

Advances and challenges in semantic communications: A systematic review

Ping Zhang^{1,2,*}, Yiming Liu^{1,*}, Yile Song¹ & Jiaxiang Zhang¹

¹State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China;

²Department of Broadband Communication, Peng Cheng Laboratory, Shenzhen 518055, China

Received 3 May 2023; Revised 27 September 2023; Accepted 29 November 2023; Published online 15 December 2023

Abstract: Inspired by the recent success of machine learning (ML), the concept of semantic communication introduced by Weaver in 1949 has gained significant attention and has become a promising research direction. Unlike conventional communication systems, semantic communication emphasizes the precise retrieval of conveyed meaning from the source to the receiver, rather than focusing on the accurate transmission of symbols. Thus, semantic communication can achieve a significant gain in source data compression, alleviate communication bandwidth pressure, and support new intelligent services, which is envisioned as a crucial enabler of future sixth-generation (6G) networks. In this review, we critically summarize the advances made in semantic information and semantic communications, including theory, architecture, and potential applications. Moreover, we deeply explore the major challenges in developing semantic communications and present the development prospects, aiming to prompt further scientific and industrial advances in semantic communications.

Keywords: semantic communications, semantic information, 6G, artificial intelligence, deep learning

INTRODUCTION

The development of information and communication technology (ICT) has a significant impact on modern society. In the past 70 years, communication systems have mainly relied on Shannon's information theory [1] for design and development. The evolution of mobile communication systems from the first generation (1G) to the fifth generation (5G) has been characterized by capacity expansions and technological advancements. However, these advancements have primarily focused on increasing the physical dimension of information transmission while approaching the limits of Shannon's information theory. For example, some evolutionary technologies have been proposed to enlarge the system capacity, including millimeter wave/terahertz communications, massive antenna arrays, and high-order modulation, etc. Nevertheless, these solutions face inevitable bottlenecks in the cost of hardware and the complexity of wireless communications.

The sixth-generation (6G) network is expected to bring a new level of connectivity and intelligence to the physical and digital worlds, supporting machine-intelligence services and the seamless interaction between humans and devices in Metaverse [2]. These new applications impose a set of challenging requirements, such as extremely high data rates, ultra-low latency, ultra-dense connectivity, substantially high energy and

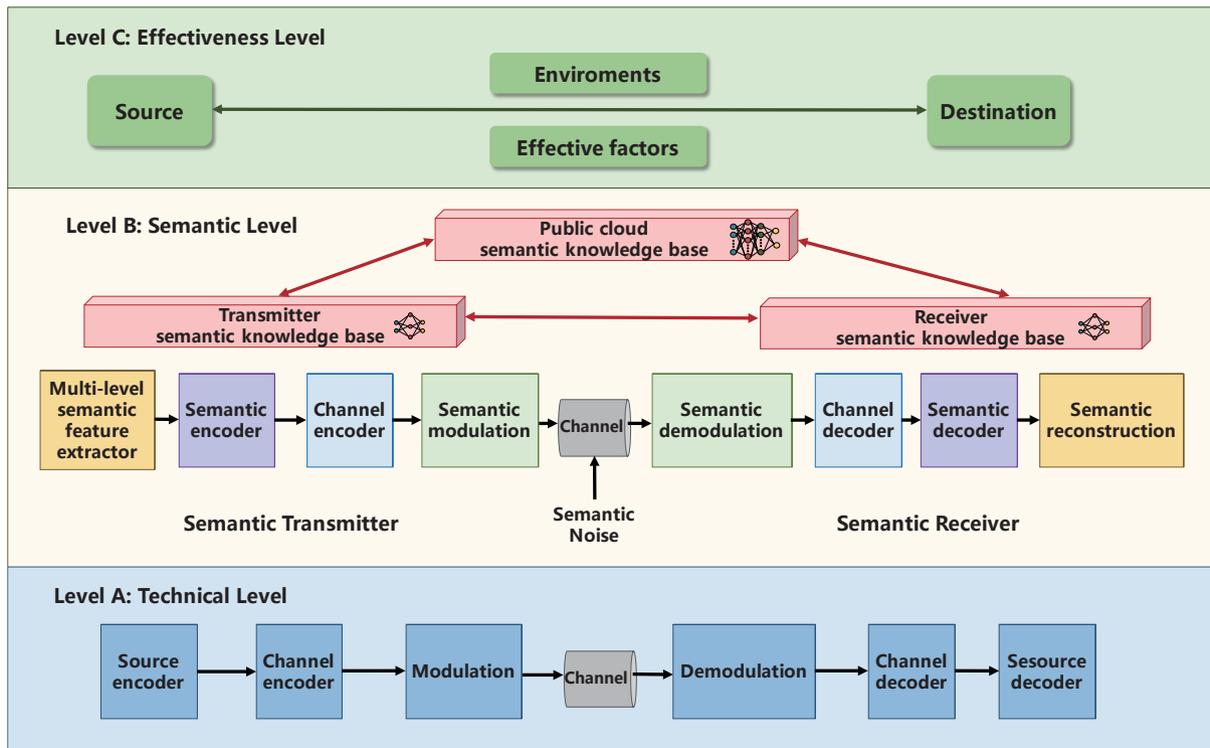


Figure 1 Three levels of communication.

spectral efficiency, and a high degree of intelligence. In addition, the implementation of new services and applications, such as autonomous driving, intelligent healthcare, and intelligent factories, has led to a need for machines to communicate with each other to execute tasks efficiently. In such cases, the goal of communication among machines is not always to reconstruct the exact message but to facilitate the receiver to make the correct inference, decision, and action at the appropriate time and within the correct context [3]. Therefore, there is a need for a paradigm shift in the design of 6G communications to effectively address these challenges and meet the requirements of these new applications.

In 1949, Shannon and Weaver [4] first identified three levels of problems within the broad subject of communication, as shown in Figure 1:

- (1) Technical level: How accurately can the symbols of communication be transmitted?
- (2) Semantic level: How precisely do the transmitted symbols convey the desired meaning?
- (3) Effectiveness level: How effectively does the received meaning affect conduct in the desired way?

Shannon's classical information theory (CIT) is centered around the technical level of communication and has made significant strides in establishing a rigorous mathematical theory based on probabilistic models. CIT defines information as the means of reducing uncertainty, and mutual information in the entropy domain is utilized to measure the amount of information transmitted. In contrast, semantic communication focuses on semantic level and emphasizes the meaning of messages rather than their syntactic representation, as depicted in Figure 1. By processing source messages to extract their semantics and transmitting only the relevant information, semantic communication has the potential to significantly reduce the amount of transmitted data while preserving the original semantics. Additionally, the semantic receiver can recover

meanings or perform intelligent tasks from received signals, even in the presence of intolerable bit errors at the technical level. As a result, semantic communication can substantially reduce bandwidth usage, enhance communication reliability, and naturally meet the demands of future networks for intelligence, conciseness, and personalization, which thus becomes a preeminent driver of the dramatic leap forward for 6G networks.

