A NLG framework for user tailoring and profiling in healthcare

Simone Balloccu[1], Steffen Pauws[2], and Ehud Reiter[3]

[1] University of Aberdeen
s13sb9@abdn.ac.uk
[2]
Philips Research, Tilburg University
steffen.pauws@philips.com
[3]
University of Aberdeen
e.reiter@abdn.ac.uk

Abstract. Communication in healthcare can improve therapy adherence and patient engagement. Research into healthcare-oriented Natural Language Generation (NLG) systems suggests that tailoring to user profile can improve overall effectiveness. However lots of systems adopts a single or small group of user profiles, thus overlooking that user subsets may have different needs and that these could evolve over time. In this paper we conceptualize a framework that can produce customised healthcare reports by extracting meaningful data insights and producing a final text which varies in content and terminology according to the user profile and traits. The dietary domain will be used to show a working example.

Keywords: NLG · natural language · data-to-text · healthcare · clustering · user modelling · tailoring · diet coaching

1 Introduction

Communication is of critical importance in healthcare. It is crucial between clinicians as its absence can cause medical errors [1], frustration between team members [2] and ultimately affects patient safety [3]. At the same time communication between specialists and patients has been highly correlated with therapy adherence [4] and better health outcomes [5]. Healthcare communication includes textual data [6] which can be expensive and time consuming to produce manually. Natural Language Generation (NLG) is the software-driven process of transforming data into a text [7] [8], thus offering an alternative to handwriting, allowing the scheduled production of medical reports and reducing manual intervention by clinicians. Many healthcare-related tasks have been addressed by exploiting NLG [9], including report automation, decision support, behavioural change. NLG systems can include a tailoring phase which is defined as changing text characteristics in order to better reach certain individuals, thus delivering customised content [10]. Research has shown that tailoring can potentially improve NLG healthcare system effectiveness. [11]. Some systems achieve tailoring by modeling one ideal user type, which in real-world can be unrealistic. Moreover, even in the case in which different user profiles are exploited, these are typically fixed. We argue that this excludes some interesting dynamics. A user could develop new behaviours or acquire previously absent knowledge which could make another profile more suitable for them. Motivated by the PhilHumans consortium (https://www.philhumans.eu/) goals under which this paper was born, namely the establishment of advanced and innovative patient-health devices interactions, we illustrates the conceptualization of a system that profiles the user based on various characteristics to produce a tailored report. The profiles contains some traits that the user can change in order to always get the most suitable report. The structure of this paper is as follows: section (2) details relevant works in literature, concerning NLG in healthcare; section (3) models the framework and shows its workflow with a simulated example; section (4) discusses future developments.

2 NLG in healthcare

The literature describes a wide range of systems that exploit NLG techniques to address specific healthcare tasks. This section includes some relevant studies, distinguishing those which involve more sophisticated user profiling and those which doesn't.

2.1 Single-user tailored systems

Some NLG systems produces some sort of "general" text that does not involves subjective reader choices in order, content or terminology decisions. typically, such texts inherently assume that a single "ideal user" model can lead to an overall satisfactory text that fits any given reader. Some early works include trauma related reports carberry1997generating and clinicians documentation [12]. Other works investigated pervasiveness and the psychological component of tailoring [13]. There were studies regarding the Neonatal Intensive Care Unit (NICU) system which aimed to extend the audience [14] to include parents [15] via the introduction of an affective element; text quality improvement by domain expert intervention [16] was inspected as well.
These types of systems simplify user profiling, leading to "standardized" reports that cannot be modified according to individual reader needs. In some

cases this makes sense, as some healthcare subdomains follow a standardized and inflexible communication form, hence limiting the usefulness of tailoring. Other domains, such as diet coaching, could instead take advantage of user characteristics in order to produce more focused and engaging text.

2.2 Multiple-user tailored systems

Several works have investigated tailoring reports with a wider user profiling i.e. reports built ad-hoc for specific user groups. Tailoring effectiveness was investigated or conceptualized in various dynamics: mammography recommendation [17]; content enrichment by clinicians [18]; patient clustering for quality of life perspectives text generation [19]; smoking cessation letters [20]; health-risk oriented driving conduct reports [21]. Such systems focus on the user's needs, thus allowing the generated text to be customised accordingly. However we want to point out that the reader preferences could evolve over time, given freshly acquired knowledge about the domain or just natural user interest progression. An example could be a patient which initially focuses on a single aspect of their therapy but then asks further detail.

