allergies	:
POC Date	: 8/7/2014	Code	:
Shift	: 7:00a - 11:00a ( 9:27a)	Status	:
Room #	: 016	Physician	:
MR#	: Unknown	Other	: Hypertension, Diabetes mellitus.
Gender	: UnKnown		

	Current Rating	Expected Rating	NIC Tally	Label Change Status
■ Risk for Ineffective Cerebral Tissue Perfusion				■ A ▶ A
● Risk Control: Cardiovascular Health	2	(4)		● A ▶ A
▲ Medication Management			1	▲ A ▶ A
▲ Medication Administration: Oral			1	▲ A ▶ A

Figure 2: Part of HANDS POC for Patient 149.

able language generator that takes the user’s linguistic style into account and generates restaurant recommendations. To the best of our knowledge, there are no existing systems that generate comprehensible and personalized hospital-stay summaries for patients. Moreover, the combination of the four different factors (patient health literacy, motivation to self-care, strengths and concerns, and the patient’s familiarity with the health issue) that guide our personalization process has not been explored before.

### 3 Dataset

Our dataset consists of the doctor’s discharge note and shift-by-shift update of the nursing care plan for 60 patients. Discharge note is an unstructured plain text document that usually contains details about the patient, along with other information like the diagnosis, findings, medications, and follow-up information. However, no uniform structure of discharge note is known to be followed by all physicians (Doyle, 2011). Figure 1 shows around 5% of our discharge note for Patient 149. On the other hand, nurses record the details in a standardized electric platform called

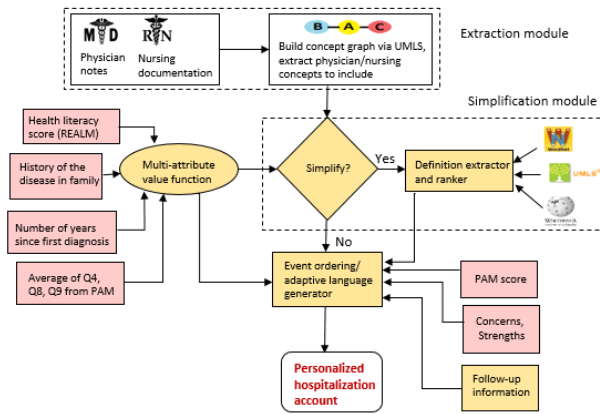


Figure 3: The inputs, output, and workflow of our personalized summary generation system.

HANDS(Keenan et al., 2002), which uses structured nursing taxonomies: *NANDA-I* for nursing diagnosis (Herdman, 2011), *NIC* for nursing intervention (Butcher et al., 2013), and *NOC* for outcomes (Strandell, 2000). Figure 2 shows 15% of a HANDS plan of care (POC) for Patient 149.

## 4 Approach

The workflow of our personalized summary generation system is shown in Figure 3. The *Extraction module* is responsible for extracting concepts from physician and nursing documentation and exploring relationship between them. The functioning of this module is explained in Section 4.1. *Simplification module* distinguishes *Complex* terms from *Simple* terms and provides simple explanations to *Complex* terms. This module is explained in Section 4.2. Most of the remaining components of the workflow play a role in producing personalized content and are explained in Section 4.3.

### 4.1 Exploring relationship between physician and nursing terms

We use MedLEE (Friedman et al., 2004), a medical information extraction tool for extracting medical concepts from the discharge notes and nursing POC. MedLEE maps the concepts to the Unified Medical Language System (UMLS) vocabulary (NIH, 2011). UMLS is a resource that includes more than 3 million concepts from over 200 health and biomedical vocabularies. The knowledge sources provided by UMLS allow us to query about different concepts, their Concept Unique Identifier (CUI), meaning, definitions, along with the relationships between concepts. We begin with

Category	Features
Lexical features	Number of vowels, consonants, prefixes, suffixes, letters, syllables per word
Counts of each POS type	Number of nouns, verbs, adjectives, prepositions, conjunctions, determiners, adverbs, numerals (using Stanford parser)
Vocabulary based	Normalized frequency of the term in Google n-gram corpus, presence of the term in Wordnet
UMLS derived features	Number of semantic types, synonyms, CUIs that are identified for the term; if the term is present in CHV; if the entire term has a CUI; if the semantic type of the term is one of the 16 semantic types from (Ramesh et al., 2013)