The concept of semantics was originally introduced in the field of semiotics [5]. In 1938, Morris [6] distinguished semiotics as three types of signs: syntactic signs, semantic signs, and pragmatic signs. Since then, many researchers have further developed the theory and proposed methods to address the challenges in the field of semantic information and transmission. Semantic information theory (SIT) primarily focuses on establishing the fundamental limits and boundaries of semantic communication systems. The most nucleus concern within the theory of semantic information is the meaning contained in the information [5]. Referring to the information entropy in CIT, Carnap *et al.* [7,8] first developed the SIT based on propositional logic and gave a mathematical expression of semantic information entropy. Later, other researchers investigated the semantic information entropy based on the field of situational logic [9], the concept of truth-likeness [10], the fuzzy set [11], etc. In addition to focusing on semantic uncertainty, some researchers consider semantic similarity to be a viable approach for quantifying semantic information. Floridi [12] developed a theory of strongly semantic information (TSSI) and corrected the problem of infinite information of sentences with semantic contradictions using the relative distance of semantics to measure the amount of information. Recently, Zhang *et al.* [13] proposed the concept of semantic base (Seb), which is used to represent semantic information in a high-dimensional space. Since research in the field of semantic communications is still in its infancy stage, a comprehensive and consistent theory regarding semantic communications has not yet been established.

With the advancements in artificial intelligence (AI) and computing power technology, there is now an opportunity to develop communication systems that can process semantic information. Recent advancements in AI have shown great potential in wireless communication scenarios, providing a viable path for undertaking semantic coding tasks for semantic communications. Researchers have tried to use the powerful feature learning and feature representation capabilities of deep learning models to model the semantic features of information sources, and have achieved a series of excellent research results [14,15]. In terms of semantic information representation of text sources, the recent text semantic encoding methods based on Transformer [16], represented by GPT [17] and BERT [18] models, have proven successful in many natural language processing (NLP) tasks. At the same time, different model structures, such as encoding-auto-encoding, decoding-autoregression, encoding-decoding, etc., have been developed. In terms of semantic information representation of image sources, deep learning models represented by convolutional neural networks (CNN) have been proven to be able to effectively learn multi-granular semantic information of images and can be widely used in classic computer vision tasks such as image classification and object recognition [19,20]. In terms of semantic representation of multi-modal data, researchers have also carried out research on multi-modal semantic representation models and methods of images and texts, taking multi-modal semantic alignment as a constraint. Based on this, in semantic communication systems, deep learning-based semantic extraction can be integrated into the communication architecture, which allows only the information of interest to the receiver for transmission, rather than raw data, thereby alleviating bandwidth pressure and enhancing privacy preservation by reducing the redundant data to be exchanged. Moreover, the proliferation of ubiquitous computing devices connected to wireless networks has led to the convergence of

communication and computing, providing a promising approach for efficient system design in the development of semantic communication.

The rapid expansion of semantic communication has received significant attention in recent years owing to its wide-ranging applicability. Numerous comprehensive surveys aim to address various aspects of semantic communication, including network architecture, theoretical analysis, potential technology, and future applications. Yang *et al.* [21] provided a comprehensive survey for the implementation of semantic communication in 6G and divided semantic communication into three categories: semantic-oriented communication, goal-oriented communication, and semantic-awareness communication. Lan *et al.* [22] provided a framework of semantic communication in the era of machine intelligence and defined three different areas of semantic communication, i.e., Human-to-Human (H2H) semantic communication, Human-to-Machine (H2M) semantic communication, and Machine-to-Machine (M2M) semantic communication, by identifying the involved subjects and objects. Qin *et al.* [23] provided a comprehensive overview on the principles and challenges of semantic communication, including semantic theory, principles, frameworks, and performance metrics of semantic communications. Iyer *et al.* [24] discussed the incorporation of semantic communication within intelligent wireless networks. Lu *et al.* [25] categorized the critical enabling techniques by explicit and implicit reasoning-based methods and elaborated on how they evolve and contribute to modern content & channel semantics-empowered communications. Liu *et al.* [26] discussed various technologies related to semantic communications, including semantic extraction, semantic coding, and semantic segmentation. They also provided a summary of solutions aimed at enhancing the efficiency, robustness, adaptability, and reliability of semantic communications. Getu *et al.* [27] provided a comprehensive survey for semantic and goal-oriented communication. Li *et al.* [28] discussed the representative techniques and use cases of some fundamental components in the ubiquitous semantic Metaverse. Shi *et al.* [29] proposed a novel architecture based on federated edge intelligence to support resource-efficient semantic-aware networks.

The above work provides a comprehensive survey of semantic communication in different areas and the related AI-based technical details. There is still a lack of a systematic survey article that provides a unified framework of semantic communication in the area of Seb theory, Seb-based semantic transmission, and semantic communication empowered “Intellicise” wireless network. In this paper, we first provide the general framework of Seb-based semantic communication and the related technical work. From Seb-based end-to-end semantic communication architecture, we also discuss the semantic communication for “Intellicise” networks, the new goal-oriented applications, and Metaverse. Finally, we discuss research challenges to pave the pathway to Seb-based semantic communication. Our main contributions are summarized as follows.

- (1) We discuss the recent advancements in investigating SIT and developing Seb-based semantic communication architecture for wireless networks.
- (2) We explore the applications of semantic communications and highlight some benefits of semantic communication for new applications and use cases.
- (3) We outline the open problems and key challenges, and potential solutions to promote further research in semantic communications.

The outline of this review is shown in Figure 2. Section “The overview of semantic communications” presents overviews of semantic communications, including SIT and semantic communication architecture. Section “The applications of semantic communications” goes into the applications of semantic commu-

