Personalizing Weekly Diet Reports

Elena Monfroglio and **Luca Anselma** and **Alessandro Mazzei**

Dipartimento di Informatica, Università di Torino

elena.monfroglio@edu.unito.it, luca.anselma@unito.it, alessandro.mazzei@unito.it

Abstract

In this paper we present the main components of a weekly diet report generator (DRG) in natural language. The idea is to produce a text that contains information on the adherence of the dishes eaten during a week to the Mediterranean diet. The system is based on a user model, a database of the dishes eaten during the week and on the automatic computation of the Mediterranean Diet Score. All these sources of information are exploited to produce a highly personalized text. The system has two main goals, related to two different kinds of users: on the one hand, when used by dietitians, the main goal is to highlight the most salient medical information of the patient diet and, on the other hand, when used by end users, the main goal is to educate them toward a Mediterranean style of eating.

1 Introduction

The diet has a huge impact on the health of people, and a number of studies have tried to apply artificial intelligence techniques to this domain. On the one hand, there is a growing interest in applying computational techniques in all the aspects of food production, monitoring, consumption (Min et al., 2019). On the other hand, diet is one of the main factors influencing human health and it has been studied in the field of health informatics (Mazzei et al., 2020; Balloccu et al., 2020).

In the domain of health, it has been shown that one of the main risk factors in the onset of chronic diseases lies in the adoption of an unhealthy diet (Jayedi et al., 2020). Specifically, following a Mediterranean diet provides many health benefits (Godos et al., 2019; Schwingshackl et al., 2017; Galbete et al., 2018). However, following a diet is often difficult both for the specific complexity of the domain, and for the human tendency to transgress on eating behaviors (Anselma et al., 2017). A diet can be seen as a set of quantitative or

qualitative rules and constraints, and technological tools could support users both in keeping track of the historical data and of the user progress, and obtaining motivation by means of their educational and persuasive roles. A virtual dietitian that reasons about eaten meals and that communicates through natural language suggesting corrective actions can be helpful in this task. The Multimedia Application for Diet Management (Anselma and Mazzei, 2015, 2018, 2020) (MADiMan[1]) was born in 2015 in order to build a virtual dietitian that is able to: (i) let the user choose the meal to eat through a mobile application, (ii) analyze the ingredients of the recipe and their quantity through the NLU module, (iii) evaluate the compatibility of the chosen meal with the principles of a diet through the Reasoner module, (iv) determine what the consequences of eating a particular dish are, (v) show these consequences to the user through natural language with messages for educational and informational purposes and motivating users to pursue their goals.

A recent development of MADiMan (Mazzei et al., 2020) concerns the integration of the *Mediterranean diet score* (*Med Score* henceforth) originally proposed in (Stefanadis, 2006). By using a food ontology, MADiMan is able to reason both (i) on macronutrient-based constraints typical of medical diets (e.g., *eat 0.8 g of proteins per kilogram of body weight per day*), and (ii) on food-based constraints typical of Mediterranean diet (e.g., *use daily olive oil in cooking*). The Med Score (0-55) is based on the specific scores (0-5) obtained over the consumption of 11 food categories during a week. Some categories prescribe to eat no more than a limit (e.g., no more than 2 portions of red meat per week), and others not less than a limit (e.g., not less than 5 portions of fish per week).

The MADiMan system includes modules that accompany the user in real-time in the contingent choices of individual meals, but a drawback is the

---

[1] http://di.unito.it/madiman

absence of a summary that allows the user to consolidate the results obtained at the end of the week. Furthermore, the implementation design lacks a proper personalization of the messages, since the NLG module does not take into account any personal data/preferences or the emotional state of the users.

This work has the intent to fulfill this limitation by producing a longer weekly report that educates the user. This automatic report is built with a higher degree of personalization, by formalizing different *user models*, in order to support different types of users who can access the platform and to implement the related communication strategies. Consequently, the information flow analyzed so far is enriched with a long report, which is sent to the user in the form of an e-mail on a weekly basis, in order to represent the habits held in the past week and to suggest which behaviors to encourage for the future and which ones to avoid.