Table 1: Features that are extracted for modeling complexity

the nursing concepts (because they are lesser in number as compared to physician concepts) and explore UMLS to identify the physician concepts that are either directly related to the nursing concept or are related through one intermediate node. We restrict ourself to only one intermediate node because going beyond that will lead to reaching up to around 1 million terms in the UMLS graph (Patel et al., 2007), which is not useful for our study. Hence, as shown in Figure 3, the input to this module are all the medical concepts present in the physician and nursing documentation and the output is a list of medical concepts, which comprises of all the concepts from the nursing POC, concepts from the discharge note that are either directly related to a nursing concept or are related through an intermediate concept, and the intermediate concepts themselves. All the concepts that have been explored in this step are candidates for inclusion in our summary. These concepts are then sent to the *Simplification module*, which is briefly described in Section 4.2. For more details on the *Extraction module*, please refer to (Di Eugenio et al., 2014).

### 4.2 Simplification

The *Simplification module* functions in two steps: 1) it determines whether a concept is *Simple* or *Complex*, and 2) it provides the simplest available definition to a *Complex* concept. Since the existing metrics for assessing health literacy (REALM, TOFHLA, NAALS) and reading level (Felsch, Fry Graph, SMOG) work only on sentences and not on terms (that might consist of a single word or multiple words like *arrhythmia*, *heart failure*), we set out to develop a new metric for determining term complexity. Our training dataset consists of 600 terms; 300 of which were randomly

Given: i) term  $T$   
 ii) linear regression function  $LR$  from Section 4.2

- 1) Assign variable  $D=0$
- 2) Extract all the features for  $T$
- 3) Supply feature values to  $LR$  to obtain a score  $Z$
- 4) If  $Z < 0.4$ ,  $D=0$   
 If  $Z > 0.66$ ,  $D=1$   
 If  $Z >= 0.4$  and  $Z <= 0.66$ :  
     If semantic type of  $T$  falls in our shortlisted types,  $D=1$   
     Else  $D=0$
- 5) If  $D=1$ :
  - i) Obtain definitions for  $T$  from UMLS, Wikipedia, Google
  - ii) Extract the medical terms in each definition
  - iii) For each term, repeat step 1, 2, and 3 and get a score.
  - iv) Add scores of the terms in each definition
  - v) Append the definition with the least score to  $T$
  - vi) Return  $T$
 Else return  $T$  with no definition attached

Figure 4: Algorithm for simplifying a term  $T$ .

selected from the Dale-Chall list<sup>1</sup>, while the remaining 300 terms were randomly chosen from our database of 3164 terms that were explored in Section 4.1. We labeled all the terms from the Dale-Chall list as *Simple* and the terms from our database were annotated as *Simple* or *Complex* by two non-native English speaking undergraduate students who have never had any medical conditions (Cohen’s Kappa  $k=0.786$ ). We assume that non-native English speakers without medical conditions are less familiar with any kind of medical term as compared to native English speakers without medical conditions. Disagreements between the annotators were resolved via mutual consultation. The remaining 2564 terms from our database were used as the testing data.

We extracted all the features enlisted in Table 1 for our terms in training and test dataset. We then used a two step approach for developing our metric. In the first step, we performed linear regression on the training dataset with *Complexity* as the dependent variable. This helped us to identify the features that do not contribute to the complexity of a term. It also provided us with a linear regression function (which we will call as  $LR$ ) that includes only the important features. In the second step, we performed clustering on the test dataset, using

<sup>1</sup>This list consists of 3000 terms that are known to be understood by 80% of 4th grade students

You were admitted for acute subcortical cerebrovascular accident. During your hospitalization, you were monitored for chances of ineffective cerebral tissue perfusion, risk for falls, problem in verbal communication and walking. We treated difficulty walking related to nervous system disorder with body mechanics promotion. Mobility as a finding has improved appreciably. We provided treatment for risk for ineffective cerebral tissue perfusion with medication management and medication administration. As a result, risk related to cardio-vascular health has reduced slightly. We worked to improve verbal impairment related to communication impairment with speech therapy.[...] With your nurse and doctors, you learned about disease process and medication.

Follow-up: Can follow-up with General Neurology clinic and Medicine clinic as outpatient if desired.