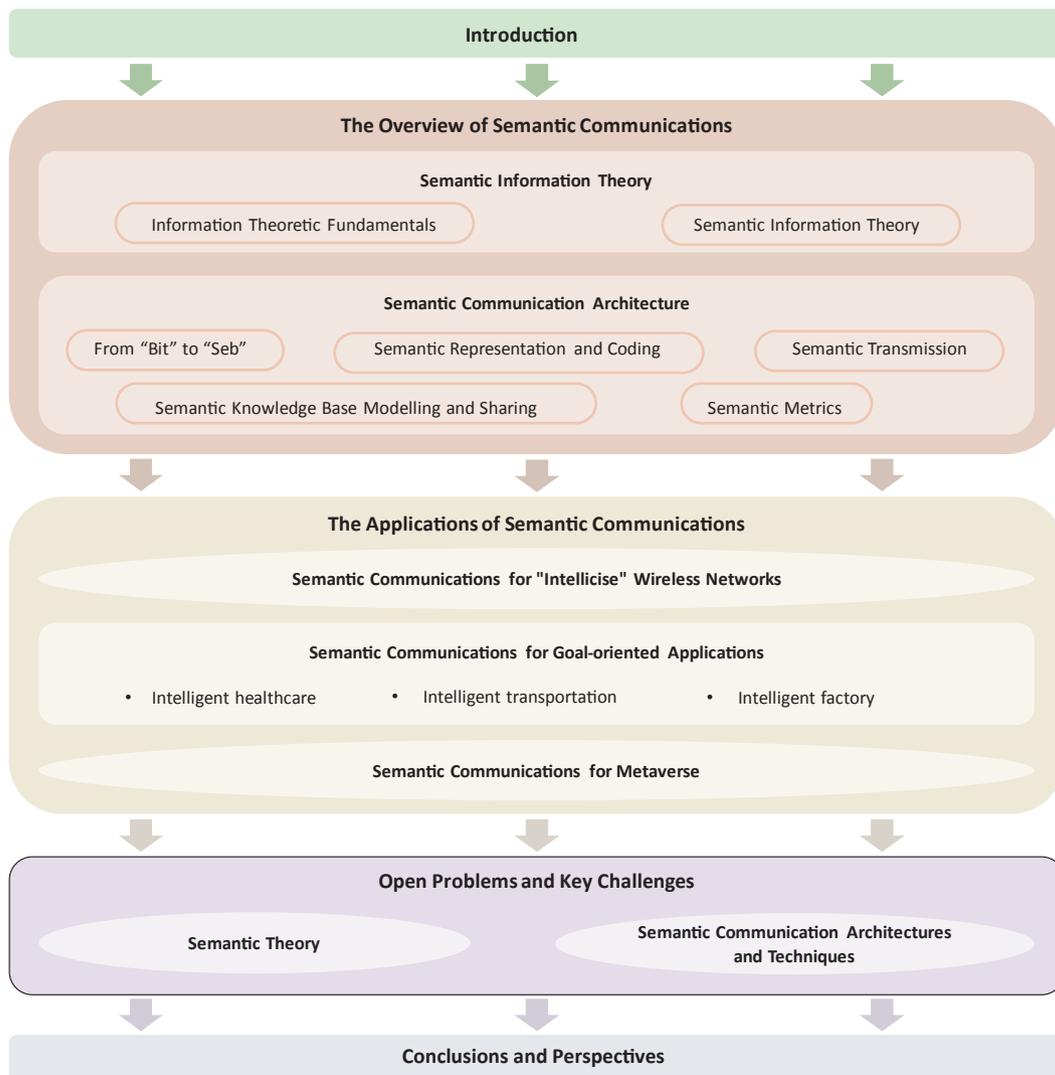


Figure 2 Roadmap of the overview.

nications, such as “Intellicise” wireless networks, goal-oriented applications, and Metaverse. The open problems and key challenges of developing semantic communications are outlined in Section “Open problems and key challenges”. Finally, Section “Conclusions and perspectives” concludes this review.

THE OVERVIEW OF SEMANTIC COMMUNICATIONS

In this section, we provide a detailed overview of SIT and the Seb-based semantic communication architecture. First, we explore the fundamentals of information theory and how it incorporates semantic information. We then discuss the evolution from traditional Bit-based communication to Seb-based semantic communication, including semantic representation and coding, semantic transmission, semantic knowledge base (SKB) modeling and sharing, and semantic metrics.

Semantic information theory

The concept of “semantic” is inherently ambiguous, which makes it challenging to provide a clear definition. In general, the concept of “semantic information” refers to the “meaning” or content conveyed by a message or communication. Semantic information has significant implications in multiple fields, including biology, cognitive science, and philosophy. In ICT fields, SIT mainly aims to build the fundamental boundaries of semantic communication systems. The most nucleus concern in the theory of semantic information is the meaning contained in the information [5]. Researchers have long been interested in formulating a broadly applicable and unified theory of semantic information. The development of SIT can be broadly categorized into two distinct views within the philosophy of information [30]. One perspective on SIT aims to quantify semantic uncertainty, similar to Shannon’s theory, by counting message occurrences to measure semantic-agnostic uncertainty. The second perspective considers the measurement of semantic similarity, which requires a novel approach to defining meaningful information.

(1) Information theoretic fundamentals

Shannon’s pioneering work on information theory provided answers to two fundamental questions in digital communication theory: what is the ultimate data compression achievable (the entropy H) and what is the ultimate transmission rate of communication (the channel capacity C). Shannon introduced the concept of entropy as a measurement of information. Utilizing probability theory, the amount of information conveyed in a message or signal is quantified in terms of bits. Information entropy not only provides a quantitative measure of the amount of information but also establishes the theoretical limits for data compression and transmission rate. In fact, the average code length for practical encoding cannot be less than the theoretical lower bound, which is given by the information entropy. For discrete variables, let $X \in \{x_1, x_2, \dots, x_n\}$ be a source with probability mass function $p(x_i), x_i \in X$. The entropy $H(X)$ of X is defined as follows:

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i). \quad (1)$$

The channel capacity refers to the maximum amount of mutual information that can be transmitted through the channel. Mutual information can be perceived as the reduction of uncertainty in one random variable due to the knowledge of the other. Considering the input X and the output Y , the channel capacity is defined by

$$C = \max_{p(x)} I(X; Y), \quad (2)$$

where $I(X; Y)$ is the mutual information between X and Y .

The lossy source coding theorem, known as the rate distortion theory, can be used to determine the minimal number of bits per symbol, measured by the rate R for a given distortion D^* . Similarly, the rate distortion function $R(D^*)$ represents the lower bound of the transmission rate for a given maximum average distortion D^* , which is defined by

$$R(D^*) = \min_{D \leq D^*} I(X; Y), \quad (3)$$

where $D = \sum_{x,y} p(x)p(y|x)d(x,y)$ is the distortion, $d(x,y)$ is the distortion metric with $d(x,y) = 0$ if $x = y$.

(2) Semantic information theory

Researchers have been developing methods to quantify semantic information for several decades. One approach is semantic entropy, which aims to measure semantic uncertainty similar to Shannon’s theory by

counting message occurrences to measure semantic-agnostic uncertainty. In 1952, Carnap and Bar-Hillel [7] first explicated the concept of semantic entropy of a sentence using logical probability rather than statistical probability. Given sentence s on the evidence e , they introduced the degree of confirmation $c(s, e)$ and gave the definition of semantic entropy as follows:

$$H(s, e) = -\log c(s, e). \quad (4)$$

The degree of confirmation $c(s, e) = \frac{m(e, s)}{m(e)}$, where $m(e, s)$ and $m(e)$ represent the logical probability of sentence s on the evidence e and that of the evidence e , respectively.

Later, Choi *et al.* [30] defined the semantic entropy of a message or sentence s as

$$H_S(s) = -\log_2(m(s)), \quad (5)$$

where $m(s)$ is the logical probability of s . This facilitates the calculation of the probability of a sentence or clause being true through logical probability.

In addition to logical probability, the definitions of semantic entropy based on the fuzzy sets are investigated. The meanings of information commonly used in daily life often possess fuzziness. For instance, words such as “tall”, “short”, “fat”, “thin”, “almost” and “nearly” are semantic descriptions that are fuzzy and ambiguous. While the statistical probability is insufficient for describing these concepts, fuzzy set theory can be used for qualitative and quantitative analysis. In particular, the characteristic functions of fuzzy sets are not binary functions, which introduces double uncertainty, both randomness and ambiguity. In 1972, De Luca and Termini [11] introduced the concept of probability information entropy onto fuzzy sets and gave the definition of fuzzy entropy. The authors in Ref. [31] defined semantic entropy using the membership degree of the semantic concepts. In Ref. [32], the authors used Rényi entropy [33] to measure semantic information and proposed a multi-grained definition of semantic information for different levels of communication systems. Besides measuring semantic information based on semantic uncertainty, some researchers have focused on measuring semantic similarity, which has led to new ways of defining meaningful information. Floridi [12] used the relative distance of semantics to measure the amount of information, which resolved the paradox in Carnap and Bar-Hillel’s proposal [8].

The definition of channel capacity, as provided by Shannon, refers to the upper bound of error-free data transmission that can be achieved over a noisy communication channel with a given bandwidth and noise level. Analogous to Shannon’s information theory, the semantic channel capacity C for discrete memoryless channels can be derived as [30]

$$C_S = \sup_{P(X|S)} \{I(X; Y) - H(S|X) + \overline{H(Y)}\}, \quad (6)$$

where $I(X; Y)$ is the mutual information between the set of transmitted messages X and the set of received messages Y , $H(S|X)$ measures the semantic ambiguity, $H(Y)$ denotes the average entropy of received symbol.