3 Proposed framework

This section details the proposed system at a conceptual level. In order to clarify the approach functioning a simulated working example will be provided. Behavioural change in nutrition domain will be considered. That is the generation of text reports that inform the reader about the key points of their eating habits in order to encourage the end users to perform a behavioural change in order to improve their diet and lifestyle.

3.1 Architecture

The system architecture consists of three main components:

- **Data Analysis Module:** responsible for processing, aggregating and interpreting the user's raw data in order to obtain a series of insights that will form the report's backbone.
- **User Modeling Module:** responsible for interpreting the user's preferences and needs, in order to obtain a tailored and coherent report structure.
- **Planning and Realisation Module:** terminal part of the architecture that produces the final text, filtered and tailored according to the needs of the reader.

The **Data Analysis Module** (DAM) deals with data reading, computation and interpretation. Although it may contain some semantic information (such as named entities), it doesn't go further and it is not actively involved in the generation of natural language. This allows the system to keep the Data Layer isolated from NLG logic, as this is suggested practice in the reference architecture [22]. **User Modeling Module** (UMM) contributes to document planning and micro-planning. It stores user information/preferences and communicates them to the **Planning and Realisation Module** (PRM), which finalizes document planning, micro planning and performs realisation.

3.2 Data Analysis Module

The first step in generating the report is extracting the data that the user might care about. These data will be filtered later according to its profile, but within this component every possible insight should be extracted. DAM implementation should vary according to each given domain but a mockup is given in the following subsection to clarify the overall logic. The mockup regards diet coaching and takes an existing data source to extract the insights.

DAM for dietary coaching

The source of the data for for the example will be MyfitnessPal [23], a popular calorie counter app which relies on a detailed food database. Once a free account has been created and the app has been filled with data, it can be scraped accordingly, allowing the developer to obtain:

- **calories:** in terms of ideal daily intake, daily/average excess or shortage, days in which the user went closer/further to/from the target
- **nutrients:** in terms of individual target for each nutrient, daily/average, top N problematic nutrients and responsible foods.

It should be kept in mind that DAM implementation should include domain expert contribution. For example evaluating nutrient excess or shortage is not trivial and should be guided by dieticians or nutritionists. The result of this first phase could be a structured file , for example in JSON format, that will then be used by the Planning and Realisation Module to extract the data of interest: we will refer to this as "Insights spec".

DAM simulated output

The following is a simulated example of an Insights spec for an imaginary user, considering calories balance and 5 nutrients (fat, protein, carbohydrates, sugar and sodium). The mockup is provided in JSON form:

```
{
  "calories" : [
    {"perc_shift" :  4},
    {"effects" : ["weight increase", "slow digestion", "sleep disorders"]},
    {"clos_d" : [
      {"day" : "Monday"}, {"shift" : -2}]
    },
    {"far_d" : [
      {"day" :"Friday"}, {"shift" : 30}]
    }],
  "abn_nutrients" : [
    [{"name" : "sugar"},
     {"perc_shift" : 100},
     {"food" : "Coca-Cola"},
     {"anomaly_eff" : ["pale skin", "anxiety", "fatigue"]},
     {"recovery_eff" : ["weight loss", "less dental plaque", "higher perceived energy"]},
     {"suggestions" : "replace sugary drinks with tea or coffee"}],
    [{"name" : "sodium"},
     {"food" : "Pringles chips"},
     {"perc_shift" : 50},
     {"anomaly_eff" : ["nausea", "headache", "seizure"]},
     {"recovery_eff" : ["memory improvement", "less bloating", "lower blood pressure"]}]
     {"suggestions" : "avoid salty snacks and opt for fruits"}]
     ]
}
```