Figure 5: Part of the summary for Patient 149

the 600 terms from our training dataset as cluster seeds. This process resulted in 3 clusters. Out of the 600 cluster seeds, 70% of those in Cluster1 had *Simple* label; 58% of those in Cluster2 had *Simple* label and 42% had *Complex* label; while 79% of those in Cluster3 had *Complex* label. This indicates the presence of three categories of terms: some that can be identified as *Simple* (Cluster1), some that are *Complex* (Cluster3), and the rest for which there is no clear distinction between *Simple* and *Complex* (Cluster2). For the terms in each of these clusters, we further supplied feature values to  $LR$  and analyzed the corresponding scores. We found that across all clusters, 88% of the terms labeled as *Simple* have scores below 0.4 while 96% of the terms whose score was above 0.66 were labeled *Complex*. For the terms whose score was between 0.4 and 0.66, no clear majority of *Simple* or *Complex* labeled terms was observed in any of the clusters. These observations led to the development of our metric, whose functioning is summarized in Figure 4 and is explained in detail in (Acharya et al., 2016). Hence, the *Simplification module* takes a medical concept as input and determines whether it is *Simple* or *Complex*. For the concepts that are identified as being *Complex*, the simplest definition is extracted from the knowledge sources and is appended to the concept, while the *Simple* concepts are directly sent to the language generator.

### 4.3 Personalized summary generation

#### 4.3.1 Summarizing hospital-stay information

We summarize the information from the discharge note and HANDS POC by using a language gen-

eration approach. We use a Java based API called SimpleNLG (Gatt and Reiter, 2009), which uses the supplied constituents like the subject, verb, object, tense, and produces a grammatically correct sentence. It can also compute inflected forms of the content and can aggregate syntactic constituents like phrases and sentences together. The medical concepts with or without definitions appended to them (i.e the output of the *Simplification module*), along with suitable verbs are our input to SimpleNLG. The group of *NANDA-I*, *NIC*, and *NOC* terms are explained in exactly the same order as they are present in the HANDS note. For each group, we begin by explaining the diagnosis (*NANDA-I*), followed by the treatment that was provided (*NIC* term), and the outcome of the treatment (*NOC*) i.e how effective the intervention was in treating the problem. The *NANDA-I* term is used as the subject of the sentence, followed by the physician terms that are directly connected to it or the intermediate concepts that were extracted while exploring the relationship in Section 4.1. The *NIC* intervention is supplied as the object for the diagnosis and a verb "treat" is used for this purpose. We use the current and expected rating values that are associated with each *NOC* concept and use the percentage improvement to decide on the adverb for the sentence. Some portion of the summary generated for Patient 149 is shown in Figure 5. The terms that are underlined in Figure 5 were determined as being *Complex* by our metric and have a definition appended to them. The definitions can be displayed in different forms (like tool-tip text or footnote) depending upon the medium in which the summary will be presented.

We also provide the follow-up information that was mentioned by the physician in the discharge note, if any. Since the patient follow-up information may appear as a separate section or may be spread across various other sections, we use 67 keywords and 15 regular expressions to algorithmically recognize such information. The last paragraph in Figure 5 shows the follow-up information that was obtained for Patient 149.

#### 4.3.2 Personalizing the summary

So far, we have a reasonable summary that contains the important content from both the physician and nursing documentations. However, our summaries still do not include the patient's perspective. Studies have shown that patients' perspective is essential for patient education (Shapiro,

Dear Patient 149, we are sorry to know that you were admitted for acute subcortical cerebrovascular accident. Cerebrovascular accident is a medical condition in which poor blood flow to the brain results in cell death. Dealing with this issue must have been tough for you, we hope you are feeling much better now.

During your hospitalization, we provided treatment for difficulty walking related to nervous system disorder and risk for ineffective cerebral tissue perfusion. We worked to improve verbal impairment and risk for falls.

We can understand that you have to make changes in your way of living, diet and physical activity as a result of your health condition. You have said that you are concerned about your family and friends. We are very glad to know that you have sources to support you and it is really good that you are working on this. Being committed to solving this problem is so important.

Follow-up: Can follow-up with General Neurology clinic and Medicine clinic as outpatient if desired.

Figure 6: Version of a summary for a patient with low PAM score and low level of health literacy.