The main research goal of this paper is to evaluate the impact of personalization on the quality of automatically generated weekly diet reports. With this aim, we first describe the main design choices in the DRG system and then we give the results of a preliminary evaluation of the system.

The paper is structured as follows: in Section 2 the concept of user model is introduced and its implementation is described. In Section 3, we describe DRG, a multilingual (Italian/English) generator that follows the typical modules of an automatic NLG system, in relation to the persuasiveness and the educational impact of the generated messages. In Section 4, we provide the results of an initial evaluation of the system and, finally, in Section 5 we conclude the paper describing some ongoing developments.

2 The User Models

In the domain of e-Health, personalization can play a role for achieving some form of engagement (Di-Marco et al., 2007), and user models play a key role in personalizing automatically generated messages. For the diet domain, a user model contains both personal information about the health status (e.g. weight) as well as user's preferences on specific topics. In particular, DRG has been designed by considering two specific categories of users, that are the *patients* and the *dietitians*. The personalization of the messages is based on the different goals that these two kinds of users have. Note that

in the first case the personalization needs to consider just the patient user model, but in the second case the personalization needs to consider both the dietitian (the message addressee) and the patient (the message topic).

On the one hand, the messages generated by DRG for *patients* have to be informative, motivational and educational. The final goal of the system is to educate the patients toward a better understanding of the Mediterranean diet principles using an emotional engaging language based on some psychological heuristics. On the other hand, the messages generated by DRG for *dietitians*, that are medical specialists on nutrition (in some cases physicians), should be as short as possible, should contain information just on bad behavior of the patient, and should use a technical lexicon without emotional content. Note that we decided to not communicate information on the good behavior to the dietitians since we think that in a support system for an expert is more important to produce a summary of the problems. However, if a dietitian prefers otherwise, it is possible to adopt a different policy by changing the DRG configuration.

On the basis of these differences, the patient user model contains: (1) numerical personal/medical information on the user, storing sex, age, weight, height, and BMI (Body Mass Index); (2) a 1-to-4 point scale for representing the stress level based on the DASS-21 questionnaire (Lovibond and Lovibond, 1995); (3) a Boolean variable representing the interest of the user for food sustainability, that is a sort of sensitivity to environmental issues. Using this source of information we can produce a specific personalization for the specific patient. In contrast, with the aim to produce a *technical* message, all the dietitians, for a specific patient, will read the same message.

3 The DRG Architecture

The DRG Architecture (Figure 1) follows the standard modular architecture of symbolic NLG (Reiter and Dale, 2000; Reiter, 2007). The generation flow starts from numerical data representing the weekly diet of a patient. The diet reasoner, a module of the MADiMan system, produces and stores in a relational database the information regarding the dishes eaten during a week, their recipes, their nutritional values and their Med Scores. Also the user model information of the various users are stored, in the same relational database.
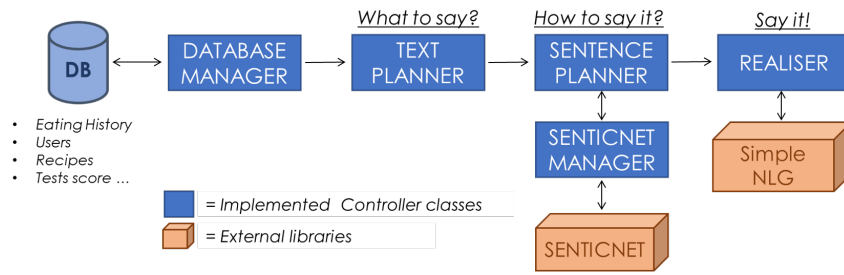
Figure 1: The DRG architecture.

