Human Systems Engineering and Design (IHSED 2026), Vol. 237, 2026, 304–311
https://doi.org/10.54941/ahfe1008170

A Multimodal AI System Concept for Large-Scale Consumer Opinion Analysis and Dual-Quality Detection Across European Markets

Maciej Niemir[1,2] and Jakub Gołąb[1]

[1]Łukasiewicz Research Network - Poznan Institute of Technology, Poznań, Poland
[2]Poznan University of Technology, Faculty of Engineering Management, Poznań, Poland

ABSTRACT

This paper presents the concept of a multimodal artificial intelligence system designed for large-scale acquisition, integration, and analysis of consumer opinions to support the detection of the so-called dual quality phenomenon in products available on European markets. The proposed approach combines multilingual natural language processing, image analysis, automatic speech recognition, as well as product identification and entity-linking mechanisms. The system covers heterogeneous data sources, including online stores, marketplaces, blogs, forums, social media platforms, and audio-video materials. It uses dedicated data acquisition components, structured data standards, and methods based on large language models for processing unstructured content. A key stage of the pipeline is the transformation of source data into source-product-opinion relations, which are subsequently subjected to anonymization, deduplication, credibility assessment, sentiment analysis, and detection of signals potentially related to dual quality. An essential element of the concept is the aggregation of products originating from diverse sources and markets, despite inconsistent names, missing identifiers, and language differences. This process combines automatic methods with expert validation in a human-in-the-loop model, enabling the correction of product assignments, interpretation of ambiguous content, and adjustment of the scope of analysis to the objective of a given study. The proposed concept addresses the needs of institutions responsible for market surveillance by supporting the prioritization of regulatory actions and the selection of products requiring further analysis, including laboratory testing. The system is not intended to unequivocally confirm the occurrence of dual quality, but rather to indicate cases in which multilingual and multimodal consumer data provide grounds for in-depth verification.

Keywords: Dual quality detection, Consumer opinion mining, Multimodal data processing, Product entity linking, AI in product data quality, Sentiment analysis, Data aggregation, Market surveillance

INTRODUCTION

The progressing globalization of markets and the intensification of cross-border product distribution within the European Union have led to significant changes in the functioning of consumer markets. At the same

Received May 12, 2026 Revised June 23, 2026 Accepted July 10, 2026 Available online September 25, 2026

time, the phenomenon of the so-called dual quality of products, consisting in differences in the composition or quality of products sold under the same brand in different countries, has become an issue of growing interest for regulatory institutions and consumers. In market practice, however, identifying such cases requires access to distributed, multilingual, and often unstructured data, which significantly hinders their systematic analysis.

One of the most valuable sources of knowledge about actual consumer experiences is opinions published on the Internet. They include both textual content posted in online stores, forums, and blogs, as well as audiovisual materials published in social media. However, such data are characterized by high heterogeneity, multimodality, and linguistic variability, which means that their effective use requires advanced artificial intelligence methods and appropriate integration mechanisms. Existing approaches to consumer opinion analysis focus mainly on text processing and sentiment analysis, rarely considering the full multimodal context and the problem of unambiguously linking opinions to specific products.

In response to these challenges, this paper proposes the concept of a multimodal artificial intelligence system enabling the analysis of consumer opinions in the context of detecting the dual quality phenomenon in products. The paper focuses on presenting its high-level architecture and main components, omitting a detailed evaluation of models, which remains the subject of further research.

RELATED WORK

Research on the automatic detection of quality problems in products based on Internet data focuses on three areas: identification of quality signals, assessment of data credibility, and analysis of consumer opinions.

The first area of research concerns the identification of complaints and quality problems, using aspect extraction and topic modeling, including knowledge graph-based approaches (Zheng et al., 2025b), which are also effective in the analysis of multilingual data (Singh & Saha, 2023). The second area focuses on detecting fake reviews and anomalies, which is important for assessing the credibility of input data (Quyyam & Yu, 2025; Lee et al., 2022; Sharma & Singh, 2023). The third stream includes opinion and sentiment analysis, currently developed mainly with the use of transformer-based models (Syamsuar & Marcello, 2024). These methods enable opinion analysis at the product-aspect level (Xu et al., 2023) and the identification of trends and topics in data (Zheng et al., 2025a, 2025b).

In this context, the only work directly addressing the dual quality problem is the publication by Brzeziński et al. (2025), in which the authors proposed the detection of so-called "dual quality mentions" in consumer opinions using transformer-based models. This approach, however, focuses exclusively on textual data analysis and does not include full data integration or a scalable system architecture.