Fig. 1: Simulated Insights spec in diet coaching domain

The file only stores data regarding the patient's eating habits: all other personal information is stored within the UMM. The two top-level elements of the file are `"calories"` and `"abn_nutrients"` which contain information about the caloric and nutritional balance respectively. They both share the key `"perc_shift"` which refers to eventual excess/shortage assumptions (in percentage form). The calorie part of the mockup contains the `"effects"` key and the days in which the user came closer to/further from the ideal value (`clos_d` / `far_d`) along with related percentage. Inside the nutritional section of the JSON there's an array containig the two most abnormal nutrients, that is the two nutrients which presents the most severe deviation from the ideal value. In this case it contains sugar and sodium, with the respective shifts. Moreover for every nutrient there's a set of adverse effects (`anomaly_eff`) like in the previous part, but with additional ones (`recovery_eff`) which are related to the opposite scenario. Finally each nutrient contains the `suggestions` field which provides a way to help fixing the problem. Their use regards the report tailoring and shall be clarified in the next sections. Overall the file shows that there's a slight anomaly in calorie intake (+4%) and a consistent anomaly in terms of sugar (+100%) and sodium (+50%) assumption.

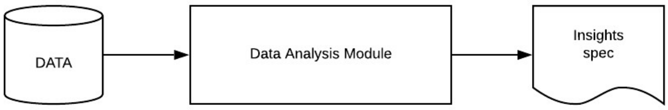

Fig. 2: Conceptual scheme of the Data Analysis Module

3.3 User Modelling Module

Some application domains, such as the dietary one, can contain heterogeneous user sets. Therefore it is unrealistic to expect a standard report structure to always be satisfactory. Different users could be interested in different kinds of nutrients, could care to different degrees about small changes, or report frequency. Furthermore numeracy and literacy can influence the reading capabilities [24] of certain users, thus introducing the need of a simplification strategy. The UMM stores user preferences and communicate them to the PRM:

- **Insights subset:** which insights should be selected.
- **frequency:** report delivery frequency
- **technical level:** whether a simplification should be adopted or not.

The above information should be provided by the user, for example with a form. Hence a data collection phase is necessary in which the user compiles it, so that the system can tailor the text accordingly. The following subsections show a mockup form.

User data collection for dietary coaching
This first section of the form provides the UMM a way to communicate to DAM the data that are necessary to calculate user's calories/nutrients goals. This kind of inter-framework co-operation is not mandatory: for example, in case DAM scrapes MyFitnessPal, these type of data is requested by the app during user registration and is thus readily avaiable.

Physical details

Name: *Paul*

Gender:

- ☑ Male
- ☐ Female

Age: *26*

Height (cm): *183*

Weight (Kg): *86*

Phisycal activity level:

- ☐ Sedentary (Little to no exercise)
- ☑ Light (exercise/sport 1-3 days a week)
- ☐ Moderate (exercise/sport 3-5 days a week)
- ☐ Active (exercise/sport 6-7 days a week)
- ☐ Heavy (very hard exercise / physical job / 2x daily training)

Fig. 3: Physical user data form mockup

User engagement via report tailoring

The following mockup asks for personal traits in order to understand how to customize the text for readibility/engagement optimization.

User experience details

1) In short what are the reasons for you to use this system?

I want to lose weight and gain self-confidence

2) Highest level of education:

Primary School

3) Are you interested in getting informed about your calorie intake?

- ☑ Yes
- ☐ No

4) What kind of nutrients are you interested in getting informed about?

- ☐ Carbohydrates
- ☐ Protein
- ☐ Sugar
- ☐ Fat
- ☐ Sodium
- ☑ All of the above

5) Are you interested in minimal changes in your calories/nutrients intake *(you should consider this option if you're following a really strict diet)*?

- ☐ Yes
- ☑ No

6) Which one of the following sentences do you find to be more understandable?

- ☐ Your calorie assumption is 25% more than what it should be.
- ☑ Your calorie assumption is around a quarter more than what it should be.

7) Are you interested in knowing possible consequences of your eating behaviour (such as adverse effects)?

- ☑ Yes
- ☐ No

8) How frequently would you like to know about your progress?

- ☐ Daily
- ☑ Weekly
- ☐ Monthly
- ☐ Other: every ___ days/weeks

Fig. 4: Experience customization form mockup

The whole questionnaire is designed in order to apply various tailoring techniques [25]. 1) increases engagement by exploiting the *descriptive feedback* mechanism, which inserts user's personal needs into the text. This can be combined with other tailoring strategies the makes the final text less "artificial" and more personal. These includes:

- **Identification**: putting the recipient's name into the text
- **Raising expectation**: emphatising that the report has been written specifically for them
- **Content matching**: inform the user that specific suggestions has been formulated based on their own personal data.