Following (Reiter, 2007), we divided the generation process into three macro-phases implementing the specific generation tasks. The *text planning* phase implements the selection of information units to communicate (content determination) and the order in which they appear (text structuring). The creation of these information units follows the idea to aggregate together semantically equivalent information. The different categories of food are aggregated based on their scores over four different possible values: *very good*, *good*, *bad* and *very bad*. Text structuring decides in which order the information should be presented: following the so-called *sandwich technique*, we communicate the units following the *very good, bad, good, very bad* order, that is alternating a positive and a negative communication. In Figure 2 we report an example of text plan. Note that the text plan contains the Med Score value, computed by the reasoner, rather than the frequency of consumption of each food category.

The *sentence planning* phase is responsible for building the syntactic structures of the messages. So, starting from the sequence of information unit produced in the text planning, a rule-based sentence planner decides both the syntax and the lexical items of the sentences. We defined a fixed schema based on a sequence of ten elements: (a) greetings, (b) Med Score, (c) encouragement, (d) very good score, (e) bad score, (f) good score, (g) very bad score, (h) best and worst dish of the week, (i) environmental impact, (j) educational notion on the Mediterranean diet. For each element (a-j), the sentence planner will use a specific quasi-tree, that is a sort of unordered and unlexicalized dependency tree (Anselma and Mazzei, 2020). The quasi-tree will be instantiated, producing a complete structure ready for realization, considering both the text plan and the user model. For instance, greetings (a) depend on the age, whilst the best/worst dishes (h), as well as the environmental impact (i), are not provided for dietitians. Moreover, for patients with a high level of stress the system does not provide information on the "very bad" category in order to not exacerbate their stress. For instance, in Figure 3 a sentence plan generated for dietitians is presented.

```
{"language":"English",
 "user name":"Giulia",
 "user age":44,
 "user gender":"F",
 "user stress":0,
 "domain knowledge":0,
 "Med score":31,
 "last Med score":26,
 "very good":[{"name":"cer","score":5},
              {"name":"veg","score":5},
              {"name":"fish","score":5},
              {"name":"oil","score":5}],
     "good":[{"name":"pot","score":3},
             {"name":"leg","score":3},
             {"name":"poul","score":3}],
      "bad":[{"name":"fru","score":2}],
 "very bad":[{"name":"rmeat","score":0},
             {"name":"dairy","score":0}],
 "best dish":206,
 "worst dish":288,
 "total environment score":1000.82,
 "bad category environment":"rmeat"}
```
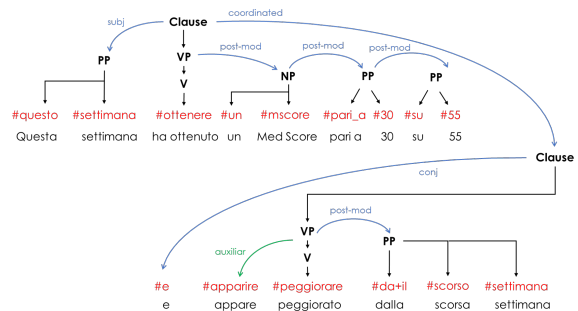
Figure 2: An example of text plan.

Figure 3: A sentence plan for the sentence "This week he has got a Med Score of 30 out of 55 and it seems to have gotten worse since last week." (translation from Italian).

The final process of the pipeline is the realization phase that accounts for function words insertion and inflection. Following the previous implementation of MADiMAn and, in order to build a bilingual Italian/English generator, for this phase we used the Italian porting (Mazzei et al., 2016) of the SimpleNLG realizer (Gatt and Reiter, 2009).

In Section 3.1, we give some details on the lexicalization that personalizes the sentences on the basis of some emotions.

3.1 Using SenticNet for lexicalization

In the field of NLG a number of works consider the use of affective strategies for the realization of an emotionally engaging text (de Rosis and Grasso, 2000; Mahamood and Reiter, 2011).

