Personalized Language-Oriented Semantic Communication

Giordano Cicchetti, Eleonora Grassucci, Danilo Comminiello [1]

Abstract

In the emerging research field of Semantic Communication, the focus is on making network communications more effective and efficient. While prior research has primarily concentrated on improving transmission efficiency, limited efforts have addressed the overall effectiveness of communication. In this work, we propose a novel framework for personalized content dissemination in broadcast communication systems. In the envisioned system, the sender transmits either a textual description or a compressed latent representation of the content. At the receiver side, a personalization module leverages large language models to dynamically manipulate the received semantics, exploiting a user interest database. Subsequently, generative models recreate the final content conditioned on the manipulated semantics, ensuring that the regenerated content is both semantically consistent with the original and tailored to individual user interests. This communication paradigm aims to enhance overall system effectiveness and maximize user engagement through relevance and personalization. Potential applications include targeted advertising, personalized news or media delivery, and adaptive educational content.

1. Introduction

In recent years, the field of communication theory has undergone a paradigm shift with the emergence of Semantic Communication (SC), a novel approach that focuses on transmitting the meaning of data rather than the raw data itself (Qin et al., 2024; 2021; Seo et al., 2022; Gündüz et al., 2022; Luo et al., 2022). This shift is driven by the increasing demand for more efficient and effective information exchange, especially in bandwidth-constrained environments where transmitting full-resolution content is often impractical (Pei et al., 2025; Grassucci et al., 2023; Barbarossa et al., 2023; Han et al., 2022).

A promising area of research within SC is language-oriented semantic communication, which compresses the original data into textual descriptions to be transmitted over the channel as semantic vectors (Nam et al., 2024). Describing data through text substantially reduces network load, as transmitting a few words is far more efficient than sending raw multimedia data. However, textual representations alone are often too coarse or abstract to capture the full richness of multimodal content. To overcome this limitation, (Cicchetti et al., 2024) proposed augmenting the text with a compressed latent representation of the original content to better guide content reconstruction at the receiver. In parallel to data transmission efficiency, personalized semantic communication has emerged as a paradigm focused on tailoring transmitted content to individual user preferences or contexts, thereby enhancing relevance and communication effectiveness (Wang et al., 2024; Peng et al., 2024; Chen et al., 2024).

In this work, we aim to integrate those two complementary research directions by extending existing frameworks with a highly tailored personalization module at the receiver side. This module leverages Large Language Models (LLMs) (Das et al., 2025) to modify received semantics based on an individual user interests database available at receiver side. Generative models, such as Latent Diffusion Models (LDMs) (Rombach et al., 2021), can be conditioned on various forms of semantic input, and a key advantage is that the generated output naturally reflects the imposed semantics. Building on this capability, we feed the manipulated semantics into powerful generative models capable of producing final content that adheres to the intended semantics and aligns with the user's preferences and context, thereby enhancing the effectiveness of communication. In summary, our contributions aim to broaden the scope of semantic communication by introducing personalization as a fundamental design principle. Potential applications may include targeted advertising, personalized media delivery, and adaptive educational content, everything while maintaining high transmission efficiency.

[1]Dept. of Information Engineering, Electronics, and Telecommunication, Sapienza University of Rome, Italy. This work was supported by the European Union under the Italian National Recovery and Resilience Plan (PNRR) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program RESTART). Correspondence to: giordano.cicchetti <giordano.cicchetti@uniroma1.it>.