2) is does not involve tailoring but regards a future development that will be better clarified later in the paper. 3) and 4) regards content selection by checking which data are relevant to the user. For example a diabetic patient could be more interested in sugar, while a bodybuilder could want to track proteins. 5) regards slight value anomalies: previous Insights spec contained a 4% calorie excess which is generally neglibigle for the common user but could be important for others. An example could be a bodybuilder preparing for a competition, which typically have to follow really strict eating programs. 6) assesses users numeracy/numerical preferences by letting them choose between a sentence that contains the pure numerical value, and one is converted into a more ("a quarter more") readable information. This allows to obtain this information without explicitly asking users to assess their own numeracy, which could be stressful. 7) addresses those users that could feel frightened by reading adverse effects. 8) is used for report delivery frequency.

UMM sample output

The result of UMM processing should be a structured file that PRM uses to construct the final text. It should contains the user traits and specify them with a proper encoding. Below is a possible UMM output, (JSON format):

```
{
  "username" : "Paul",
  "motivation" : "I want to lose weight and gain self-confidence",
  "calories" : "true",
  "excluded_nutrients" : [],
  "precise_reporting" : "false",
  "low_numeracy" : "true",
  "report_effects" : "true"
  "freq" : "w"
}
```

Fig. 5: Simulated UMM output based on previous form

It can be seen that the JSON doesn't keep any of the information regarding age, sex, weight etc... which are meant to be delivered to the DAM in a similar structure. Therefore only those information that are necessary to tailor the text are kept, in order to be fed to the PRM. From now on the result of the UMM will be called "Tailoring spec".

Changing user profile
As stated before, the user preferences could change over time, either temporarily or permanently. In this case the user must be provided with a tool that allows them to edit his/her choices accordingly. Ideally this could be achieved by storing the user preferences on a database, and presenting them on a related web-page. Here the user could change any given detail in order to update his/her profile and obtain a different report.

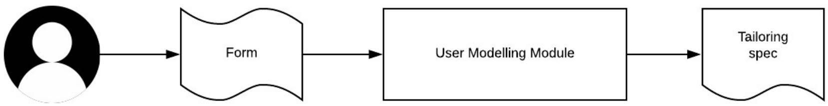

Fig. 6: Conceptual scheme of the User Modelling Module

3.4 Planning and Realisation Module

The last component of the system receives the Insights and Tailoring spec from DAM and UMM. Toghether these two files allow the PRM to filter the insights and re-elaborate them to achieve the ideal text structure. Two sub-components are exploited to improve text quality:

- **Variation:** used to choose from several equivalent frames to express a concept. If this component is used, it is possible to randomize the choice in order to make specific components of the report less predictable, and to obtain a less monotonous reading experience.
- **Quantifier:** responsible for paraphrasing numerical data into verbal concepts to facilitate understanding. Using this component a data like "125%" can be substituted by a more understandable "a quarter more".

The Variation Component works quite simply: the component shuffles between compatible frames. As an example, the following is the mockup output of three possible iterations of the Variation Component:

Hi Paul, this is your weekly diet report.
Greetings Paul, the following is your weekly diet report.
Hi Paul, let's see how you did during this week.

Fig. 7: Three examples of frame variation for a generic greetings message

The Quantifier ideally works with some kind of data structure that maps numerical thresholds to the respective semantic aliases. A couple of implementation requirements arises as well: for each mapping there should be a certain tolerance, hence values such as 155% will still fall as "about a half more". Moreover, the case in which a value falls between two elements but outside of both tolerances must be handled. A possible solution could be to provide a mid-translation, by putting "between" and the two mappings. For example, consider a 63% value that falls between 50% ("half") and 75% ("three quarters") and a 5% tolerance.This results in two ranges [45-55%] and [70-80%] that doesn't include the given value. The translation would be "between a half and three quarter".

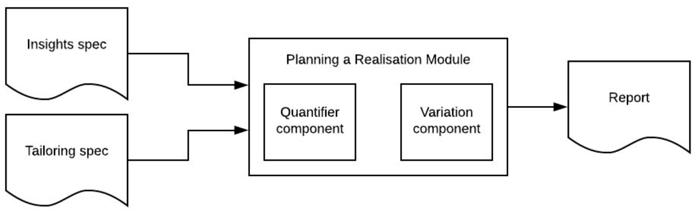

Fig. 8: Conceptual scheme of the Planning and Realisation Module

Final report sample in dietary domain
The following is a mockup showing tailored report. It follows the data and user preferences that were proposed in the previous sections. Tailoring strategies are highlighted into the report. Namely:

- identification
- descriptive feedback
- raising expectation
- content matching

Hi Paul, this is your weekly diet report. You told us that you want to lose weight and gain self confidence . So we wrote this in order to help you out.