To give different emotional nuances to the final messages, we decided to use an emotional lexicon. SenticNet is a multilingual knowledge base designed for text sentiment analysis and provides a list of 150 000 lemmata, each one with different types of information, including primary and secondary emotions. Crucially, we used SenticNet for associating the stress level contained in the user model with emotion types. In SenticNet the lemmata are associated with emotions via type and polarity as described in the *Hourglass of Emotions* (Susanto et al., 2020) model. Specifically, the emotions are classified in four categories (introspection, temper, attitude and sensitivity), each one with six different polarity levels. Our idea is to associate with each stress level a specific type of emotion to mitigate the stress. We stipulate that, in correspondence to the stress levels, the types of emotions will be selected in this specific ascending order: sensitivity, introspection, attitude and temper. In the case of more lemmata with a same type of emotion, the *SenticNetManager* algorithm will prefer the lemma with highest polarity. In this way, we constrain DRG to select the least negative term. Thus, we built a DRG emotional lexicon for English and Italian by intersecting the original SimpleNLG lexicon with the SenticNet lexicon. Moreover, for each leaf of the quasi-trees, we defined a specific *synset* of words belonging to the DRG emotional lexicon. In this way, the *SenticNetManager* will choose among the words in the synset the best one in correspondence to a specific user stress level. For instance, let us suppose that a synset of a quasi-tree contains three lemmas: *choice*, *idea* or *decision*. On the one hand, in SenticNet *choice* and *idea* correspond both to the same emotion type, that is temper, that will be selected by *SenticNetManager* in the case of high stress; since *choice* has a higher polarity value will be preferred over *idea*. On the other hand, *decision* is related to introspection and it will be selected in case the stress is medium-low.

4 Initial Evaluation of DRG

We are aware that message personalization does not always correspond to an effective improvement for the end user (Reiter et al., 2003). So, in order to evaluate DRG, we performed two different evaluations. A first preliminary evaluation consisted in submitting a number of Italian and English texts generated by DRG to an adjunct professor of dietistic (henceforth, the *expert*). The evaluation was set up by generating ten different reports simulating the diet of ten patients. These simulations consist in randomly selected dishes from a database of recipes recovered from well-known web sites (e.g. BBC Food). For eight simulations, DRG generated a report personalized for the patient, and for two simulations DRG generated a report personalized for dietitians. By considering the specific user for which the text is generated, the expert had to evaluate a report in terms of: (i) readability, that consists in the linguistic quality of the report, (ii) accuracy or content quality and (iii) usefulness, that is the effective educational support that the system could provide to the patient. The general feedback of this first evaluation was positive, with a good level for all the three measurements. However, the expert suggested to improve the system in three directions: (i) to integrate the Med Score with information about macro/micronutrients (i.e. cholesterol, proteins, etc.); (ii) to provide more details about the ingredients; (iii) to enhance the personalization considering the patient's BMI.

A second preliminary and still ongoing evaluation was performed only for Italian language to have the patients' feedback. Similarly to the first evaluation, DRG generated four texts for four different patients, on the basis of a simulation consisting of randomly selected dishes. Moreover, we built a baseline text by simply listing all the information contained in the text plan (cf. Figure 2). In Table 1 we report an example of text generated by DRG and the corresponding baseline text.

The evaluation was set up in the form of an online form with the testing hypothesis that the users would prefer highly personalized report over the