For semantic communications, the SKB is deployed to reduce semantic source entropy. D’Alfonso [10] exemplified the positive contribution of SKBs. Specifically, with the help of KB, even when the receiver is unable to decode a semantic message directly, the message could be correctly inferred given a set of logical relations, thus leading to possibilities of lossless semantic compression with fewer encoded bits. In that

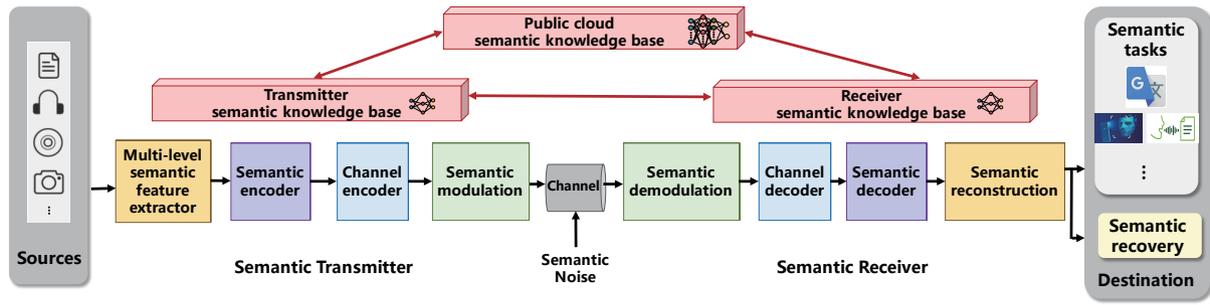


Figure 3 Semantic communication architecture.

regard, Choi *et al.* [30] further investigated the uncertainty and quantification of SKB. Specifically, they defined the knowledge entropy of a knowledge base as the uncertainty of answers it derives. Therefore, the knowledge entropy is expressed as the average semantic entropy of a query (message) x computable from the knowledge base K as follows:

$$H(K) = \frac{1}{|K|} \sum_{x \in K} H(x), \quad (7)$$

where $H(x)$ is the semantic entropy of the message x and $|K|$ denotes the size of that knowledge base.

The lossy source coding theorem states that when the source code rate is greater than $R(D^*)$, a source code can always be found whose average distortion does not exceed D^* . Among them, $R(D^*)$ is the minimum average mutual information between the input and output under the fidelity criterion. Referring to the lossy source coding theorem, the concept of semantic channel capacity with average semantic distortion and limited semantic distortion has been proposed. The rate-distortion function in the semantic communication system can be defined as [34]

$$R(D_s, D_a) = \min I(S; \bar{X}, \bar{S}), \quad (8)$$

where D_s is the semantic distortion between source X and recovered information \bar{X} at the receiver, and D_a is the distortion between semantic representation S and received semantic representation \bar{S} .

Recent research on SIT shows great promise for semantic communication; however, it is still evolving and is not yet a complete theory. Particularly, the SKB is a critical component in the reasoning process, research on SKB is currently scattered, and key issues such as the impact of a shared SKB on communication and the quantitative modeling of semantic flow in partially-shared SKBs require further exploration. In addition, research on general SIT, including semantic security and robustness, efficiency and generalization trade-offs, and semantic computing, will be important areas for future research.

Semantic communication architecture

As previously discussed in the introduction, conventional communication systems mainly focus on the technical level, which is only the first level of the three levels of communication. To further improve the communication efficiency, semantic communication has emerged as a promising solution and gained significant attention. As depicted in Figure 3, the architecture of semantic communication generally comprises the following crucial components.

The Seb-based SKB SKB is a fundamental component of semantic communications, which is an inherent knowledge network model that provides relevant semantic knowledge descriptions for communication participants such as transmitting and receiving users [35]. The SKB is constructed and shared by both the transmitter and receiver, and it supports the semantic encoding and decoding of source messages in semantic communications. For end-to-end semantic communication, the SKB can be categorized into three types: source, channel, and task knowledge bases. These knowledge bases provide a multi-level semantic knowledge representation of source data, propagation environment, as well as task requirements, respectively. Furthermore, in the context of semantic-oriented wireless networks, SKB can also be divided into two types, namely private local SKB and public cloud SKB. It is noteworthy that different users and devices may possess diverse SKBs, which are influenced by various factors such as background, environment, communication history, etc. Moreover, each user maintains and updates their own private SKB to store individualized local semantic knowledge and information. The public cloud SKB also requires dynamic updates and consensus among the majority of users to share knowledge updates.

The semantic transmitter The semantic transmitter comprises several key modules, including the multi-level semantic feature extractor, semantic and channel encoder, semantic modulation module, etc. The multi-level semantic feature extractor processes and extracts useful and relevant information from various types of source data based on the knowledge descriptions in its SKB. The objective of the semantic feature extractor is to maximize the compression of semantic information by identifying and consolidating similar semantics, eliminating redundant semantics, and correcting corrupted semantics. As different semantic features correspond to different levels of importance at the receiver, the semantic and channel encoder compresses and removes irrelevant information, while processing and providing unequal protection based on the channel conditions. Finally, the semantic messages are modulated by the transmitter into a signal for transmitting reliably over a noisy channel.

The semantic receiver The semantic receiver typically comprises several components, such as the semantic demodulation module, semantic and channel decoder, and semantic reconstruction module, among others. Typically, the process of semantic demodulation is the inverse of semantic modulation. The semantic and channel decoder alleviates physical and semantic noise in the received signal transmitted over an unreliable channel and then recovers the multi-level semantic features. Finally, the reconstructed message is obtained by combining the different levels of semantic features using the reconstruction module.

The semantic noise Semantic noise refers to interference that can distort the process of semantic communication, leading to errors in correctly identifying and interpreting semantic information. The semantic noise can be introduced during the process of semantic communication, including encoding, transmission, and decoding. During the encoding stage, semantic noise may arise due to the incorrect identification of entities and their relationships in the signal at the transmitter. Additionally, channel fading and noise can also cause the loss of transmitted semantic messages, resulting in semantic distortion. In the decoding stage, semantic noise can be generated due to errors in the interpretation of semantic information by the receiver or misunderstandings by users. Therefore, it is crucial to minimize semantic noise in order to ensure accurate and effective semantic communication.

The environment semantics The environment semantic information conveyed in wireless channel environments can be leveraged to reduce the cost of acquiring channel state information (CSI), which is crucial for the efficient implementation of communication systems. By taking into account the prior knowledge of

the wireless environment and channel, it is possible to represent the environment at both macro and micro scales, catering to the information needs of different channel characteristics. The large-scale level includes layout and global environment representations for large-scale parameters (LSP), while the small-scale level comprises local and target representations for small-scale parameters (SSP), and line-of-sight (LOS) blockage. In a recent study [36], the concept of propagation environment semantics (PES) was proposed, which is a set of propagation environment semantic symbols (PESSs). The PESSs are deconstructed environment representations at a semantic level and can be utilized for the prediction of wireless environments.

In order to facilitate a semantic communication architecture, we will focus on key methods such as semantic representation and coding, semantic transmission, and SKB modeling and sharing. In addition, we discuss semantic metrics to evaluate the effectiveness of these methods.

(1) From “Bit” to “Seb”

Considering the concept of “bit” often utilized in existing syntactic-based communications, a foundational framework for the representation and quantification of semantic communications is also essential. In this context, Seb was proposed in Ref. [13], typically referring to a foundational or fundamental structure that holds and organizes the meaning of symbols, words, or concepts. Seb acts as the cornerstone of a semantic space, and its integration with this semantic space facilitates a concise representation of diverse multimodal communication sources. Sebs are typically depicted as vectors, containing both explicit and implicit semantic features of the source. This vectorized representation aligns with the notion of model-based transmission, wherein AI semantic models are transferred. This process inherently involves the dissemination of AI-driven semantic capabilities.