You did a great job on calories; you had some problems with sugar and sodium.

Going a bit deeper, Monday was your best day! You calories were about perfect then. Friday gave you some problems, as you ate about a third more than what you should.

Your sugar consumption is around twice more than what it should be, and much of it came from Coca Cola. We know it's hard to change what you eat, but sugar excess is bad for your health and can cause pale skin, anxiety and fatigue. Try having less sugary foods as it would turn in weight loss, less dental plaque and more energy.

Your sodium consumption is around a half more than what it should be. A lot of it came from Pringles chips. Keep in mind that sodium excess can lead to nausea, headache and seizures. Less salty foods means memory improvement, less bloating and lower blood pressure.

Paul, we came up with some suggestions based on your needs . Regarding sugar, you could replace sugary drinks with tea or coffee. Also sodium will definitely be lower if you avoid salty snacks and opt for fruits instead.

Fig. 9: Full report sample

It can be seen how UMM impact on the final results. All the DAM insights were kept, according to user choice. The 4% calorie excess is ignored (as requested) and substituted by "you did a great job on calories", while the side effects of eating behaviours are kept. The numeracy of the user is addressed, as there's no numerical data inside the text. The report adopts a generally positive accent, in line with Affective-NLG (A-NLG) strategies [26]. This is aimed at boosting user confidence, by lightening the weight of wrong behaviours (as in "you had some problems" when dealing with significant nutrient excesses) and reinforcing the positive ones ("you did a great job on calories").

Changing user preferences

For sake of completeness suppose that the user changes its preferences. We'll consider the case in which Paul realises that reading about eating behaviour effects stresses him; moreover he will become very precise and picky, thus deciding that he wants to see the actual numbers and every minimal change. Then he proceed to change his previous choices regarding these characteristics and this results in a change of the Tailoring spec, which structure is now the following:

```
{
  "username" : "Paul",
  "motivation" : "I want to lose weight and gain self-confidence",
  "calories" : "true",
  "excluded_nutrients" : [],
  "precise_reporting" : "true",
  "low_numeracy" : "false",
  "report_effects" : "false"
  "freq" : "w"
}
```

Fig. 10: Updated Tailoring spec based on user change

Overall the structure is almost the same, but the value of the `low_numeracy` key is now "false" in order to disable the Quantifier; `precise_reporting` changed to "true" so that the values are not rounded; `report_effects` is now "false" to avoid frightening the user with potential adverse effects. The following is another mockup that shows how the report would look like:

Hi Paul, the following is your weekly diet report. We know that you want to lose weight and and gain self confidence . So we wrote this in order to help you out.

You did a great job on calories, with just a slight overeating once in a while (4% on average); you had some problems with sugar and sodium.

Going a bit deeper, Monday was your best day! You calories were about perfect then, with just a little shortage (2%). Friday gave you some problems, as you ate about 30% more than what you should.

Your sugar consumption is around 100% more than what it should be, and the main culript seems to be Coca Cola.

Your sodium consumption is around 50% more than what it should be. A lot of it came from Pringles chips.

We know that following a diet is stressful but reducing sugar and sodium will help you reaching your goals.

Paul, we came up with some suggestions based on your needs . Regarding sugar, you could replace sugary drinks with tea or coffee. Also sodium will definitely be lower if you avoid salty snacks and opt for fruits instead.

Fig. 11: Full report sample (2)

All the changes have been addressed, as it can be seen in the mockup. The minimal values are now exposed (4% in average calories and 2% in Monday details). Moreover it can be seen how the adverse effects are completely gone, and just replaced by a generic advice that changing eating habits will "help you reaching your goals". Lastly it can be seen how some frames of the report are different. Here the variation have been used, like "You told us that you want to lose weight..." that became "We know you want to lose weight...". Another example is "We know that changing what you eat..." that now is "We know that following a diet is stressful...". The rest of the report look pretty much identical.