1993) and that engaging the patients in their own care reduces hospitalizations and improves the quality of life (Riegel et al., 2011). Our work on personalizing patient summary is motivated by these studies. We expect that including patient-specific information such as social-emotional status, preferences, and needs in a summary will encourage patients to read and understand its content, and will make them more informed and active in understand and improving their health status.

There are four different factors that guide our personalization process: health literacy, patient engagement level, patient's familiarity with the health issue, and their strength/concerns. We also introduce several parameters, whose values depend upon the response given by the patient to these four factors.

**A) Health literacy:** Health literacy is the measure of an individual's ability to gain access to and use information in ways that promote and maintain good health (Nutbeam, 1998). We use the Rapid Estimate of Adult Literacy (REALM) (Davis et al., 1993) test for assessing the health literacy of the patients. REALM consists of 66-itemed word recognition and pronunciation test. Depending upon how correctly a participant pronounces the words in the list, a score is provided. This score tells us whether the health literacy level of the patient is of *third grade or below*, *fourth to sixth grade*, *seventh to eighth grade*, or of *high school level*.

Dear Patient 149, you were admitted for acute subcortical cerebrovascular accident. Cerebrovascular accident is a medical condition in which poor blood flow to the brain results in cell death. During your hospitalization, you were monitored for chances of ineffective cerebral tissue perfusion, risk for falls, problem in verbal communication and walking.

We treated difficulty walking related to nervous system disorder with body mechanics promotion. We provided treatment for risk for ineffective cerebral tissue perfusion with medication management and medication administration. We worked to improve verbal impairment related to communication impairment with speech therapy. We treated risk for falls by managing environment to provide safety.

As a result of these interventions, mobility has improved appreciably. Risk related to cardiovascular health has reduced slightly. On the other hand, communication and fall prevention behavior have improved slightly. With your nurse and doctors, you learned about disease process, medication and fall prevention.

We appreciate your efforts in making changes in your way of living, diet and physical activity for maintaining your health. Keep up the good work. We are very glad to know that you have sources to support you. We hope that you feel better so that you can spend time with your family and friends and return back to your work.

Follow-up: Can follow-up with General Neurology clinic and Medicine clinic as outpatient if desired.

For more information on cerebrovascular accident, please refer to the following website: <https://www.healthline.com/health/cerebrovascular-accident>

Figure 7: Version of a summary for patient with high PAM score and high level of health literacy.

**B) Patient engagement level:** In order to determine how motivated a patient is in taking care of his/her health, we use a metric called Patient Activation Measure (PAM) (Hibbard et al., 2005). PAM consists of 13 questions that can be used to determine the patient's stage of activation. We represent patients at stage 1 or 2 as having *low PAM* and those at stage 3 or 4 as having *high PAM*.

**C) Strengths/concerns of the patient:** We are also interested in identifying the patient's sources of strength and how the disease has affected their lives. For this purpose, we have conducted interviews of 21 patients with heart issues. These interviews are open-ended and the patients are asked to talk about their experiences since they were first diagnosed. We used a pure inductive, grounded theory method for coding the interviews of 9 patients. We found several categories of strength/concern that most of the patients frequently mention, such as: a) priorities in life, b) changes in lifestyle because of the health issue, c)

means of support, and d) ability to cope up with health issues. These topics as well as the possible responses that were collected from the interviews will be phrased as multiple choice questions and will be used for eliciting the strengths and concerns of the patients in real time.

**D) Patient's familiarity with the health issue:** We are also interested in capturing the patient's disease-specific knowledge. After thoroughly analyzing all the patient interviews, we noticed that patients who have either been having the health issue for some time, or have a history of the health issue in the family, use more disease-specific terminologies during their conversation. Based on this observation, we introduced two parameters: *number of years since first diagnosis*, and *history of the health issue in the family*.

Hence, before a patient gets discharged, s/he will take the health literacy test and answer the following: 1)13 questions from PAM, 2) questions regarding their strengths and concerns, and 3) questions that will assess the patient's familiarity with the health issue. We also introduce a parameter called *health proficiency*, whose value depends upon 4 other parameters: *health literacy*, *number of years since first diagnosis*, *history of the health issue in the family*, and *self-efficacy score* (i.e. an average of the scores for questions 4, 8, 9 from PAM). Currently, these 4 constituent parameters are combined in such a way that *health proficiency* can have a value of 1, 2, or 3. We provide maximum weightage to *health literacy*, while the remaining three features are equally weighted. We have developed several rules that take the scores for all the four parameters into account and assign a value to *health proficiency*. Based on the value of *health proficiency*, we make decisions on whether to include more or less details about the medical procedures in the patient summary. Similarly, depending upon whether a patient has *high PAM* or *low PAM*, we decide on whether more or less empathy should be included in the summary.