Generally, given a type of source information, the generation of Sebs should also consider the specific communication intents and the semantic knowledge. These factors determine the granularity of the Sebs. For an example of communication intents, more Sebs are required for image reconstruction-oriented communication than for image classification task-oriented communication. This is because the former requires a more detailed representation of the entire image, whereas the latter may be specific to classification only. As the basic unit of semantic representation, the generation of Seb is still an unsolved issue. Zheng *et al.* [37] proposed Seb-based semantic communication framework for image transmission, where Sebs are generated for images and employed at both transmitter and receiver, significantly enhancing the transmission efficiency.

(2) Semantic representation and coding

Semantic representation involves the interpretation and logical representation of the meaning conveyed by source symbols. In future mobile communication systems, making full use of the source/channel characteristics to extract semantic information can further reduce the coding rate and effectively improve spectrum utilization. Traditional signal processing methods, such as Fourier transform, wavelet transform, and discrete cosine transform (DCT), are based on the transformed basis that has no relation to semantics, thereby separating the internal relationship of different modal signals.

Benefiting from the rapid development of AI technology, researchers have made efforts to model the semantic features of information sources by leveraging the feature learning and representation capabilities of DL models. The fundamental idea of this approach is to pre-train a semantic representation model (Pre-Training Model, PTM) on a large-scale and readily available dataset to address general tasks and subsequently fine-tune the trained semantic representation model for specific tasks. This approach offers two

advantages: firstly, the semantic model trained on general tasks possesses robust representation capabilities, making it widely applicable; secondly, when the pre-trained model is adapted to specific tasks, only a small amount of labeled data is needed to achieve optimal model performance. Based on this foundation, it has become possible to extract semantic information from different modal sources, including text, audio, image, video, etc.

1) Semantic representation and coding for text The semantic information of text refers to grammatical information, word meanings, logical expressions among words, etc. The conventional approach to semantic analysis in NLP is grammar-driven, such as using concrete syntax trees. Griffiths *et al.* [38] proposed a probabilistic approach to semantic representation, which models the probability of word occurrences in different contexts explicitly. Hinton *et al.* [39] developed the stochastic neighbor embedding (SNE) method, which aims to preserve the local structure of high-dimensional data in a lower-dimensional space by modeling the probability distribution of pairs of neighboring points. The recent advancements in NLP have facilitated the development of semantic representation for text. Word2Vec is a class of models for generating word vectors [40] that use neural networks to learn word vector representations from large text corpora. Based on this idea, Pennington *et al.* [41] proposed the GloVe model, and Bojanowski *et al.* [42] proposed the fastText word vector model. However, these methods belong to context-independent static word vector representations, which fail to describe syntax information. To capture the aspects of word meaning that are dependent on context, the ELMo model was proposed in Ref. [43], which used vectors derived from a bidirectional long short-term memory (LSTM) network for language modeling. The Transformer model with a self-attention mechanism was proposed in Ref. [16], which extracts characteristics of input sentences in parallel, achieving lower computational complexity and paving the way for a new era of advancements in NLP. Building on the Transformer model, more powerful pre-trained language models such as GPT [17] and BERT [18] have been proposed.

Recently, many researchers use the powerful feature extraction capabilities of DL to develop joint semantic source-channel coding. The authors in Ref. [44] developed a neural network architecture for joint source-channel coding of text that uses GloVe model to extract semantic information. The authors in Ref. [45] proposed a semantic communication system based on DL, which utilizes the Transformer model with a fixed attention mechanism for transmitting textual information. Additionally, a flexible semantic communication system based on the universal Transformer (UT) was proposed in Ref. [46], which incorporates an adaptive recurrence mechanism into the Transformer to overcome the limitations of the original fixed structure. The authors in Ref. [47] proposed an ECSC system that utilizes context information within and between sentences, incorporating self-attention, segment-level relative attention, a gate mechanism, and Transformer-XL for semantic representation and recovery in text transmission.

2) Semantic representation and coding for audio The semantic information of audio refers to the meaning or content conveyed by the audio signal, such as speech, music, environmental sounds, or any other acoustic events. It involves understanding and analyzing the audio signal to extract relevant information, such as speech recognition, speaker identification, or sound event detection, among others. Similar to the semantic representation and coding for text, the DL-based solutions for extracting and coding semantic features in audio data have been explored. With the goal of reducing the amount of transmitted data, most of these studies focus on transmitting only the text-related semantic features extracted from speech signals, while ensuring that the text transcription of the speech signals can be reconstructed at the receiver. The authors in

Ref. [48] proposed an end-to-end semantic communication system for speech signals that uses a DL approach and an attention mechanism with squeeze-and-excitation (SE) networks. In Ref. [49], an autoencoder based on the Wav2Vec architecture was proposed, which includes two convolutional neural networks (CNNs) in the encoder, namely feature decomposer (FD) and audio generator (AG), respectively. In Ref. [50], an end-to-end DL-based transceiver was proposed for semantic speech transmission. The transceiver extracts and encodes semantic information for speech recognition tasks and produces transcriptions at the receiver. The model includes soft alignment and redundancy removal modules to reduce semantic redundancy, a semantic correction module to correct transcriptions, and a connectionist temporal classification (CTC) alignment module to extract additional speech-related information, thereby enhancing speech reconstruction.

3) Semantic representation and coding for image/video With regard to image/video sources, semantic representation and coding typically involve the extraction of semantic information and the generation of images based on deep semantic representation. With the emergence of large-scale datasets and the improvement of computing power, data-driven neural network models have shown huge potential in the extraction of semantic information, such as AlexNet [51], VGG [52], ResNet [53], EfficientNet [54], etc. In terms of image generation based on deep semantic information, the VAE model [55] uses the Kullback-Leible (KL) divergence to measure the similarity between the generated distribution and the real distribution through variational approximation. The generative adversarial network (GAN) proposed by Goodfellow *et al.* [56] generates images through the zero-sum game between the generator and discriminator, producing images with highly similar syntactic information to real images. Similar to the semantic representation and coding for text and audio sources, DL-based semantic feature extractions and joint source-channel coding also received significant attention in the context of images/videos. Bourtsoulatzé *et al.* [57] proposed a joint source and channel coding approach for wireless image transmission. Yang *et al.* [58] proposed a wireless image transmission scheme using adaptive deep joint source-channel coding, which utilizes a policy network to optimize the trade-off between signal quality and transmission rate. Zhang *et al.* [59] proposed a multi-layer semantic-aware communication system that incorporates a feature extractor capable of extracting multiple levels of semantic information. Additionally, Dai *et al.* [60] proposed a method that utilizes non-linear transformations to extract semantic features from the source, which can then be used as side information for source-channel coding.

4) Semantic representation and coding for multimodal data The transmitted information may include one or more modalities and involve conversions between different modalities. By exploring the relationship between different modalities, complementary information can be combined and redundant information can be eliminated. The authors in Ref. [61] studied a semantic communication system for visual question answering (VQA), where users transmit images and texts to obtain information about the images. Like previous work on image and text transmission, the image transmitter utilizes a pre-trained ResNet-101 network, while the text transmitter employs a Bi-LSTM network. However, further investigation is needed to design the decoder that merges the correlated information from both users and answers the visual questions. The authors in Ref. [62] presented a unified semantic encoding framework for both image and text transmitters using Transformer and introduced a new semantic decoder network that includes a query module and an information fusion module. To address the issue of updating models when tasks change or multiple models need to be stored, the authors in Ref. [63] proposed a unified DL-enabled semantic communication system (U-DeepSC) that can handle multiple tasks with various modalities. A multi-exit architecture in U-DeepSC

provides early-exit results for simple tasks, and a unified codebook for feature representation reduces transmission overhead by transmitting only indices of task-specific features. The authors in Ref. [64] proposed a cooperative task-oriented communication method for transmitting multi-modal data from multiple end devices to a central server. The proposed method utilizes the transmission results of low-rate modalities to control the transmission of high-rate modalities, aiming to reduce the transmitted data amount. Taking a human activity recognition (HAR) task as an example, the proposed method is evaluated for monitoring videos and acceleration data in a smart home environment.