Additional considerations: document planning and realisation
The previous mockups adopt the following fixed order for its elements:

- An introduction that greets the user
- informations regarding the calorie assumption
- informations regarding nutrients assumption, optionally including side and behavioural-change effects.
- suggestions on how to correct wrong behaviours.

The justification for this choice is related to the fact that NLG systesm commonly adopts chronological ordering when dealing with document summarization. In our case there's a lack of text corpus, which means that the informations doesn't have a timestamp to refer to. We argue that, in the case of calorie apps like MyFitnessPal, the app workflow can be compared to the chronologic appearance of a corpus. When the user logs into the app the first information that is showed is the calorie balance, followed by the nutrients details. Hence the reason of the chosen order. The second detail that must be addressed is the text generation technique. The proposed framework is conceptual, hence this detail hasn't been definitely chosen but our guess is that the better choice would be a rule and template-based system. This is because of: the mentioned lack of a corpus on which a machine learning system could be trained; the robustness of templates compared to E2E and other rather new techniques, as works like [27] and [28] evidenced; the characteristic of template and rule-based systems to be fine-tuned to meet specific criteria depending on the domain, which is typically not as easy to achieve with technologies like neurals architectures; the fact that accuracy [4] is of paramount importance in health care.

4 Future Developments

The following are some of the ideas regarding the possible future developments of the proposed framework:

4.1 Clinical team inclusion

User engagement is one of the key points of the system. A possible way to improve it could be to study a possible link between the system itself and the clinical team behind the user's therapy. This would make the approach less "artificial", turning it into a link between patient and specialist and allowing

[4] https://ehudreiter.com/2019/09/26/generated-texts-must-be-accurate/

communication between the parts within the report itself. An example could be a specialist-generated message at the end of a report about a therapy that is progressing positively such as

Nurse Smith is very pleased at your progress, and sends her congratulations.

4.2 Stress monitoring introduction

In the context of diet reports, the idea of monitoring user stress is being considered. This involves two main casistics, which are detailed in the next two subsections.

Stress influence on eating habits
Several studies have shown a possible correlation between stress and various food related issues like food choices [29], metabolism [30] and obesity [31]. Statistical analysis in the NeuroFAST project [5] dataset showed a correlation between stress and coffee assumption [32]; moreover, executing a preliminary analysis on the same dataset (during framework conceptualization) hinted at the possible existence of a stress-eating correlation. Monitoring this factor can be critical in helping the user improve his/her lifestyle.

Stress influence on text comprehension
It is known that stress impacts working memory negatively and therefore hampers reading abilities and text comprehension [33], [34]. This leads to the hypothesis that a stressed user could face reading difficulties, thus leading to the need of a simplified report. Further investigation in such a phenomenon is left for the future.

4.3 Reporting therapy change and progresses

The proposed mockups consider data related to one week only. This is very limiting in an area where the entire system should be focused on the recognition of users' behavioural change. Therefore it is of critical importance to expand the system in this direction, adding suitable methodologies to compare different data series. In this way it could becomes possible to track user progress and focus the content of the report on the most critical ones.

[5] https://neurofast.gu.se/

4.4 Applying the system to other domains

The mockup examples considered the dietary domain, but framework's generality opens up possible developments in other healthcare topics. One of these could be the obstructive sleeping apnea (OSA). In this context, the effectiveness of Continuous Positive Airway Pressure (CPAP) as a treatment has been demonstrated in several studies such as [35], [36], but at the same time it has been observed that patients struggles to meet adherence criteria that would allow them to experience long term benefits. Some effective adherence improvement methodologies are avaiable, which showed significant influence on adherence in 1-3 months therapy period [37], still there's a need to keep patient adherence as high as possible. Moreover, stress have been linked to bad CPAP adherence [38], and this represents a further reason for introducing this parameter into the system. Sleep quality improvement has been already suggested as a well fitting domain for NLG [9] and user-tailored messages showed some promising results in CPAP [39].

4.5 User clustering

UMM mockup contained a question addressing education level that isn't used in tailoring. That information is actually based on the hypothesis that education level could be related to reading preferences of the individuals, and could allow the system to cluster the different users together in order to use a reduced form that asks less information and then tries to obtain the remaining ones with clustering based on common criteria (one of which could be education level).

5 acknowledgement

This research was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 812882.