	Italian version	English version
DRG	Ciao, Davide. Questa settimana hai ottenuto un Med Score pari a 23 su 50 e, inoltre, sei peggiorato dalla scorsa settimana. Non mollare! La quantità di patate, pesce ed olio era quasi eccellente. E inoltre hai fatto un lavoro fantastico con cereali e verdura. La merenda di venerdì era una scelta eccellente perché il piatto King Ranch Chicken Casserole ha una buona quantità di cereali e verdura. Gli esperti sconsiglierebbero il piatto Creamy Au Gratin Potatoes che hai mangiato il lunedì scorso a colazione perché la quantità di latte e derivati non è buona. Ricorda: una pessima dieta uccide più del fumo.	Hi Davide. This week you got a Med Score of 23 out of 50 and, furthermore, you have not improved since last week. Do not give up! The amount of potatoes, fish and oil was almost excellent. Furthermore, you've done a fantastic job with cereal and vegetables. Friday's snack was an excellent choice because the King Ranch Chicken Casserole dish has a good amount of grains and vegetables. Experts would advise against the Creamy Au Gratin Potatoes dish you ate for breakfast last Monday because the amount of milk and derivatives is not good. Remember: a bad diet kills more than smoking.
Baseline	Questa settimana hai ottenuto i seguenti punteggi: - Med Score: 23 su 50 - Carne rossa, latticini e carne bianca: 0 su 5 - Legumi e frutta: 1 su 5 - Pesce: 3 su 5 - Olio e patate: 4 su 5 - Cereali e verdura: 5 su 5 - Migliore piatto: King Ranch Chicken Casserole - Peggior piatto: Creamy Au Gratin Potatoes	This week you obtained the following scores: - Med Score: 23 out of 50 - Red meat, diary and poultry: 0 out of 5 - Legumes e fruit: 1 out of 5 - Fish: 3 out of 5 - Oil and potatoes: 4 out of 5 - Cereal and vegetables: 5 out of 5 - Best dish: King Ranch Chicken Casserole - Worst dish: Creamy Au Gratin Potatoes

Table 1: The text generated by DRG and the corresponding baseline text used for the evaluation. The Italian version is on the left and the English version, not used for evaluation, is on the right.

baseline. The form presents a user description, the baseline text and the DRG text (using a Latin square arrangement), and asks to evaluate the readability, the accuracy and the usefulness of each text by means of a 7-point Likert scale. Four pairs of reports along with a user description are shown. The four different cases were constructed by varying both the weekly dishes (randomly extracted) and the type of user for whom the report is generated. Currently, only five testers participated to the second evaluation, as reported in Table 2. We are aware that the small number of testers cannot guarantee statistically significant results. However, we can speculate that the readability score confirms the appealing of personalization in the linguistic quality of the text.

5 Conclusions and Ongoing Work

In this short paper we presented the main properties of DRG, that is a symbolic natural language generation system for building weekly report on Mediterranean diet. We are still evaluating our sys-

tem by using the procedures described. Moreover, in order to have a more realistic and significant feedback on DRG, we are going to involve some students in dietistic in a form-based evaluation.

As a future work, we intend to design a more complete comparative evaluation based on an ablation strategy. We intend to generate different versions of the report by excluding/exploiting the various components of DRG. In particular, we want to evaluate the contribution of the emotional lexicon in a A/B test.

References

Luca Anselma and Alessandro Mazzei. 2015. Towards diet management with automatic reasoning and persuasive natural language generation. In *Progress in Artificial Intelligence - 17th Portuguese Conference on Artificial Intelligence, EPIA 2015, Coimbra, Portugal, September 8-11, 2015. Proceedings*, number 9273 in Lecture Notes in Computer Science, pages 79–90. Springer. ISBN 978-3-319-23484-7.

Luca Anselma and Alessandro Mazzei. 2018. Designing and testing the messages produced by a virtual dietitian. In *Proc. of 11th International Conference on Natural Language Generation (INLG 2018)*, pages 244–253. ACL.

Luca Anselma and Alessandro Mazzei. 2020. Building a Persuasive Virtual Dietitian. *INFORMATICS*, 7(3):1–27.

	Readability	*Accuracy*	*Usefulness*
DRG	5.65	5.3	5.45
Baseline	5.45	5.75	5.55

Table 2: Preliminary evaluation results (average over 7-point Likert scale).

Luca Anselma, Alessandro Mazzei, and Franco De Michieli. 2017. An artificial intelligence framework for compensating transgressions and its application to diet management. *Journal of Biomedical Informatics*, 68:58–70.

Simone Balloccu, Steffen Pauws, and Ehud Reiter. 2020. A NLG framework for user tailoring and profiling in healthcare. In *Proceedings of the First Workshop on Smart Personal Health Interfaces co-located with 25th International Conference on Intelligent User Interfaces (IUI 2020)*, volume 2596, pages 13 – 32. CEUR-WS.

