CS 491 Senior Design Project

Project Specification Document

CONTENTA

GROUP #T2325

Ömer Oktay Gültekin | 21901413
Oğuz Kuyucu | 21902683
Barış Tan Ünal | 22003617
Mert Ünlü | 22003747
Alperen Utku Yalçın | 22002187

Supervisor: Fazlı Can
Innovation Expert: Tağmaç Topal
Instructors: Mert Bıçakçı & Atakan Erdem
16.11.2023

ABSTRACT

This report outlines the specifications for the development of an innovative mobile application focused on enhancing consumer awareness and safety in the realm of packaged food consumption. The application utilizes cutting-edge image recognition technology, allowing users to scan the ingredients section of food products through their mobile phone cameras. By employing advanced algorithms, the application identifies potentially harmful substances, including allergens such as aspartame, monosodium glutamate, or E621. To personalize the user experience, individuals can input information about their allergies, height, and weight, enabling the application to pinpoint allergenic ingredients and provide insights into potential health complications.

To ensure accuracy in ingredient detection, the application integrates language processing algorithms, addressing challenges such as misinterpretation caused by wrinkles or folds on packaging. Initially designed for Turkish and/or English, the application is poised for expansion to cater to a global audience through the incorporation of a seamless translation API. In the later stages, the application has the potential to broaden the scope of products to cosmetics or cleaning goods, too. This comprehensive report details the technical specifications and features essential for the successful development and deployment of this innovative mobile application, advancing the landscape of informed and health-conscious food choices for users worldwide.

Keywords: packaged food, cosmetic products, human health, ingredient analysis, image-processing, language processing algorithms.

TABLE OF CONTENTS

LIST OF FIGURES

1.0 INTRODUCTION

People buy products on a daily basis from many platforms and the number of products within the market increases whenever a new item is introduced to the public. When buying these products customers have to be careful not to consume products that might be affecting their health negatively and filter any kinds of content they may not desire to consume by reading the ingredients section of the product package. Even if the customer pays attention and reads the section, the process of understanding the ingredients on the flow is particularly difficult for the customers if they have little knowledge of the ingredient itself, its background, and its potential risks. This problem is evident in cases where people who are intolerant to specific ingredients have to avoid consuming some products [1], or in cases where customers wish to avoid the government-approved but potentially harmful content within the products.

Fig. 1: Rear Face of a Cranberry Juice Bottle [2].

An example is “Acesulfame Potassium”, which is an ingredient used as a sweetener in some food products (such as the Cranberry Juice seen in Figure 1) but it has concerns of cancer as well as risks to pregnant people, so it is understandable that some people may wish to not consume such an ingredient [2][3]. Our mission is to make users aware of the contents of their food and make sure they consume what is right for them. Our product app *Contenta* aims to inform the customers quickly, and efficiently while aiming to aid their day-to-day product consumption with transparency on what they consume.

Fig. 2: Nutrition Facts and Ingredients of a Snack Bar [4].

1.1 DESCRIPTION

Contenta is a mobile application that enables users to scan the ingredients of packaged products and gives information about the substances within the products. The application aims to read, detect, and list the ingredients that are written on the packages in an easily readable format and warn or give information to people about their concerns, risks, and backgrounds; as well as help them follow along with any diets that exclude the diet's unfavorable ingredients. The healthcare app aims to produce fast, accurate results to aid in increasing customer healthcare awareness and speed up the shopping experience for people who wish to know what is within the products they consume.

1.1.1 USER PROFILE

The users first download the app and when they open the app for the first time, they are greeted with a quick explanation of how to use the application. The users then create their profile by giving information about their allergies, their possible health issues, whether there is someone pregnant in the household, or whether they wish to not consume a specific ingredient, etc. Furthermore, the users can choose to follow along with any diet plans or food preferences and can cancel at any time. The app will advise the users whether it is a good idea or a bad idea to consume a product based on the user profile preferences.

1.1.2 SCANNING THE INGREDIENTS

When the app is opened, there will be several different options on the screen, and there will be a button “Scan Ingredients” that opens the camera with a click. The screen will display the back camera’s vision and the user will pick up the product they wish to scan, and take a photo of the ingredients section of this product. Then, the information will be processed and on the screen, there will appear the contents and their explanations. Each content will be listed according to its safety level and categorized with a color indication (red for potentially unsafe, yellow for neutral, green for healthy to consume). The users can click on the listed ingredients to view further information about how the ingredient is produced, what it contains, what risks it has, what health benefits it has, and some articles about the ingredient. Furthermore, depending on the preferences of the user, such as their allergies and their pre-set diet plan, some of the ingredients may be advised or unadvised.