(3) Semantic transmission

Conventional physical modules are typically optimized independently. For example, modulation and signal detection are designed to minimize bit-error-rate (BER), and channel estimation is aimed at minimizing the mean-squared error (MSE) between estimated and real channels. However, for semantic communication, transmit features have different importance levels and are semantically correlated, which can be exploited for semantic recovery. In the process of semantic transmission, the semantic transmitter first extracts the semantic-related features of the source message, encodes and modulates the semantic information properly with the support of SKB, considering both source semantic and channel characteristics to combat interference and noise during transmission. Then, the semantic receiver performs semantic demodulation, semantic decoding, and semantic fusion while deciding whether to either reconstruct the source message or directly execute intelligent tasks based on the users' requirements. In the following, we discuss semantic modulation, semantic hybrid automatic repeat request (HARQ), channel estimation, and semantic-aware resource allocation in order to exploit the inter-correlation among features and optimize the overall system performance in semantic communication.

1) Semantic modulation Unlike conventional modulation techniques such as quadrature amplitude modulation (QAM), which are content-unaware and treat all bits or features equally, semantic modulation takes into account the importance and semantic correlation of transmitted features or bits. Although DL-based solutions have been well applied to semantic representation and coding and used to replace the conventional separate source and channel coding/decoding modules, it is still challenging to adopt DL-based modulation for wireless communications. DL-based encoders usually output real numbers, which need to be mapped to discrete constellation symbols. However, this mapping is equivalent to a non-differentiable function, making it difficult to apply application neural networks directly. There are two main methods for implementing semantic modulation in wireless communications. The non-neural-network-based approaches usually employ a quantizer to convert the encoder's continuous output into a sequence of discrete symbols that can be transmitted through digital communication systems. The neural network-based methods typically use an extra neural network to generate the likelihood of constellations. The authors in Ref. [65] proposed a joint coding-modulation (JCM) method that utilizes a neural-network-based stochastic encoder along with a random coding-based modulator. The neural network is trained to learn the transfer probability from source data, while the random coding generates the actual modulated symbols based on the learned transfer probability. Neural-network-based methods have shown promise in semantic modulation, but there are still open issues to address and they may not fully optimize overall system performance.

2) Semantic HARQ The use of HARQ with acknowledgment (ACK) feedback is crucial for successful semantic transmission in varying channel conditions. However, conventional HARQ methods that use channel coding for forward error correction (FEC) and cyclic redundancy check (CRC) for error detection are

designed at the bit level, whereas semantic similarity is more important than bit errors in semantic transmission. The authors in Ref. [66] proposed an end-to-end deep source-channel coding scheme for sentence semantic transmission using HARQ to improve the reliability and efficiency of the semantic transmission by retransmitting only the corrupted parts of the received message. Furthermore, the semantic transmitter can adaptively carry different amounts of semantic content according to the channel information. The authors in Ref. [67] proposed an adaptive bit rate control scheme for semantic communication systems with incremental knowledge-based HARQ. The proposed scheme adjusts the bit rate of the encoder based on the HARQ feedback information, allowing the encoder to generate a bit stream with a lower or higher rate depending on the channel conditions. This approach helps to achieve a balance between the transmission rate and the error rate, thus improving the overall system performance.

3) CSI Feedback CSI plays a critical role in communication systems, but the feedback of CSI consumes a substantial amount of valuable transmission resources. Data hiding is a potential DL method for semantic transmission and is exploited to remove transmission payload in CSI feedback. The authors in Ref. [68] proposed a DL-based hiding framework for downlink CSI acquisition to eliminate the feedback overhead. The CSI is embedded in digital images, which has minimal impact on image semantics. Additionally, CSI feedback can help the semantic channel coding to assign important information to the subchannels with high signal-to-noise ratios (SNRs). By treating the channel matrices as images, CNN-based solutions are effective for channel estimation and CSI feedback. Xu *et al.* [69] proposed a deep joint source-channel coding framework for the CSI feedback, including a non-linear transform method to compress CSI and an SNR adaption mechanism to adapt to wireless channel variations. The authors in Ref. [70] proposed an adaptive CSI feedback scheme for precoding, which improves effectiveness by adjusting the feedback overhead. Specifically, they developed a performance evaluator to predict the reconstruction quality of each image. This enables the proposed scheme to adaptively decrease the CSI feedback overhead for transmitted images with high predicted reconstruction qualities in the joint source-channel coding system.

4) Semantic-aware resource allocation Initial work on resource allocation in semantic-aware networks optimizes resource allocation for semantic communications based on specific tasks and scenarios. In the context of text semantic communication, semantic spectrum efficiency was defined, and the joint optimization resource allocation in terms of channel assignment and the number of transmitted semantic symbols was studied in Ref. [71]. To enhance the quality of user experience (QoE) for different users and services, the authors in Ref. [72] further designed a semantic-aware resource allocation algorithm for multi-unit and multi-task networks, which achieved optimal QoE by adjusting the transmission of the number of transmitted semantic symbol, channel allocation, and power allocation. Knowledge graphs were used to extract semantic features from text data in Ref. [73], where the semantic similarity was taken as the performance evaluation index of the semantic system. They jointly optimized the resource block allocation and the transmission features using the sum of the semantic similarity of all users in the network as the optimization objective. For image semantic communication scenarios, Liu *et al.* [74] considered the importance of semantic features to establish a resource allocation model and maximized the success probability of tasks by optimizing bandwidth allocation, power, and semantic compression rate. The authors in Ref. [75] studied the joint optimization problem of semantic compression rate, bandwidth resources, and transmission powers. Considering the personalized features of users in semantic communication, the authors in Ref. [76] proposed a semantic information processing framework based on user interest and studied the power allocation problem to

improve semantic communication performance using game theory. The authors in Ref. [77] proposed a new semantic transmission performance metric, i.e., system throughput in the message, and then used a Transformer-based model to obtain the mapping relationship between the bit transmission rate and this performance metric. They constructed a resource allocation model to optimize base station selection and bandwidth allocation. The authors in Ref. [78] studied the joint optimization problem of user selection, communication resource arrangement, and semantic feature selection. They proposed an algorithm based on multi-agent reinforcement learning to minimize the delay of semantic information transmission. Moreover, to jointly optimize the communication and computational resources, the authors in Ref. [79] studied the problem of semantic information extraction and computing resource allocation based on rate splitting in semantic communication systems.

(4) Semantic knowledge base modeling and sharing

The SKB is an important module introduced by semantic communication. It provides guidance for semantic information processing at both the transmitter and receiver, and the shared SKB between them enables semantic transmission. The SKB not only “perceives” the semantic features of the source, but it also relates to specific transmission tasks and conditions, enabling unequal transmission in the presence of CSI. Additionally, slight errors at the semantic level can be easily identified and corrected using the existing knowledge in the SKB. As mentioned previously, the SKBs for wireless communication include source, channel, and task SKB. In the following, we deeply discuss semantic knowledge modelling as well as semantic knowledge updating and sharing in terms of source, channel, and task SKBs.