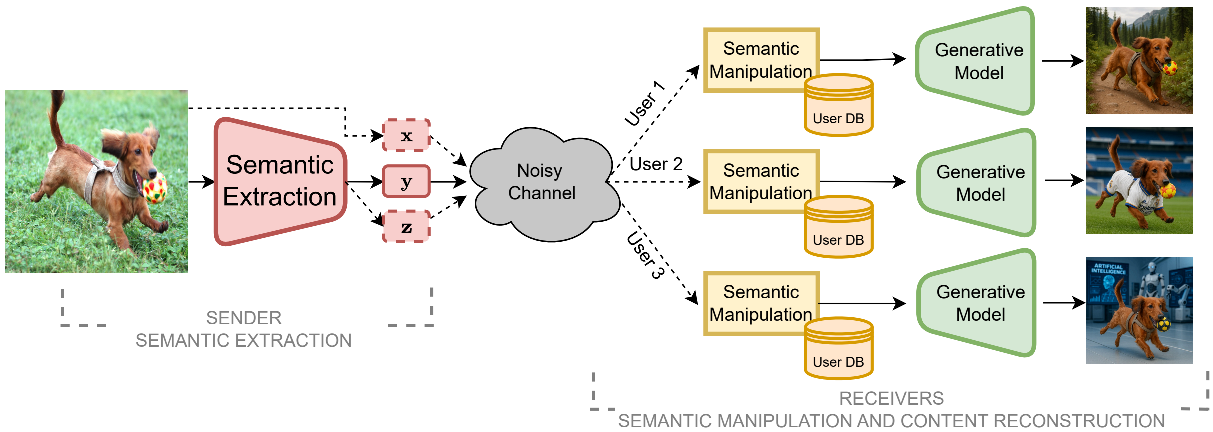

Figure 1. **Overview of the proposed personalized semantic communication framework.** The sender extracts semantic information from raw content and transmits it over a noisy channel. Each receiver employs a user-specific personalization module, leveraging a user interest database and a large language model to adapt the received semantics. The final content is reconstructed using a generative model, ensuring both semantic fidelity and personalization.

2. Proposed Framework

In this section, we present the conceptual design of the proposed semantic communication framework, which aims to adapt transmitted content based on network conditions and user interests.

2.1. Semantic Extraction

Building on prior research, semantic information can be extracted from raw data $\mathbf{x}$ using various methodologies. In this work, we consider two primary approaches:

- **Latent Representation via Neural Encoding.** A neural encoder $\mathcal{E}$ processes the input data $\mathbf{x}$ to produce a latent vector representation $\mathbf{z} = \mathcal{E}(\mathbf{x})$, capturing high-level semantics in a compressed format.

- **Data-to-Text Generation.** Data-to-text models are employed to generate natural language descriptions from input data. Given an input $\mathbf{x}$, the model produces a textual output $\mathbf{y}$ defined as a sequence of Y words:

$$\mathbf{y} = \mathrm{I2T}(\mathbf{x}) = (y_1, y_2, \ldots y_j, \ldots, y_Y),$$

  where y_j denotes the j-th word in the sequence, each composed of a variable number of characters.

We design the semantic extraction module to be dynamic and adaptive to the state of the network. When the network operates under ideal conditions (e.g., low traffic and high bandwidth), the full data $\mathbf{x}$ may be transmitted. As network conditions degrade, the communication system adapts by transmitting progressively less detailed representations of the content. We define three operational modes:

1. **Ideal Network Conditions.** Transmit the complete original data along with its textual description, therefore the semantic information ($\mathbf{s}$) equals the data: $\mathbf{s} = (\mathbf{x}, \mathbf{y})$.

2. **Moderate Network Degradation.** Transmit both the latent representation and the textual description, resulting in $\mathbf{s} = (\mathbf{z}, \mathbf{y})$.

3. **Poor Network Conditions.** Transmit only the textual description, resulting in $\mathbf{s} = \mathbf{y}$.

This flexible approach ensures that meaningful communication is preserved under varying transmission capabilities.

2.2. End-User Profiling

For each user, the framework maintains a user-specific profile database that captures individual interests to personalize the semantic content. In this work, we assume a simplified model in which a pre-existing database is available for each user. This database includes a set of high-level semantic interest categories:

$$I = \{c_1, c_2, \ldots, c_n\}, \quad c_i \in C,$$

where C represents the universal set of possible interest categories. For instance, categories might include broad themes such as sports, technology, health, finance, or entertainment, allowing the system to be aware of the user-specific domains of interest.

2.3. LLM-Based Semantic Manipulation

Upon receiving the semantic information $\mathbf{s}$, the system always has the possibility to retrieve the textual caption $\mathbf{y}$

that describes the semantics of the original content. Moreover, the receiver is equipped with a Large Language Model (LLM) that is exploited to manipulate received caption $\mathbf{y}$ in accordance with the interests of the end-user. Let I_i denote the interest set for the i-th user. The LLM processes both the received textual caption $\mathbf{y}$ and the user's interest profile to produce a personalized output as follows:

$$\hat{\mathbf{y}} = \mathrm{LLM}(\mathbf{y}, I_i),$$

where $\hat{\mathbf{y}}$ is a semantically enriched version of the original text $\mathbf{y}$, incorporating elements aligned with the user's preferences I_i.