1.1.3 SCANNING THE NUTRITION FACTS

The users also are able to scan the ingredients of the nutrition facts, much like how they scan the ingredients. The nutrition values of the food items are then evaluated by the app and whether the food is healthy or unhealthy according to your preferences is displayed. According to their calorie intake information, users may receive further feedback and information on the food products, if they wish to.

1.1.4 WRITING / READING INFORMATIVE BLOGS

As the awareness about the contents of packaged goods increases among healthy customers, the number of blogs posted about food ingredients and their possible risks is also increasing. Since there is already a demand for acquiring knowledge from experts about the details of customer food consumption, our app also gives users access to such informative helpful blogs. In our blogs section, experts can post their ideas, preferences, advice, and details about products/ ingredients. Customers can easily open up the blogs of these writers and get more details, as well as gain perspectives on their shopping experience by deciding for themselves if they should consume a specific product.

1.2 HIGH-LEVEL SYSTEM ARCHITECTURE & COMPONENTS OF THE PROPOSED SOLUTION

Contenta app architecture comprises a front-end interface for users to initiate ingredient scanning procedures through the device's camera. After scanning, the app triggers an on-device parsing algorithm to extract and categorize ingredients from the product label. This image processing algorithm is running on the device, not on the cloud. This parsed data is then cross-referenced with an

embedded database containing information on hazardous ingredients. The app leverages the device's computational power, minimizing reliance on external servers and ensuring real-time processing for a better user experience.

The system architecture encompasses a user interface layer for seamless interaction, an on-device parsing module for ingredient extraction, and a local database for storing information about dangerous products. The app's self-sufficiency in processing ensures data privacy and minimizes latency, providing users with a quick and reliable means of identifying potentially harmful ingredients in scanned products.

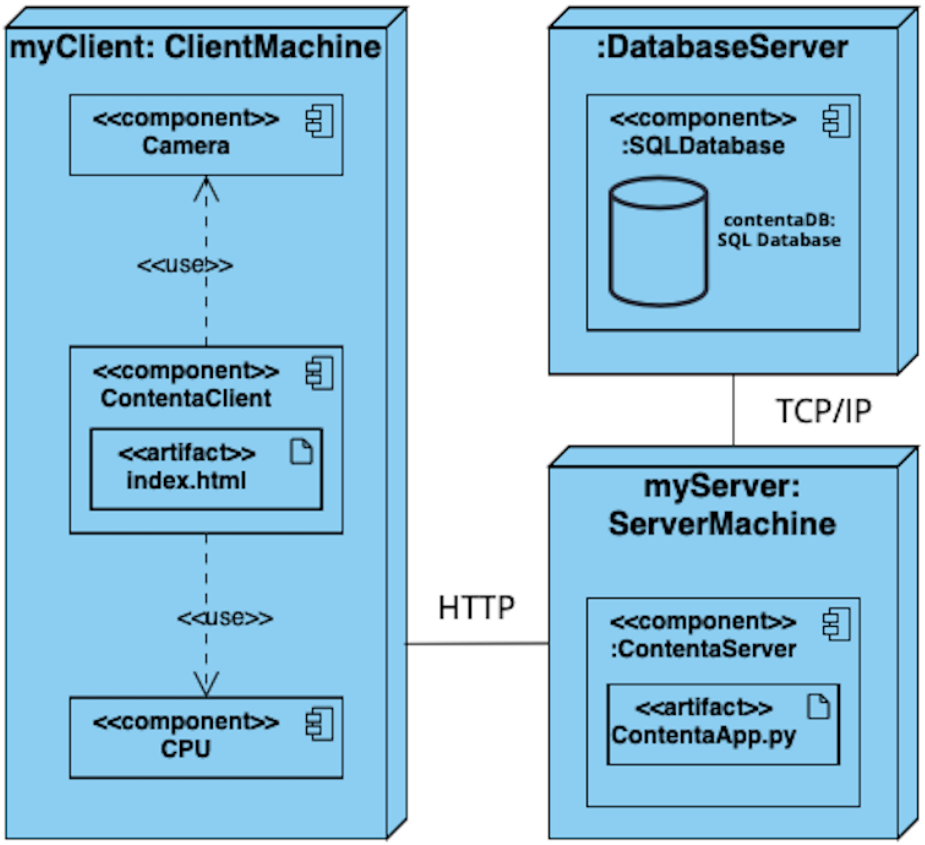

Fig. 3: Deployment diagram illustrating the components of the high-level system.