Within the above areas, there is a lack of solutions integrating opinion analysis with product identification and multimodal data processing, particularly in the context of dual quality. The system concept proposed in this paper addresses this gap by integrating multimodal methods with a human-in-the-loop approach.

SYSTEM CONCEPT

The proposed system is a conceptual analytical platform which, from the perspective of institutions responsible for market surveillance, may serve as a tool supporting the selection of cases for in-depth analysis. Its task is not to automatically determine the occurrence of dual quality, but to indicate areas where there are grounds justifying further actions, such as expert analysis, contact with the manufacturer, or laboratory testing.

The system combines structured and unstructured textual and audiovisual data originating from heterogeneous sources and then processes them within a common analytical pipeline. The central element of the concept is the transition from distributed and ambiguous source data to initially structured product-opinion relations. These relations constitute the basis for analysis, aggregation, and reporting. Due to the incompleteness of product data on the Internet, differences in product naming, and the lack of unambiguous identifiers, the system combines automatic methods with expert validation in a human-in-the-loop model.

The general structure of the proposed solution is presented in Figure 1. The following subsections describe the main stages of the pipeline: data acquisition, multimodal extraction, analytical processing, product aggregation, and reporting of results.

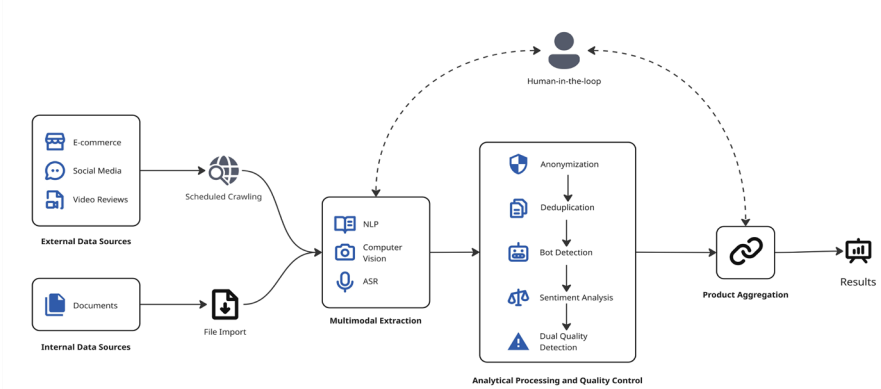

Figure 1: Conceptual pipeline of the proposed multimodal AI system.

Data Acquisition

In the context of detecting dual quality signals, it is particularly important that data originate from many independent sources, cover different markets, and reflect actual consumer experiences.

The system distinguishes two basic types of sources: external sources, acquired directly from the Internet through web scraping, and sources provided by the user. The starting point consists of keyword-based queries, such as product names, barcode numbers, or category phrases entered by the user. On this basis, the system uses Internet search engines to retrieve potential data sources for automatic processing; alternatively, a source may be indicated directly by the user.

For online stores or shopping platforms, the acquisition process focuses on extracting all available information about products - names, categories, descriptions, identifiers, detailed attributes - and related consumer opinions. Such sources usually have a repeatable page structure, and the result of the process is typically data in a product-opinion structure.

Data acquisition from unstructured or weakly structured sources, such as blogs, discussion forums, textual product reviews, or selected social media content, is different. In such cases, the entire page is retrieved together with its structure and metadata.

Audiovisual sources constitute a separate category, including, among others, product reviews published on platforms such as YouTube, TikTok, or Instagram. In this case, the system retrieves not only the video material itself, but also accompanying metadata: title, description, thumbnail, comments, and other available contextual elements that may support assessment of the material's usefulness and the overall context of the acquired information.

An alternative data acquisition channel consists of sources provided directly by the system user. These may include both structured files, such as spreadsheets, and video files. In this case, it is important to preserve the relation between the product, opinion, and source of origin so that repeated imports do not result in duplication.

Multimodal Extraction

The next stage of the presented system concept is the extraction of information from the acquired data and bringing it into a form suitable for further analytical processing. The main goal of this layer is therefore to create the basic source-product-opinion relation.

The extraction method depends on the type of source. As mentioned earlier, in the case of online stores and marketplaces, the product-opinion relation is usually structured and clear. The presence of structured data, such as tags compliant with the schema.org standard, is an additional facilitation, as it allows specific fields to be directly identified without the need to use advanced AI algorithms.