1) Semantic knowledge modelling SKBs based on knowledge graphs, labeled training data sets, and feature vectors have been applied to end-to-end semantic communication and have shown promising results. For text transmission, the authors in Ref. [80] constructed the SKB as a set of semantic triples, which are used by the sender and receiver for encoding and decoding the text information, respectively. Additionally, the authors in Ref. [81] extracted semantic triples from the text source and calculated their importance for transmission based on the quality of the channel state. For speech transmission, the authors in Ref. [82] defined the SKB as a multi-level knowledge graph and proposed a construction method based on semantic representation and abstraction. For image data transmission, the authors in Ref. [83] proposed a multi-layer semantic representation method and a collaborative reasoning mechanism for heterogeneous networks enabled by multi-layer SKB. Moreover, for multi-task requirements and multi-modal data sources, the authors in Ref. [63] proposed a cross-task shared SKB consisting of discrete semantic basis vectors and joint training with semantic-channel coding, which can reduce transmission overhead and model size while achieving comparable performance to task-specific semantic communication frameworks.

2) Semantic knowledge updating and sharing The efficient implementation of semantic communication relies on a good match between the SKB at the sender and that at the receiver. Furthermore, the SKB should be kept adaptive to the time-varying data information, including source, task, as well as channel environment data. Therefore, it is necessary to design a dynamic updating and sharing method for SKBs of semantic transmission. In semantic communication, users automatically learn and update their local knowledge base during the communication process. Additionally, each user may also maintain and update their private SKB dynamically to store their unique or partially shared private semantic knowledge and information. The public SKB also has dynamic updating capabilities but requires consensus among most users to share knowledge updates.

Table 1 Semantic metrics for different source types

Source type	Semantic communication metric	Meaning	Refs.
Text	Word error rate (WER)	The ratio of the total number of incorrectly recognized words to the total number of words in the ground truth transcript	[84]
	Bilingual evaluation understudy (BLEU)	The similarity between the machine-translated text and one or more reference translations	[85]
	Bidirectional encoder representations from transformers (BERT)-based semantic similarity	The semantic similarity between two pieces of text based on BERT language model	[18]
	Semantic similarity metric (SSM)	The semantic similarity between the transmitted sentence and the estimated sentence	[76]
	Average bit consumption per sentence	The average number of bits consumed per sentence in the wireless text transmission process	[58]
Speech	Signal-to-distortion ratio (SDR)	The ratio of the signal power to the distortion power	[86]
	Perceptual evaluation of speech quality (PESQ)	A measure of the perceived quality of speech after being transmitted through a communication channel	[87]
	Fréchet deepspeech distance (FDSD) Kernel DeepSpeech Distance (KDSD)	An evaluation for the quality of synthesized speech signals, indicating the similarity between the reconstructed speech signal and the original signal	[88,89]
	Multiple stimuli with hidden reference and anchor (MUSHRA)	The overall quality of the speech source from a human perception perspective	[90]
Image	Peak signal-to-noise ratio (PSNR)	The accuracy of the reconstructed image in terms of pixel accuracy	[91]
	Structural similarity (SSIM)	The accuracy of the reconstructed image in terms of structural similarity	[92]
	Learned perceptual image patch similarity (LPIPS)	The perceptual similarity between two images	[93]
	mean intersection over union (mIoU)	The accuracy of a model by comparing the intersection and union of the model's output results and annotated results	[94]
	Image semantic similarity (ISS)	The similarity between two images based on their semantic content	[70]
Video	Motion-based video integrity evaluation (MOVIE)	The distortion in video caused by motion errors	[95]
	Fusion-based video quality assessment (FVQA)	A comprehensive evaluation combined several quality indicators to judge the overall quality of the video	[96]
	Video quality metric (VQM)	The measure for the quality of video based on perceptual image quality factors	[97]
	Video quality model for variable frame delay (VQM_VFD)	A variation of VQM that takes into account variable frame delays in the video	[98]
	Video multi-method assessment fusion (VMAF)	A metric combines multiple objective quality metrics to predict subjective quality scores	[99]

(5) Semantic metrics

Semantic communication emphasizes the “provisioning of the right and significant piece of information to the right point of computation (or actuation) at the right point in time” [3]. Therefore, the quality of semantic communication is closely related to the characteristics of the information source and focuses on indicators of transmission effectiveness rather than just accuracy at the bit or block level. As shown in Table 1 [18,58,70,76,84–99], we discussed the semantic metrics for different source types.

1) Semantic metrics for text sources The evaluation of semantic communication quality for text sources includes objective metrics such as WER and BLEU, as well as subjective metrics such as BERT-based semantic similarity using deep neural networks (DNNs). WER measures text reconstruction accuracy by calculating the percentage of words that are incorrectly recognized by the system compared with the ground

truth transcript. However, WER cannot account for synonyms or semantic similarity. BLEU is a commonly used metric for evaluating the quality of machine translation output and calculates the similarity between the machine-translated text and one or more reference translations. BERT-based semantic similarity, based on bidirectional Transformer DNNs and cosine similarity, provides scores that are more closely aligned with subjective perception compared with objective metrics like WER and BLEU. In order to account for the linguistic ambiguity where a word may have different meanings in different contexts, the authors in Ref. [76] introduced an SSM to measure the degree of semantic similarity between the transmitted and estimated sentence. The authors in Ref. [66] proposed a wireless text semantic communication metric to evaluate the system's performance from a communication perspective, which measures the average bit consumption per sentence. They also applied this metric to evaluate their proposed text semantic transmission with HARQ.

2) Semantic metrics of speech sources The evaluation of semantic communication quality for speech sources encompasses both objective and subjective metrics. Objective metrics such as SDR, PESQ, FDSO, and KDSO are used to evaluate the perceptual quality of speech. The SDR is the ratio of the signal power to the distortion power. PESQ is a widely used metric that measures the perceived quality of speech after being transmitted through a communication channel. It is used to simulate mean opinion score (MOS) subjective ratings of speech and is composed of the PESQ and a perceptual analysis measurement system. The scoring range is from -0.5 to 4.5 . FDSO and KDSO are metrics that evaluate the quality of synthesized speech signals, indicating the similarity between the reconstructed speech signal and the original signal. The lower the score, the higher the similarity between the reconstructed speech signal and the original signal. On the other hand, subjective metrics like MUSHRA are used to assess the overall quality of the speech source from a human perception perspective. It uses a blind listening test to score the subjective quality of a reconstructed speech signal by setting several degraded speech signals, including a lossless speech signal as an upper reference and a fully degraded speech signal as a lower reference. The subjective ratings are averaged to obtain the subjective quality evaluation of the reconstructed speech signal. In addition, in cross-modal tasks such as speech-to-text, metrics such as WER and BLEU for text can also be used as evaluation metrics for speech communication quality.

3) Semantic metrics for image/video sources The evaluation of semantic communication quality for image sources includes objective metrics and subjective evaluation metrics. Objective metrics such as PSNR and SSIM measure the accuracy of the reconstructed image in terms of pixel accuracy and structural similarity, respectively. Additionally, subjective evaluation metrics such as LPIPS and mIoU, focus on semantic similarity based on DNNs and subjective perception, respectively. The authors in Ref. [78] proposed the use of ISS to evaluate the performance of cooperative image semantic communication networks by capturing the correlation between the image's meaning and its corresponding semantic information. For video sources, quality evaluation of semantic communication involves metrics used for image sources to evaluate the semantic communication quality of each frame and video quality assessment metrics such as the MOVIE index, FVQA, VQM, VQM_VFD, and VMAF. In addition, the accuracy of downstream task indicators is also considered in evaluating the quality of semantic communication for video sources.