As one can see from the Deployment Diagram above, there are roughly three main units in our high-level system: client, server, and database. The application component at the client will be using the camera for scanning the packets and the CPU for running the parsing algorithms. It will communicate with the server, which has a link to the database, to get the latest set of harmful

substances and possibly update the client-related data.

1.3 BUSINESS MODEL

Contenta will be released as a free application at the beginning. The main revenue is projected to come from ads in its first months in the market. We will try our best to promote *Contenta* and gain as many users as possible in the first 3 to 6 months with the ad-based free application model. During the summer of 2024, our team hopes to expand the area of products that Contenta can analyze so that a premium version can be released with a new update. With a monthly subscription for a fixed fee, premium users will be able to scan a wide range of products from cosmetics to hygiene goods. Whether the functionality of scanning nutrition facts and obtaining the Nutri-Score will be included in the free or premium version is subject to change.

1.4 CONSTRAINTS

1.4.1 IMPLEMENTATION CONSTRAINTS

- Library Constraints: There are few resources and APIs at our disposal for AI image processing, NLP text correction, and missing text filling (currently what we predict are necessary). Their accuracy is our constraint. The image processing accuracy of reading the food packaging label information correctly will create a constraint on how successfully the app will function.
- AI Storage Constraint: We have different AI technologies in use, and we have to decide between either storing the AI models within the app or having a server do the necessary computing. Both have downfalls as well as benefits.
- Internet Access Constraint: The user's mobile phone will need to have access to the internet in order for the app to display any information about the packages.

1.4.2 ECONOMIC CONSTRAINTS

It is costly to release an application in the IOS App Store, whereas it is free for the Google Play Store and our budget is minimal. The database, website, and any other components have to be chosen from less costly options.

1.4.3 ETHICAL CONSTRAINTS

We wish to collect as little data as possible, as we were instructed in our course seminars. Moreover, we wish to perform the ingredient scanning without having a product-based database so we wish to avoid storing each product. Instead, we wish to store limited user data and ingredient data, while giving appropriate references to articles about any ingredients. This is to minimize the falsehood of any information given.

1.4.4 HEALTH CONSTRAINTS

This is an application about health. Therefore, it needs to give correct outputs. It may lead to severe problems in case of failure. For example, if a product has an egg and someone with an egg allergy consumes the product by relying on the application's reporting that there is no egg in the product, it may lead to a severe crisis. Besides careful implementation, consulting medical doctors may be necessary.

1.4.5 TIME CONSTRAINTS

We have to have a complete running system by mid-May 2024 and release it to the market as soon as possible. Any development lifecycles beyond our first product will enable us to update the existing application, therefore we need to produce an end product by the end of the term. Furthermore, we are students who are working on the project within the school period. We will need to prioritize only the “food products” category at the first stages of our app and move to the other fields of packaging if there is a feasible time to implement other categories such as cosmetics, healthcare, and so on.

1.5 POSSIBLE RISKS

- Possibility of users’ personal data leakage: Data needs to be stored securely. Otherwise, it may be subject to GDPR or KVKK.
- Possibility of failure of scanning labels perfectly: As mentioned above, this may lead to severe outcomes like allergy crisis. Therefore, an approach similar to *Tesla*’s should be adopted: users should be warned that they shouldn’t solely rely on applications especially if they have special condition(s) like allergy.

1.6 OPTIONAL ADDITIONAL FUNCTIONALITIES

If the implementation of the main functionalities takes shorter than expected, we plan to add translation support to *Contenta* where the users will be able to scan an ingredient list of any mainstream language and have the analysis of the ingredients in English. Broadening the scope of products beyond food is not our goal until the final presentations, but if the implementation stage goes much faster than we expected, we might proceed with that.