The phrases that have been used for expressing empathy and encouragement, and the statements for reinforcing patient participation have been derived from the literature on physician-patient and nurse-patient communication (Keller, 1989; Cassell, 1985), as well as some online sources.<sup>2,3,4</sup>

<sup>2</sup>[www.thedoctors.com/KnowledgeCenter/PatientSafety/Appendix-2-Examples-Empathetic-Statements-to-Use](http://www.thedoctors.com/KnowledgeCenter/PatientSafety/Appendix-2-Examples-Empathetic-Statements-to-Use)

<sup>3</sup>[myheartsisters.org/2013/02/24/empathy-101](http://myheartsisters.org/2013/02/24/empathy-101)

<sup>4</sup>[www.kevinmd.com/empathy-patient-interactions.html](http://www.kevinmd.com/empathy-patient-interactions.html)

We have also collected samples of statements from working nursing professionals. Figure 6 shows the personalized version of the summary in Figure 5 for a patient with low health literacy and *low PAM*. Similarly, Figure 7 shows the high health literacy and *high PAM* version of the same summary. As seen from the two samples, the low health literacy version provides information about the health issues of the patient (see second paragraph in Figure 6) and does not include further details about the interventions, while the high literacy version includes details of the health issues, interventions that were done, and the outcomes of the interventions (see second and third paragraph in Figure 7). The patient's response to the questions on their strengths and concerns (as discussed in Section 4.3.2) are described in the third and fourth paragraph of Figure 6 and Figure 7 respectively. For patients with *low PAM* (see Figure 6), we include empathetic phrases like "we are sorry to know that", "we can understand that" and highlight the importance of patient participation with sentences like "Being committed to solving this problem is so important". For *high PAM* patients, we appreciate their efforts in taking care of themselves and include encouraging sentences like "Keep up the good work".

## 5 Evaluation and Results

We have performed two qualitative evaluations of our summaries, where we measured the coverage of medical terminologies, and obtained feedback on the content and organization of information in the personalized summaries.

### 5.1 Coverage of medical terminologies

In order to determine whether our summaries have proper representation of the important information from the physician and nursing notes, we asked a nursing student to read both the physician and nursing notes and generate hand-written summaries for 35 patients. A doctor and a nurse highlighted the important contents from 5 out of the 35 handwritten summaries, which were then compared with the corresponding computer generated versions. This evaluation showed that on average, our summaries contain 80% of the concepts that were considered as important by both the doctor and nurse. Similarly, 70% of the concepts from the entire handwritten summary are present in our automatically generated summaries.

### 5.2 Feedback on the personalized summaries

We asked three patient advisors for their feedback on our attempts for personalization. Patient advisors are a good representative of the patient views because their job role is to communicate with patients first hand. The main aspect that were evaluated are the appearance of our personalized summaries in terms of their feasibility, readability, consistency of style and formatting, and the clarity of the language used.

All the patient advisors liked the personalized summaries as compared to the original ones because they thought that it had a better flow of information. They were able to distinguish between the low literacy and high literacy version of our summaries. All of them said that they would like to get such a summary when they are discharged from the hospital. One interesting thing that we observed is that even within such a small group of evaluators that have almost similar medical knowledge and experiences, we found that there was no uniformity in the sample of summary they preferred, nor were their reasons behind choosing the particular sample alike. This further demonstrates the need for producing personalized summaries, because the preferences and interests of individuals vary from each other.

## 6 Conclusion and Future Work

In this paper, we described our efforts on summarizing information from physician and nursing documentation and simplifying medical terms. We explored different factors that can guide a personalization system for producing adaptive health content for patients. We also proposed a personalization system that can incorporate the beliefs, interests, and preferences of different patients into the same text.