Processing unstructured or partially structured sources, such as blogs, discussion forums, product rankings, advisory articles, or selected social media content, is much more difficult. In such materials, even the product name does not always appear explicitly, and the opinion may be dispersed across a longer text fragment. In addition, such content often combines informational, marketing, comparative, and subjective user statements. Therefore, the system must use modules for the analysis of unstructured

content (Plaskowski et al., 2024), whose task is to recognize product mentions, separate consumer opinions from marketing descriptions, and link these fragments to the appropriate product context.

The analysis of audio-video materials also poses a challenge. Statements in videos are often colloquial: a product or an entire range of products in a single video may be referred to by pronouns, a brand name, or a colloquial expression, without providing the full commercial name. For this purpose, in our concept, audiovisual material is decomposed into two representations. The audio track is transcribed using automatic speech recognition (ASR), which enables further processing of speech as text, while the image is analyzed by extracting selected frames to detect occurrences of the product or its packaging.

The combination of these modalities is particularly important in situations where the user says only "this product", while the actual product is visible on the screen. Product identification based on images, however, requires access to a reference product database containing names, variants, attributes, and images of packaging. If the product cannot be unambiguously matched to the reference database, the system retains the name or description extracted from the transcript. Due to the possible uncertainty of both product identification and the opinion assigned to it, the result may be passed on for validation and correction by the user.

Due to the computational cost of downloading and processing long videos, our concept also includes a preliminary assessment of the usefulness of an audio-video source. Using metadata - title, description, comments - and the video preview image, large multimodal language models can be used to estimate whether the material will contain opinions about products and then present the user with a recommendation in the form of a score together with a justification. This mechanism constitutes another point of human-system interaction.

Analytical Processing and Quality Control

The analytical processing and quality control layer handles analyzing the previously extracted product-opinion relations and preparing them for further aggregation and reporting. At this stage, the system no longer operates solely on raw content, but on structured tuples containing information about the product, opinion, source, and available metadata.

The first task of this layer is anonymization of opinion content. Since opinions may contain sensitive information, the system should mask such information while preserving the meaning of the statement and information relevant to further analysis. The system should primarily identify named entities such as personal data, account names, addresses, order numbers, locations, etc.

The next step planned in the system concept is deduplication of the identified opinions, conducted both at the level of exact text matching and semantic similarity, since opinions may be slightly modified, shortened, or rewritten. However, the source context should be preserved. An identical

short opinion appearing in separate places on the Internet does not always mean the same analytical record; it may also be an independent occurrence of the same type of consumer experience.

The system then assesses the probability of automatic origin of the content, i.e., bot detection. The bot detection module assigns an opinion a credibility score or a probability of automatic generation based on linguistic, structural, and contextual features. The result of this module does not have to lead to automatic rejection of the opinion but may serve as additional information about its credibility.

After basic quality control of the collected opinion, the data are enriched with sentiment analysis results. The module determines the sentiment of the opinion towards the indicated product, in its basic form as positive or negative. In a more advanced variant, the analysis may even be conducted at the aspect level, such as taste, smell, composition, performance, durability, packaging, or general perception of quality, depending on the product category.

Of particular importance is the module for detecting dual quality signals at the level of a single opinion. The module identifies statements indicating possible quality differences between products offered under the same or a similar brand in different markets, analyzing, among others, comparisons between countries, references to changes in composition, and differences in taste, smell, effectiveness, consistency, or packaging.

Aggregation and Reporting

The final stage of the pipeline concept is data aggregation and preparation for reporting. Its purpose is to transform the collected source-product-opinion relations into analytical groups corresponding to any scope of study selected by the user.

The main challenge here is linking product data obtained from multiple sources, where in many cases the only information available for aggregation is the product name, and that name is not consistent due to the nature of the source page, country, and lack of standards for product descriptions on the Internet (Niemir & Mrugalska, 2022). Aggregation is therefore not simple grouping by name or global identifier, but a process of creating product sets consistent with the objective of the analysis, where the user, supported by algorithms, defines them independently. The user may define a group narrowly, for example as a specific product variant, deliberately omitting sources with overly general product names, or broadly - as a product line, brand, or even category. As a result, the same dataset can be used repeatedly at various levels of detail and in different studies simultaneously.

In the proposed concept, part of the aggregation can be performed automatically - for example when global GTIN identifiers are available, or when product names in various sources are identical. However, the greatest advantage of the approach is algorithmically supported freedom of search. Linking proposals can be generated using classical ranking-based

text comparison methods, such as BM25, embedding models, or cross-encoders, which support similarity assessment despite language differences, abbreviations, and spelling errors.