2.0 DESIGN REQUIREMENTS

2.1 FUNCTIONAL REQUIREMENTS

- *Sign Up:* A user can sign up with an email and password.
- *Login:* A user can log in from different devices by using email and password.
- Forgot Password: A user can get an email to change his/her password whenever he/she forgets it.
- *Change Password:* When logged in, a user can change his/her password by using the old password.
- *Add Allergy:* A user can add an allergy from the allergies list. The application will notify the user if the user scans a food that contains an ingredient that may trigger the allergy.
- *Select Language:* A user can select a language to scan. A user can change this selection at any time.
- *Scan Ingredients:* A user can scan food's ingredients and then select the desired language area. The application will detect and inform the user about the current status of food additives in the US, UK, and EU. The application will detect if there are any ingredients that may trigger the user's allergies. The application will warn users if there are any potentially harmful substances in the product.
- *Become Premium Member:* Users can become a premium member by subscribing. Subscription fees will be determined in the later stages.
- *Terminate Subscription:* A premium user can stop his/her subscription.
- *Scan Nutrition Facts Table:* A premium member can scan the food nutrients part in the label. The application will detect the Nutri-Score (A, B, C, D, or E) of the food which will be calculated according to EU health legislation.
- *Add Unwanted Ingredient:* A user can select an additive from the list and get a warning when they scan food ingredients and the application detects the ingredient.
- *Provide Feedback:* Users can provide feedback to the developer team.

- *Delete Account:* A user can delete his/her account. His/Her personal information will be deleted in this case.

2.2 NON-FUNCTIONAL REQUIREMENTS

2.2.1 USER-FRIENDLINESS

The app's users are expected to be of all ages, so the UI should be intuitive and easy to understand even in its first-ever use. Hence, it should be similar to other healthcare apps on the internet and it should have a friendly, inviting look to it- especially since it is a healthcare app. The pages should look clean and the fonts should be big enough for the users to read the ingredients. Since the fonts on packaged food products are smaller in size, we wish for our users to have a more convenient experience and prefer to scan the ingredients to a format that they can read through easily. Also, a sweet spot between easing the mechanism of processing the image and making it easy for the customer to scan it should be found. Making the user select the part that includes the ingredients after taking a picture of the ingredients would remarkably fasten the processing stage but it should not be hard for the user to do so. The user should not wait more than 5 seconds for any operation without providing any input. Furthermore, the splash onboarding screen will give a tour of the app when it is opened the first time, easing the process of understanding the app.

2.2.2 MAINTAINABILITY

The app should have as little maintenance as possible and the end result of our production should be a self-sufficient app that can continue its functionality with as little reliance on updates as possible. The app is aimed to be non-product-based, which greatly increases its maintainability in the long term. The database will avoid holding information related to the products themselves but will be keeping user-related information, ingredient information, and blog information. The process of identifying the ingredients within the product is based solely on what is written on the package itself. Therefore if the ingredients of a product change, then the ingredients label is also changed. This means our app has all the data necessary to function as needed even when a product changes completely, without having to rely on a database update whenever a change in a product's ingredients is detected. This greatly increases the sustainability of the app as well as its accuracy in providing correct healthcare advice. The information about ingredients will be changed according to new findings, which are maintained with expert blogs and feedback. User information is maintained by themselves for the most part and we keep their data only as long as the user wishes to continue using our product. Furthermore, blogs are maintained by the experts and they can edit, remove, or add blogs as they wish. If any blog is found to contain false information, experts are informed by the customers and they should revise their

blogs accordingly. Any necessary action may be taken by the admins to remove & edit the blogs as well as take action against experts; warn them, or take the rights of contribution to the app - experts are fully responsible for their own misinformation and are penalized. The servers should not stay down for more than 8 hours and the users should be notified at least 24 hours before a planned maintenance.

2.2.3 SCALABILITY

In our implementation we have both short-term and long-term goals. We are developing the application with the intention of adding our long-term goals such as transitioning to cosmetic products after our application works on food-related products if the application gets good attention. Since we have a road map in mind, the foundations and the database of the app will be compatible with other types of product categories and ingredient categories. We plan to support up to 10,000 users at the beginning and improve it as the number of users increases.

2.2.4 PERFORMANCE

The processing speed of the images is crucial for the user experience as it will probably be the bottleneck of the whole waiting time in the application overall. We plan to make the user's CPU run the image processing algorithms. If we were to give an approximate number from the initial tests we performed, the output should be created in 5 seconds at maximum on a Snapdragon 855 or higher. Since the users' mobile phones each have different processing speeds, although the processing time may vary, it should preferably be between 3 to 7 seconds (given 2 seconds of uncertainty to our test value of 5 seconds) for more than 90% of our users.