Our next immediate goal is to further improve our personalization algorithm by performing several iterations of evaluations with nurses and patient advisors. We are also in the process of conducting a more thorough analysis of our patient interviews to identify other features that can be useful for improving the content and quality of our personalized summaries. We also plan to conduct an evaluation on a fairly large population so that we can get insights into whether the decisions taken by our algorithm on the kind of content to include/exclude in different situations aligns with that of a more general population.

## References

- Sabita Acharya, Barbara Di Eugenio, Andrew D Boyd, Karen Dunn Lopez, Richard Cameron, and Gail M Keenan. 2016. Generating summaries of hospitalizations: A new metric to assess the complexity of medical terms and their definitions. In *The 9th International Natural Language Generation conference*, page 26.
- Siddhartha Banerjee, Prasenjit Mitra, and Kazunari Sugiyama. 2015. Multi-document abstractive summarization using ilp based multi-sentence compression. In *IJCAI*, pages 1208–1214.
- Or Biran, Samuel Brody, and Noémie Elhadad. 2011. Putting it simply: a context-aware approach to lexical simplification. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 496–501. Association for Computational Linguistics.
- Bruce G Buchanan, Johanna D Moore, Diana E Forsythe, Giuseppe Carenini, Stellan Ohlsson, and Gordon Banks. 1995. An intelligent interactive system for delivering individualized information to patients. *Artificial intelligence in medicine*, 7(2):117–154.
- Howard K Butcher, Gloria M Bulechek, Joanne M McCloskey Dochterman, and Cheryl Wagner. 2013. *Nursing interventions classification (NIC)*. Elsevier Health Sciences.
- Eric J Cassell. 1985. Talking with patients: The theory of doctor-patient communication.(vol. 1).
- Asad J Choudhry, Yaser MK Baghdadi, Amy E Wagie, Elizabeth B Habermann, Stephanie F Heller, Donald H Jenkins, Daniel C Cullinane, and Martin D Zielinski. 2016. Readability of discharge summaries: with what level of information are we dismissing our patients? *The American Journal of Surgery*, 211(3):631–636.
- Terry C Davis, Sandra W Long, Robert H Jackson, EJ Mayeaux, Ronald B George, Peggy W Murphy, and Michael A Crouch. 1993. Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Family medicine*, 25(6):391–395.
- Barbara Di Eugenio, Andrew D Boyd, Camillo Lugaresi, Abhinaya Balasubramanian, Gail M Keenan, Mike Burton, Tamara G Rezende Macieira, Karen Dunn Lopez, Carol Friedman, Jianrong Li, et al. 2014. Patientnarr: Towards generating patient-centric summaries of hospital stays. *INLG 2014*, page 6.
- Chrysanne DiMarco, Graeme Hirst, Leo Wanner, and John Wilkinson. 1995. Healthdoc: Customizing patient information and health education by medical condition and personal characteristics. In *Workshop on Artificial Intelligence in Patient Education*.
- Edward Doyle. 2011. *The wrong stuff: big gaps in discharge communications*.
- Noémie Elhadad. 2006. Comprehending technical texts: predicting and defining unfamiliar terms. In *AMIA*.
- Noemie Elhadad, M-Y Kan, Judith L Klavans, and KR McKeown. 2005. Customization in a unified framework for summarizing medical literature. *Artificial intelligence in medicine*, 33(2):179–198.
- Noemie Elhadad and Komal Sutaria. 2007. Mining a lexicon of technical terms and lay equivalents. In *Proceedings of the Workshop on BioNLP 2007: Biological, Translational, and Clinical Language Processing*, pages 49–56. Association for Computational Linguistics.
- Carol Friedman, Lyudmila Shagina, Yves Lussier, and George Hripcsak. 2004. Automated encoding of clinical documents based on natural language processing. *Journal of the American Medical Informatics Association*, 11(5):392–402.
- Albert Gatt and Ehud Reiter. 2009. Simplenlg: A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation*, pages 90–93. Association for Computational Linguistics.
- T Heather Herdman. 2011. *Nursing Diagnoses 2012-14: Definitions and Classification*. John Wiley & Sons.
- Judith H Hibbard, Eldon R Mahoney, Jean Stockard, and Martin Tusler. 2005. Development and testing of a short form of the patient activation measure. *Health services research*, 40(6p1):1918–1930.
- Colby Horn, Cathryn Manduca, and David Kauchak. 2014. Learning a lexical simplifier using wikipedia. In *ACL (2)*, pages 458–463.
- Holly B Jimison, Lawrence M Fagan, RD Shachter, and Edward H Shortliffe. 1992. Patient-specific explanation in models of chronic disease. *Artificial Intelligence in Medicine*, 4(3):191–205.
- Sasikiran Kandula, Dorothy Curtis, and Qing Zeng-Treitler. 2010. A semantic and syntactic text simplification tool for health content. In *AMIA annual symposium proceedings*, volume 2010, page 366. American Medical Informatics Association.
- Gail M Keenan, Julia R Stocker, Annie T Geo-Thomas, Nandit R Soparkar, Violet H Barkauskas, and Jan L Lee. 2002. The hands project: studying and refining the automated collection of a cross-setting clinical data set. *Computers Informatics Nursing*, 20(3):89–100.
- Mary L Keller. 1989. Interpersonal relationships: Professional communication skills for nurses.