The final decision to link products remains with the user. The system presents aggregation candidates together with similarity premises, while the user accepts, rejects, or corrects the proposed assignments. Approved decisions can then be used to iteratively extend groups with further similar records. In this sense, the system gradually supports the matching process, learning from earlier user choices, but does not replace expert judgment.

The result of aggregation consists of product groups together with assigned opinions, sources, metadata, and results of our analyses. Only at this level is it possible to prepare comparative reports by country, market, source, brand, variant, time, or category. Reports may include aggregated sentiment, frequency of complaints, characteristic problems, and the number of signals potentially related to dual quality. These summaries do not automatically confirm the phenomenon of dual quality but provide an ordered basis for further expert assessment.

CONCLUSION

This paper presented the concept of a multimodal artificial intelligence system supporting the analysis of consumer opinions in the context of detecting signals potentially related to the phenomenon of dual quality in products. The proposed approach integrates textual, product-related, and audiovisual data originating from heterogeneous sources and processes them within a common analytical pipeline.

A key element of the system is the transformation of source data into source-product-opinion relations, followed by their aggregation into analytical groups corresponding to the scope of the study. Due to the inconsistency of product names, the multilingual nature of data, and the frequent lack of identifiers, this process requires combining automatic methods with expert validation.

An important contribution of the paper is framing the system as a decision-support tool rather than a solution that automatically confirms the occurrence of dual quality. The human-in-the-loop model enables correction of product assignments, validation of results, and interpretation of quality signals in the context of a specific study.

The work is conceptual in nature. Further research should include evaluation of the effectiveness of extracting product-opinion relations, identifying products in multimodal data, aggregating product groups, and detecting dual quality signals in real market monitoring processes.

ACKNOWLEDGMENT

This work was supported by the National Centre for Research and Development (NCBR) under the INFOSTRATEG III program, project INFOSTRATEGIII/0003/202100 “Produktoskop”.

REFERENCES

Brzeziński, M., Niemir, M., Muszyński, K., Lango, M., & Wiśniewski, D. (2025). Boosting Dual Quality detection with AI-based social media analysis. Information Processing & Management, 62(4), 104138. https://doi.org/10.1016/j.ipm.2025.104138

Lee, M., Song, Y.H., Li, L., (...), Yang, S.-B. (2022) Detecting fake reviews with supervised machine learning algorithms. Service Industries Journal. https://doi.org/10.1080/02642069.2022.2054996

Niemir, M., & Mrugalska, B. (2022). Identifying the cognitive gap in the causes of product name ambiguity in e-commerce. Logforum, 18(3), 9. https://doi.org/10.17270/J.LOG.2022.738

Plaskowski, D., Skwarek, S., Grajewska, D., Niemir, M., & Ławrynowicz, A. (2024). Automating Opinion Extraction from Semi-Structured Webpages: Leveraging Language Models and Instruction Finetuning on Synthetic Data. 681–688. https://www.scitepress.org/Link.aspx?doi=10.5220/0012384900003636

Quyyam, T., Yu, Q. (2025) A Comprehensive Survey of Fake Review Detection Technology. Smart Innovation, Systems and Technologies. https://doi.org/10.1007/978-981-96-0147-9_6

Sharma, K., Singh, A. (2023) A Systematic Review: Detection of Anomalies in Social Networks. 2nd International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2023 - Proceedings. https://doi.org/10.1109/ICSCDS56580.2023.10104612

Singh, A., Saha, S. (2023) GraphIC: A graph-based approach for identifying complaints from code-mixed product reviews. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2022.119444

Syamsuar, D., Marcello (2024) The Implementation of Artificial Intelligence for Online Review: A Systematic Literature Review. Proceedings of 2024 International Conference on Information Management and Technology, ICIMTech 2024. https://doi.org/10.1109/ICIMTech63123.2024.10780904

Xu, J., Zhang, J., Song, L., Gao, Y. (2023) Fine-grained sentiment analysis of online reviews based on RoBERTa-BiLSTM-CRF. Xitong Gongcheng Lilun yu Shijian/System Engineering Theory and Practice. https://doi.org/10.12011/SETP2022-2001

Zheng, L., He, Z., He, S. (2025) A topic model-based knowledge graph to detect product defects from social media data. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2024.126313

Zheng, L., Sun, L., He, Z., He, S. (2025) Dynamic product quality improvement using social media data and competitor-based Kano model. International Journal of Production Economics. https://doi.org/10.1016/j.ijpe.2025.109645