2.4. Content Reconstruction

Content reconstruction is handled by a conditional generative diffusion model (DM), which produces a reconstructed image $\hat{\mathbf{x}}$ from the available semantic inputs.

In the case in which the receiver receives only the textual description of data, the model samples from the conditional distribution of data given the modified textual description:

$$\hat{\mathbf{x}} = \mathrm{DM}(\hat{\mathbf{y}})$$

This allows for the reconstruction of meaningful and contextually relevant content based on the semantically adapted message.

If a latent representation $\mathbf{z}$ is also retained at the receiver, generation is conditioned on both inputs:

$$\hat{\mathbf{x}} = \mathrm{DM}(\hat{\mathbf{y}}, \mathbf{z}).$$

Finally, in the case where the full image $\mathbf{x}$ is available to the end user, the generative model should modify such an image so that it meets the interests of the end user expressed in the modified textual caption $\hat{\mathbf{y}}$:

$$\hat{\mathbf{x}} = \mathrm{DM}(\hat{\mathbf{y}}, \mathbf{x}).$$

3. Experimental Results

To validate our approach, we identify four distinct interest categories to represent high-level user clusters:

- I_1: {*Nature, Hiking*}
- I_2: {*Sea, Beach*}
- I_3: {*Football, Real Madrid*}
- I_4: {*AI, Technology*}

We evaluate our method for the task of image transmission. To generate and manipulate the textual captions we employ

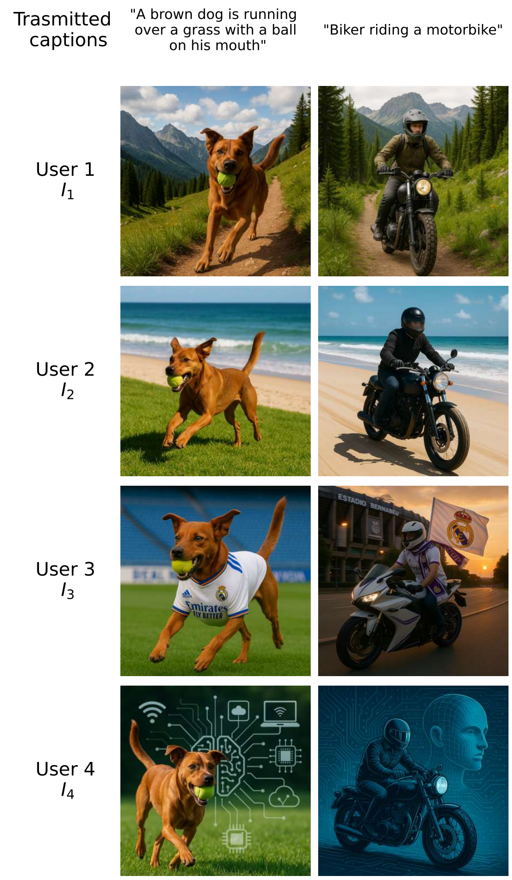

Figure 2. Visual results for the *textual caption only*. The first row shows the textual caption of the original images. The subsequent rows show the regenerated images at the receiver, after semantic manipulation has been applied. Each of these rows corresponds to a different end user, each associated with a distinct interest category I_i.

the well known GPT-4o by OpenAI. Final content generation is performed using DALL·E 3. Although these are proprietary models and not open-source, they represent the current state-of-the-art and demonstrate the quality achievable with contemporary generative AI technology.

During the simulation phase, we group the first and second operational modes described in Section 2.1 into a single category, referred to as the ***reference image available*** scenario. The second scenario corresponds to the third operational mode, where only the textual caption is available at the receiver. We refers to it as the ***text only available*** scenario.