References

References

1. Yoel Donchin, Daniel Gopher, Miriam Olin, Yehuda Badihi, Michal RNB Biesky, Charles L Sprung, Ruven Pizov, and Shamay Cotev. A look into the nature and causes of human errors in the intensive care unit. *Critical care medicine*, 23(2):294–300, 1995.
2. Lawrence K McKnight, Peter D Stetson, Suzanne Bakken, Christine Curran, and James J Cimino. Perceived information needs and communication difficulties of inpatient physicians and nurses. *Journal of the American Medical Informatics Association*, 9(Supplement_6):S64–S69, 2002.

3. Douglas Brock, Erin Abu-Rish, Chia-Ru Chiu, Dana Hammer, Sharon Wilson, Linda Vorvick, Katherine Blondon, Douglas Schaad, Debra Liner, and Brenda Zierler. Republished: Interprofessional education in team communication: working together to improve patient safety. *Postgraduate medical journal*, 89(1057):642–651, 2013.
4. Kelly B Haskard Zolnierek and M Robin DiMatteo. Physician communication and patient adherence to treatment: a meta-analysis. *Medical care*, 47(8):826, 2009.
5. Moira A Stewart. Effective physician-patient communication and health outcomes: a review. *CMAJ: Canadian Medical Association Journal*, 152(9):1423, 1995.
6. Enrico Coiera. *Guide to health informatics*. CRC press, 2015.
7. Ehud Reiter and Robert Dale. Building natural generation systems. *Studies in Natural Language Processing. Cambridge University Press*, 2000.
8. Albert Gatt and Emiel Krahmer. Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research*, 61:65–170, 2018.
9. Steffen Pauws, Albert Gatt, Emiel Krahmer, and Ehud Reiter. Making effective use of healthcare data using data-to-text technology. In *Data Science for Healthcare*, pages 119–145. Springer, 2019.
10. Rita Kukafka. Tailored health communication. In *Consumer Health Informatics*, pages 22–33. Springer, 2005.
11. Seth M Noar, Christina N Benac, and Melissa S Harris. Does tailoring matter? meta-analytic review of tailored print health behavior change interventions. *Psychological bulletin*, 133(4):673, 2007.
12. KE Campbell, K Wieckert, LM Fagan, and MA Musen. A computer-based tool for generation of progress notes. In *Proceedings of the Annual Symposium on Computer Application in Medical Care*, page 284. American Medical Informatics Association, 1993.
13. Luca Anselma and Alessandro Mazzei. Designing and testing the messages produced by a virtual dietitian. In *Proceedings of the 11th International Conference on Natural Language Generation*, pages 244–253, 2018.
14. Albert Gatt, Francois Portet, Ehud Reiter, Jim Hunter, Saad Mahamood, Wendy Moncur, and Somayajulu Sripada. From data to text in the neonatal intensive care unit: Using nlg technology for decision support and information management. *Ai Communications*, 22(3):153–186, 2009.
15. Saad Mahamood and Ehud Reiter. Generating affective natural language for parents of neonatal infants. In *Proceedings of the 13th European Workshop on Natural Language Generation*, pages 12–21, 2011.
16. Saad Mahamood and Ehud Reiter. Working with clinicians to improve a patient-information nlg system. In *Proceedings of the Seventh International Natural Language Generation Conference*, pages 100–104. Association for Computational Linguistics, 2012.
17. Celette Sugg Skinner, Victor J Strecher, and Harm Hospers. Physicians' recommendations for mammography: do tailored messages make a difference? *American Journal of Public Health*, 84(1):43–49, 1994.
18. Chrysanne DiMarco, HDominic Covvey, D Cowan, V DiCiccio, E Hovy, J Lipa, D Mulholland, et al. The development of a natural language generation sys-

tem for personalized e-health information. In *Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems*, page 2339. IOS Press, 2007.