- Judith L Klavans and Smaranda Muresan. 2000. Definder: Rule-based methods for the extraction of medical terminology and their associated definitions from on-line text. In *Proceedings of the AMIA Symposium*, page 1049. American Medical Informatics Association.
- Mark Kutner, Elizabeth Greenburg, Ying Jin, and Christine Paulsen. 2006. The health literacy of america's adults: Results from the 2003 national assessment of adult literacy. nces 2006-483. *National Center for Education Statistics*.
- Saad Mahamood and Ehud Reiter. 2011. Generating affective natural language for parents of neonatal infants. In *Proceedings of the 13th European Workshop on Natural Language Generation*, pages 12–21. Association for Computational Linguistics.
- François Mairesse and Marilyn A Walker. 2011. Controlling user perceptions of linguistic style: Trainable generation of personality traits. *Computational Linguistics*, 37(3):455–488.
- Priscilla Moraes, Kathleen McCoy, and Sandra Carberry. 2014. Adapting graph summaries to the users reading levels. *INLG 2014*, page 64.
- NIH. 2011. Umls quick start guide. Unified Medical Language System, <https://www.nlm.nih.gov/research/umls/quickstart.html>. Last accessed on 10/10/2016.
- D Nutbeam. 1998. Health promotion glossary1. *Health promotion international*, 13(4):349–64.
- Ethel Ong, Jerwin Damay, Gerard Lojico, Kimberly Lu, and Dex Tarantan. 2007. Simplifying text in medical literature. *Journal of Research in Science, Computing and Engineering*, 4(1).
- Cecile L Paris. 1988. Tailoring object descriptions to a user's level of expertise. *Computational Linguistics*, 14(3):64–78.
- Chintan O Patel, James J Cimino, et al. 2007. A scale-free network view of the umls to learn terminology translations. In *Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems*, page 689. IOS Press.
- Balaji Polepalli Ramesh, Thomas K Houston, Cynthia Brandt, Hua Fang, and Hong Yu. 2013. Improving patients' electronic health record comprehension with noteaid. In *MedInfo*, pages 714–718.
- Barbara Riegel, Christopher S Lee, Nancy Albert, Terry Lennie, Misook Chung, Eun Kyeung Song, Brooke Bentley, Seongkum Heo, Linda Worrall-Carter, and Debra K Moser. 2011. From novice to expert: confidence and activity status determine heart failure self-care performance. *Nursing research*, 60(2):132–138.
- Johanna Shapiro. 1993. The use of narrative in the doctor-patient encounter. *Family Systems Medicine*, 11(1):47.
- Matthew Shardlow. 2013. A comparison of techniques to automatically identify complex words. In *ACL (Student Research Workshop)*, pages 103–109.
- Giovanni Stilo, Moreno De Vincenzi, Alberto E Tozzi, and Paola Velardi. 2013. Automated learning of everyday patients' language for medical blogs analytics. In *RANLP*, pages 640–648.
- Corrine Strandell. 2000. Nursing outcomes classification. *Rehabilitation Nursing*, 25(6):236.
- Sandra Williams, Paul Piwek, and Richard Power. 2007. Generating monologue and dialogue to present personalised medical information to patients. In *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pages 167–170. Association for Computational Linguistics.
- Guangbing Yang, Dunwei Wen, Nian-Shing Chen, Erkki Sutinen, et al. 2015. A novel contextual topic model for multi-document summarization. *Expert Systems with Applications*, 42(3):1340–1352.