This grouping is motivated by the internal behavior of the DALL·E 3 generative model, which operates in the latent space when modifying an input image. Specifically, if the full image is available at the receiver, DALL·E 3 first compresses it into a latent representation using an internal image encoder. In our second operational mode, this same encod-

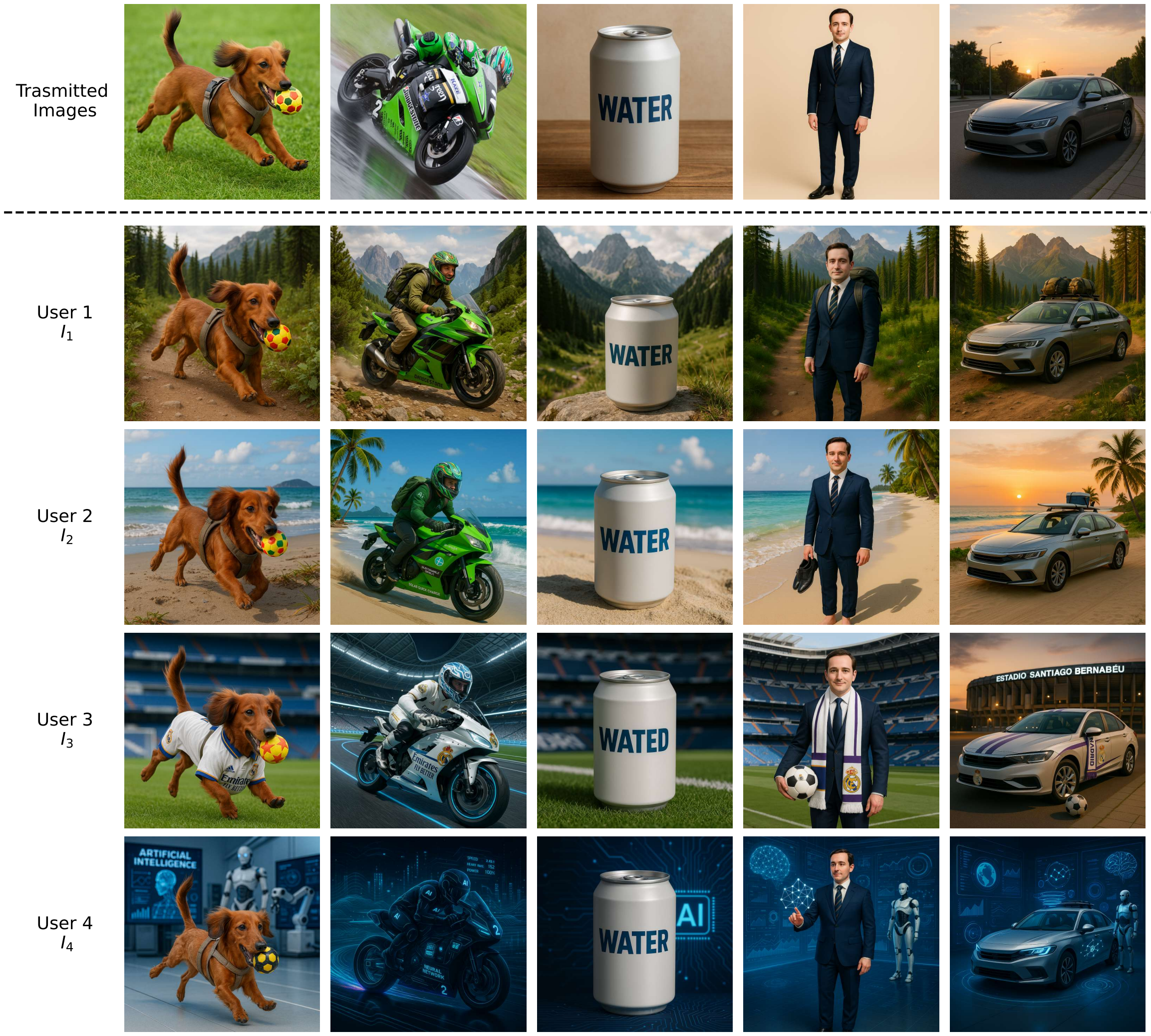

Figure 3. Visual results for the *reference image available* scenario. The first row displays the original images that the sender intends to transmit to the end users. The subsequent rows show the regenerated images at the receiver, after semantic manipulation has been applied. Each of these rows corresponds to a different end user, each associated with a distinct interest category I_i.

ing step is performed at the sender side. Thus, the distinction between these two modes lies solely in the placement of the image encoder, whether it resides at the sender or the receiver. Assuming error-free transmission, both configurations yield equivalent inputs to the generative model, allowing them to be treated as a single scenario.