19. Saar Hommes, Chris van der Lee, Felix Clouth, Jeroen Vermunt, Xander Verbeek, and Emiel Krahmer. A personalized data-to-text support tool for cancer patients. In *Proceedings of the 12th International Conference on Natural Language Generation*, pages 443–452, 2019.
20. Ehud Reiter, Roma Robertson, and Liesl M Osman. Lessons from a failure: Generating tailored smoking cessation letters. *Artificial Intelligence*, 144(1-2):41–58, 2003.
21. Daniel Braun, Ehud Reiter, and Advaith Siddharthan. Saferdrive: An nlg-based behaviour change support system for drivers. *Natural Language Engineering*, 24(4):551–588, 2018.
22. Ehud Reiter. An architecture for data-to-text systems. In *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pages 97–104. Association for Computational Linguistics, 2007.
23. David Rebedew. Myfitnesspal. *Family practice management*, 22(2):31–31, 2015.
24. Sandra Williams and Ehud Reiter. Generating basic skills reports for low-skilled readers. *Natural Language Engineering*, 14(4):495–525, 2008.
25. Robert P Hawkins, Matthew Kreuter, Kenneth Resnicow, Martin Fishbein, and Arie Dijkstra. Understanding tailoring in communicating about health. *Health education research*, 23(3):454–466, 2008.
26. Fiorella De Rosis and Floriana Grasso. Affective natural language generation. In *International Workshop on Affective Interactions*, pages 204–218. Springer, 1999.
27. Ondřej Dušek, Jekaterina Novikova, and Verena Rieser. Findings of the e2e nlg challenge. *arXiv preprint arXiv:1810.01170*, 2018.
28. Charese Smiley, Elnaz Davoodi, Dezhao Song, and Frank Schilder. The e2e nlg challenge: A tale of two systems. In *Proceedings of the 11th International Conference on Natural Language Generation*, pages 472–477, 2018.
29. Jo-Anne Puddephatt, Gregory S Keenan, Amy Fielden, Danielle L Reeves, Jason CG Halford, and Charlotte A Hardman. 'eating to survive': A qualitative analysis of factors influencing food choice and eating behaviour in a food-insecure population. *Appetite*, page 104547, 2019.
30. Friedrich Riffer, Manuel Sprung, Hannah Münch, Elmar Kaiser, Lore Streibl, Kathrin Heneis, and Alexandra Kautzky-Willer. Relationship between psychological stress and metabolism in morbidly obese individuals. *Wiener klinische Wochenschrift*, pages 1–11, 2019.
31. Susan J Torres and Caryl A Nowson. Relationship between stress, eating behavior, and obesity. *Nutrition*, 23(11-12):887–894, 2007.
32. RM Malone, K Giles, NG Maloney, CL Fyfe, A Lorenzo-Arribas, DB O'Connor, and AM Johnstone. Effects of stress and mood on caffeine consumption in shift and non-shift workers. *Proceedings of the Nutrition Society*, 74(OCE1), 2015.
33. Manpreet K Rai, Lester C Loschky, and Richard Jackson Harris. The effects of stress on reading: A comparison of first-language versus intermediate

second-language reading comprehension. *Journal of Educational Psychology*, 107(2):348, 2015.

34. Peng Peng, Marcia Barnes, CuiCui Wang, Wei Wang, Shan Li, H Lee Swanson, William Dardick, and Sha Tao. A meta-analysis on the relation between reading and working memory. *Psychological bulletin*, 144(1):48, 2018.
35. Josep M. Montserrat, Montserrat Ferrer, Lourdes Hernandez, Ramon Farre, Gemma Vilagut, Daniel Navajas, Joan R. Badia, Eva Carrasco, Juan De Pablo, Eugeni Ballester, et al. Effectiveness of cpap treatment in daytime function in sleep apnea syndrome: a randomized controlled study with an optimized placebo. *American Journal of Respiratory and Critical Care Medicine*, 164(4):608–613, 2001.
36. Nigel Mcardle, Graham Devereux, Hassan Heidarnejad, Heather M Engleman, Thomas W Mackay, and Neil J Douglas. Long-term use of cpap therapy for sleep apnea/hypopnea syndrome. *American journal of respiratory and critical care medicine*, 159(4):1108–1114, 1999.
37. W Hardy, J Powers, JG Jasko, C Stitt, G Lotz, and M Aloia. Sleepmapper: a mobile application and website to engage sleep apnea patients in pap therapy and improve adherence to treatment. *Proceedings of SLEEP-14, APSS*, 2014.
38. Gilla K Shapiro and Colin M Shapiro. Factors that influence cpap adherence: an overview. *Sleep and Breathing*, 14(4):323–335, 2010.
39. Lacroix J. Visser T. Den Teuling N. Tatousek, J. Promoting adherence to cpap with tailored education and feedback: A randomized controlled clinical trial. *Proc. Sleep 2015*, 2016.