2.2.5 PRIVACY

The application will collect personal data that may be linked to real people, storing their preferences and healthcare information as long as they wish to. Since our application is made and maintained in Turkey, we have to store data in Turkey and delete data if a user deletes their account. The terms and conditions, as well as the privacy policy of *Contenta* are confirmed by the users when they wish to use the application, and the user is notified accordingly that they accept these documents fully. Regarding the regulations and laws in Turkey, data is protected and kept according to the 6698th law of the legislation law "Mevzuat", and the health-related data is kept for the reasons of health services, and any information is terminated as soon as the user wishes to deletes their account, or when there is no need for keeping any necessary user data [5]. The user has full control of their data, and we aim to

make the process as transparent as possible by only storing the data the user wishes us to store, and the data will only be processed for the purposes of the app functionalities the user wishes to use, such as the calories of food they took will only be stored to calculate the calorie intake with respect to their diet and the processed data will be given to the user directly. Unless the user wishes us to store any data by providing us with the information themselves by interacting with the app's UI, no data will be stored and all unavailable data will be deleted permanently.

3.0 FEASIBILITY DISCUSSIONS

3.1. MARKET & COMPETITIVE ANALYSIS

3.1.1 SIMILAR APPLICATIONS

There are several applications/websites which serve similar purposes.

- *Ecomercek*: This is a website that lists products and their ingredients. They also warn users about potentially harmful substances and potential allergens in the food. However, this is a product-based application that requires maintenance every time a new product enters a market or a product's ingredients are changed. Besides, *Ecomercek* cannot be personalized for individual requirements like allergies or undesired ingredients.

- *Ingredio*: This is a scanning-based application that works similarly to the application we plan to implement. However, it cannot be personalized for individual requirements like allergies or undesired ingredients. It fails to recognize artificial flavors such as orange flavor. It lists the orange flavor as orange. It works in English and French. It doesn't classify foods as class A, B, C, D, or E according to EU legislation.

3.1.2 THE DIFFERENCE BETWEEN US AND THEM

Our team is aware of the fact that we are diving into a market that has great competition. Although the main functionality of *Contenta* is very similar to our competitors, the main architecture of the application differs a lot. Most of the applications in the market that serve similar purposes as we do have huge databases for storing the information of each and every product possible so that they can find relevant information from the barcodes. On the other hand, *Contenta* is not dependent on products, i.e., it is a text-based application. This has its advantages and disadvantages but we will be using its advantages to prevail against our competitors.

3.2 POTENTIAL TECHNICAL CHALLENGES

- Parsing the ingredients: The major challenge is to scan the ingredients. Food packages may have all shapes and they are almost never flat. Scanning non-flat surfaces is a challenge.

- Detecting artificial flavors and sentences: Detecting artificial flavors requires us to use language processing algorithms. While we project that this will not be a big problem for English, it may become a challenge in Turkish and other languages because of the lack of libraries. Also, some products include sentences instead of lists in their ingredients section. Parsing the ingredient list from sentence-like structures is a significant challenge.

4.0 PRODUCTIZATION

4.1 PROJECTED TIMELINE

We aim to have a working image-processing and parsing algorithm to a limited extent by January 2024 so that we can have a better demo presentation with our supervisor. We will be writing the Analysis and Requirements Report by the 1st of December until then. Our goal is to finish the implementation by May 2024 at the latest so that we can put *Contenta* into the Google Play Store and have some active users before our final presentations in mid-May. After the presentations, we want to further implement the additional features so that we can release the update that will bring the subscription model into life around September 2024. The following steps will be dependent on the success of *Contenta*.

4.2 POSSIBLE INVESTORS

As of mid-November, we are in talks with someone who is keen on supporting *Contenta* financially. He might partner us up with some other people who might be interested in our application. However, this requires our team to write a marketing/investor’s deck. Until December, an investor’s deck will be written so that we don’t miss the opportunity of a possible investor this early in the project.

4.3 LICENSING

Contenta is an application that carries most of its value implicitly in its idea of scanning the ingredients and possibly nutrition facts section to analyze the product, as the implementation of the application is not that complicated for an average developer. As it was suggested by our contacts in

the industry, we will take the necessary actions to license our product if it is possible. We will evaluate the cost of it and do what we can to protect the rights of *Contenta*.