Figure 2 and Figure 3 present the visual results for the ***text only available*** and ***reference image available*** scenarios, respectively. Notably, the generative model preserves the core semantic content of the original images while adapting the surrounding context, the objects, and some details to align with the proxy interest categories. This transformation goes beyond simple style transfer or background replacement; the entire scene is contextually altered, with relevant details added to better match the target domain. As a result, the final images are highly personalized, reflecting the user's specific interests. This process significantly enhances the overall effectiveness of the communication system by increasing the relevance of the content and maximizing user engagement.

4. Conclusion

In this work, we introduced a novel framework for personalized language-oriented semantic communication, aimed at enhancing the effectiveness of data transmission. By leveraging large language models and generative AI tools, the system adapts a common semantic message to reflect the individual preferences of each user, enabling highly tailored content delivery. Our experimental results, focused on image transmission, demonstrate the system capability to preserve the semantic intent of the original content while personalizing it in accordance with diverse interest profiles.

References

Barbarossa, S., Comminiello, D., Grassucci, E., Pezone, F., Sardellitti, S., and Di Lorenzo, P. Semantic communications based on adaptive generative models and information bottleneck. *IEEE Comm. Magazine*, 2023.

Chen, Z., Yang, H. H., Chong, K. F. E., and Quek, T. Q. Personalizing semantic communication: A foundation model approach. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 846–850, 2024.

Cicchetti, G., Grassucci, E., Park, J., Choi, J., Barbarossa, S., and Comminiello, D. Language-oriented semantic latent representation for image transmission. In *2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, 2024.

Das, B. C., Amini, M. H., and Wu, Y. Security and privacy challenges of large language models: A survey. *ACM Computing Surveys*, 57(6):1–39, 2025.

Grassucci, E., Barbarossa, S., and Comminiello, D. Generative semantic communication: Diffusion models beyond bit recovery. *ArXiv preprint: arXiv:2306.04321*, 2023.

Gündüz, D., Qin, Z., Aguerri, I. E., Dhillon, H. S., Yang, Z., Yener, A., Wong, K., and Chae, C.-B. Beyond transmitting bits: Context, semantics, and task-oriented communications. *IEEE Journal on Selected Areas in Communications*, 41:5–41, 2022.

Han, T., Tang, J., Yang, Q., Duan, Y., Zhang, Z., and Shi, Z. Generative model based highly efficient semantic communication approach for image transmission. *IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, 2022.

Luo, X., Chen, H.-H., and Guo, Q. Semantic communications: Overview, open issues, and future research directions. *IEEE Wireless Comm.*, 29(1):210–219, 2022.

Nam, H., Park, J., Choi, J., Bennis, M., and Kim, S.-L. Language-oriented communication with semantic coding and knowledge distillation for text-to-image generation. In *IEEE Int. Conf. on Acoustics, Speech and Signal Process.*, 2024.

Pei, J., Feng, C., Wang, P., Tabassum, H., and Shi, D. Latent diffusion model-enabled low-latency semantic communication in the presence of semantic ambiguities and wireless channel noises. *IEEE Trans. on Wireless Comm.*, 2025.

Peng, Y., Jiang, F., Dong, L., Wang, K., and Yang, K. Personalized federated learning for generative AI-assisted semantic communications. *ArXiv preprint: arXiv:2410.02450*, 2024.

Qin, Z., Tao, X., Lu, J., and Li, G. Y. Semantic communications: Principles and challenges. *ArXiv preprint: arXiv:2201.01389*, 2021.

Qin, Z., Liang, L., Wang, Z., Jin, S., Tao, X., Tong, W., and Li, G. Y. Ai empowered wireless communications: From bits to semantics. *Proceedings of the IEEE*, 112(7): 621–652, 2024.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. High-resolution image synthesis with latent diffusion models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10674–10685, 2021.

Seo, S., Park, J., Ko, S.-W., Choi, J. D., Bennis, M., and Kim, S.-L. Toward semantic communication protocols: A probabilistic logic perspective. *IEEE Journal on Selected Areas in Communications*, 41:2670–2686, 2022.

Wang, Y., Ni, W., Yi, W., Xu, X., Zhang, P., and Nallanathan, A. Federated contrastive learning for personalized semantic communication. *IEEE Communications Letters*, 28: 1875–1879, 2024.