Fiorella de Rosis and Floriana Grasso. 2000. *Affective Natural Language Generation*, pages 204–218. Springer Berlin Heidelberg, Berlin, Heidelberg.

Chrysanne DiMarco, H. D. Covvey, Peter Bray, Donald Cowan, Vic DiCiccio, Eduard Hovy, Joan Lipa, and Doug Mulholland. 2007. The development of a natural language generation system for personalized e-health information. In *Proceedings of the 12th World Congress on Health (Medical) Informatics*, volume 129, pages 2339–2340. IOS Press.

Cecilia Galbete, Lukas Schwingshackl, Carolina Schwedhelm, Heiner Boeing, and Matthias B Schulze. 2018. Evaluating Mediterranean diet and risk of chronic disease in cohort studies: an umbrella review of meta-analyses. *European journal of epidemiology*, 33(10):909–931.

Albert Gatt and Ehud Reiter. 2009. SimpleNLG: A Realisation Engine for Practical Applications. In *Proceedings of the 12th European Workshop on Natural Language Generation*, ENLG '09, pages 90–93, Stroudsburg, PA, USA. Association for Computational Linguistics.

Justyna Godos, Raffaele Ferri, Filippo Caraci, Filomena Irene Ilaria Cosentino, Sabrina Castellano, Fabio Galvano, and Giuseppe Grosso. 2019. Adherence to the Mediterranean diet is associated with better sleep quality in italian adults. *Nutrients*, 11(5):976.

Ahmad Jayedi, Sepideh Soltani, Anna Abdolshahi, and Sakineh Shab-Bidar. 2020. Healthy and unhealthy dietary patterns and the risk of chronic disease: an umbrella review of meta-analyses of prospective cohort studies. *British Journal of Nutrition*, 124(11):1133–1144.

P.F. Lovibond and S.H. Lovibond. 1995. The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33(3):335–343.

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.

Alessandro Mazzei, Cristina Battaglino, and Cristina Bosco. 2016. SimpleNLG-IT: adapting SimpleNLG to Italian. In *Proceedings of the 9th International Natural Language Generation conference*, pages 184–192, Edinburgh, UK. Association for Computational Linguistics.

Alessandro Mazzei, Antonio Lieto, Mirko Di Lascio, and Luca Anselma. 2020. Adopting the mediterranean diet score in a diet management system. In *13th International Joint Conference on Biomedical Engineering Systems and Technologies-HEALTHINF*, pages 670–676. SCITEPress.

Weiqing Min, Shuqiang Jiang, Linhu Liu, Yong Rui, and Ramesh Jain. 2019. A survey on food computing. *ACM Comput. Surv.*, 52(5).

E. Reiter, R. Robertson, and L. Osman. 2003. Lessons from a Failure: Generating Tailored Smoking Cessation Letters. *Artificial Intelligence*, 144:41–58.

Ehud Reiter. 2007. An architecture for data-to-text systems. In *Proc. of the 11th European Workshop on Natural Language Generation*, ENLG '07, pages 97–104, Stroudsburg, PA, USA. Association for Computational Linguistics.

Ehud Reiter and Robert Dale. 2000. *Building Natural Language Generation Systems*. Cambridge University Press, New York, NY, USA.

Lukas Schwingshackl, Carolina Schwedhelm, Cecilia Galbete, and Georg Hoffmann. 2017. Adherence to mediterranean diet and risk of cancer: an updated systematic review and meta-analysis. *Nutrients*, 9(10):1063.

Demosthenes B. Panagiotakos, Christos Pitsavos, Christodoulos Stefanadis. 2006. Dietary patterns: A Mediterranean diet score and its relation to clinical and biological markers of cardiovascular disease risk. *Nutrition, Metabolism and Cardiovascular Diseases*, 16(8):559–568.

Yosephine Susanto, Andrew G Livingstone, Bee Chin Ng, and Erik Cambria. 2020. The hourglass model revisited. *IEEE Intelligent Systems*, 35(5):96–102.