5.0 CONCLUSION

This report on the specifications of the *Contenta* mobile application unfolds a detailed roadmap for the development of an innovative and user-centric solution aimed at enhancing consumer awareness and safety in the context of packaged food consumption and potentially cosmetics & cleaning product usage.

The journey begins with an insightful abstract (Section 0) that succinctly encapsulates the essence of *Contenta*, introducing its core functionalities such as image recognition, allergen detection, and language processing algorithms. The report navigates through the landscape of user needs and challenges in the introduction (Section 1.0), emphasizing the necessity for a tool like *Contenta* to address the complexities of understanding product ingredients, especially for individuals with specific health concerns. Section 1.1 dives into the core description of *Contenta*, illuminating its features, objectives, and potential for expanding beyond food products to cosmetics or cleaning goods. The user-centric approach is introduced through the User Profile (Section 1.1.1), which allows individuals to personalize their experience based on allergies, health issues, and dietary preferences. The intricate process of scanning ingredients and nutrition facts is detailed in Sections 1.1.2 and 1.1.3, showcasing the seamless integration of image processing technology. The added dimension of informative blogs (Section 1.1.4) enriches the user experience, fostering a community of informed consumers.

Moving into Section 1.2, the High-Level System Architecture delineates the components of the proposed solution, elucidating the on-device parsing module, local database, and user interface. The architectural illustration in Fig. 1 provides a visual representation of the main components of the application and their communication. The Business Model (Section 1.3) outlines the strategic approach to releasing *Contenta* as a free application initially, capitalizing on ads before introducing a premium version with expanded product analysis capabilities. Section 1.4 introduces various constraints, including implementation, economic, ethical, health, and time constraints. Each constraint is meticulously addressed, laying the groundwork for risk mitigation and ensuring the ethical integrity of the application. The section on Possible Risks (Section 1.5) sheds light on potential challenges such as data leakage and the importance of user responsibility in relying on the application. Section 1.6 highlights optional functionalities, showcasing the flexibility of the project's scope.

The Design Requirements (Section 2.0) detail both functional and non-functional requirements, underscoring the importance of user-friendliness, maintainability, performance, and privacy. Section 3.0 engages in a feasibility discussion, offering a market and competitive analysis (Section 3.1) and identifying potential technical challenges (Section 3.2) such as scanning non-flat surfaces and detecting artificial flavors. The Productization section (Section 4.0) provides insights into the projected timeline, potential investors, and licensing considerations, concluding with a commitment to timely implementation and a strategic approach to protect *Contenta*'s intellectual property.

In essence, this report culminates in a comprehensive and forward-looking vision for *Contenta*, positioning it as an indispensable tool that not only addresses current market needs but also anticipates and adapts to future challenges and opportunities. The combination of technical innovation, user-centric design, and strategic business planning positions *Contenta* as a transformative force in the realm of informed and health-conscious consumer choices.

6.0 LIST OF ABBREVIATIONS

- AI: Artificial Intelligence
- EU: European Union
- GDPR: General Data Protection Regulation
- KVKK: Kişisel Verilerin Korunması Kanunu (Personal Data Protection Law of Türkiye)
- Nutri-Score: Nutrition Score (a.k.a 5-Colour Nutrition Label)
- UI: User Interface

REFERENCES

[1] M. Janani, P. Selvasekaran, M. Lokanadham, and R. Chidambaram, "Food and food products associated with food allergy and food intolerance – An overview," Food Research International, vol. 138, Part B, pp. 109780, 2020, ISSN 0963-9969, [Online]. Available: https://doi.org/10.1016/j.foodres.2020.109780.

[2] Calderone, Julia. “24 foods that artificial sweeteners are hiding in”. Business Insider. Accessed: Nov. 16, 2023. https://www.businessinsider.com/surprising-foods-that-artificial-sweeteners-are-hiding-in-2016-1.

[3] “Acesulfame Potassium”. Center for Science in the Public Interest. Accessed: Nov. 16, 2023. https://www.cspinet.org/article/acesulfame-potassium.

[4] “FDA Food Packaging Guidelines for 2023”. NewPrint. Accessed: Nov. 16, 2023. https://www.newprint.ca/blog/fda-food-packaging-regulations/.

[5] Türkiye, TBMM. (2016, Mar. 24). Kişisel Verilerin Korunması Kanunu. Accessed: Nov. 16, 2023. https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=6698&MevzuatTur=1&MevzuatTertip=5.