4) Other semantic metrics Some recent studies go further to explore the significant aspect of semantic information, such as age of information (AoI) [100] and value of information (VoI) [101]. In Ref. [102], the authors investigated the impact of age of correlated features on real-time supervised learning tasks. In Ref. [103], the authors proposed the concept of age of semantics (AoS), which measures the freshness of

semantics as the time duration since the perfect inference of the current process status. In addition, to optimize semantic communication system performance, some semantic metrics have been discussed. The authors in Ref. [71] defined semantic transmission rate as the amount of semantic information effectively transmitted per second. A definition of semantic spectral efficiency was also proposed in Ref. [71], which quantifies the rate at which semantic information can be transmitted over a given bandwidth with successful transmission. The authors in Ref. [77] proposed system throughput in message to represent network performance in the broader context of intelligent semantic communication networks.

Summary and lessons learned

Lessons learned for Seb-based SIT and semantic communications The definition and measurement of semantic information have not yet achieved a consensus, and the theoretical framework for semantic information is still in an infancy stage. Nevertheless, these theoretical researches pave the way for the research on semantic coding, semantic communication architecture design, semantic metric design, etc. The generation of Sebs is only the first step, which is followed by the selection of Sebs, where semantic theory is expected to serve as a quantitative method to evaluate the efficiency of the generated Sebs. Besides, to achieve intent-oriented transmission, the tradeoff between the efficiency and scalability of Sebs is also worth further discussion.

Lessons learned for semantic representation and coding Semantic representation and coding are essential components for facilitating efficient semantic communication by extracting features from redundant data and preserving their meaning [104]. Despite the significant progress made in this field, there is still no consensus on the unified theoretical framework to guide the representation and encoding of semantic information. Currently, semantic representation coding processes semantic information using techniques such as deep learning and neural networks, and it allows for the extraction of semantic features from multimodal data and the effective transmission of these features. It is crucial to ensure the correct interpretation of information and semantic consistency, which guarantees the effectiveness of semantic communication systems. Additionally, the complex design of DL-based models leads to extra time consumption and computational resource consumption during the training process. The establishment of a universal theoretical framework for semantic information and the development of efficient semantic representation and coding modules remain open issues that require further attention.

Lessons learned for SKB modelling and updating SKBs can provide relevant semantic knowledge descriptions for source information, and can be modeled based on knowledge graphs, labeled training data sets, and feature vectors for end-to-end semantic communications. However, existing SKBs, which are primarily designed for text and image data, often fall short when dealing with diverse data types. Moreover, the dynamic nature of wireless communication environments necessitates the matching of SKBs to adapt to evolving information. In this context, it is imperative to focus on SKBs' modelling and updating techniques, emphasizing effective information extraction, semantic knowledge updates, and the development of robust strategies. These measures are essential to ensure that SKBs remain adaptable and effective within the dynamic communication settings that characterize future wireless networks. With multi-modal data sources, diverse tasks, and dynamic channels becoming the norm, there is an increasing need for the development of efficient multi-agent SKBs. Achieving a deep integration of AI into communication technology and the

innovative construction of a multi-level SKB framework continue to pose critical challenges that need to be addressed to unlock the full potential of these advanced communication systems.

Lessons learned for semantic transmission The semantic-driven transmission paradigm departs from traditional methods by encoding messages into semantic features and adapting semantic waveforms to dynamic channel conditions. However, there are still technical challenges, such as semantic noise, semantic distortion, and intricate resource allocation. For fast adaptation in changing transmission environments, transfer learning techniques can be harnessed within the AI semantic transmission model. These techniques enable the model to converge rapidly after a few training iterations when faced with new channel conditions. This approach ensures stable and efficient semantic information transmission in intelligent channel environments. Moreover, considering the semantic distortion in transmission process, the incorporation of the HARQ solution should be well investigated. Finally, the application of reinforcement learning provides a promising approach for dealing with resource allocation problems of high complexity.

Lessons learned for semantic metrics As the integration of AI and communications continues to evolve, the evaluation metrics attach significant importance in terms of designing semantic communications. State-of-art work utilizes specialized metrics to evaluate system performance since the meaning of information, the effectiveness of the system, and the significance of semantic features attach different levels of importance in various applications. However, it is hard to compare the performance or effectiveness of semantic communications when they are characterized by different metrics. Therefore, if possible, further exploration of unified metrics for semantic communications is needed for system optimization.

Semantic communication is a promising paradigm that can revolutionize wireless communication systems. However, there are still limitations in the research landscape, including an ambiguous definition of a semantic communication system, and a lack of scalable frameworks for developing semantic communication networks. The development of more rigorous and well-defined technical foundations is necessary to fully realize the potential of semantic communication.

THE APPLICATIONS OF SEMANTIC COMMUNICATIONS

As the core representative technology of the inherent concision of information, semantic communications play a fundamental role in future wireless networks and ecosystem. In this section, we discuss the applications of semantic communications, including “Intellicise” wireless networks, goal-oriented applications, and Metaverse. “Intellicise” wireless networks form an intelligent ecology, reaching a state of self-optimization, self-balance, and self-evolution for wireless communications [13]. Semantic communications play a fundamental role in constructing and developing “Intellicise” wireless networks by enhancing the efficiency of signal processing, meaning delivery, and network management. Moreover, semantic communication and “Intellicise” wireless networks will open a new era of goal-oriented services and applications, encompassing domains like intelligent healthcare, intelligent transportation, and intelligent factories. In such cases, the goal of communication among machines is not always to reconstruct the exact message but to facilitate the receiver to make the correct inference, decision, and action at the appropriate time and within the correct context [3]. Finally, the concept of the Metaverse has grown in popularity as “the successor to the mobile Internet”. Metaverse over wireless networks is an emerging use case of 6G [105]. Semantic communications

offer a feasible way of reducing end-to-end latency by ignoring irrelevant details and transmitting the information after understanding. Additionally, semantic-aware “Intellicise” wireless networks have become increasingly important in supporting the distributed interaction requirements of terminals and clouds in the Metaverse, enabling real-time synchronization.

Semantic communications for “intellicise” wireless networks

The upcoming 6G networks will facilitate the integration of physical and digital spaces, leading to a significant expansion of connections and the emergence of innovative applications with diverse service demands. These changes will undoubtedly pose formidable challenges to the future network. To address these challenges, the concept of “Intellicise”, which stands for intelligence-endogenous and primitive-concise, was first proposed in Ref. [106]. The “Intellicise” network emphasizes the use of native intelligence and endogenous simplified architecture to achieve system optimization. With nodes possessing intelligence, the network itself will gradually evolve towards a native intelligent system, ultimately achieving a state of self-evolution, self-optimization, and self-balance. Therefore, the development of “Intellicise” wireless networks is essential to support the integration of physical and digital space, facilitate efficient coordination between various nodes, and enable the provision of intelligent services that can meet the changing demands of the future network posed by the upcoming 6G networks. In this context, semantic communication has shown great potential in extracting and transmitting the meaning of information and is envisioned as a promising technology for “intellicise” wireless networks. In the following, we will explore the benefits of the implementation of semantic communication for “Intellicise” wireless networks.

Semantic communications for “Intellicise” transmission The significant advancements made in ICT technology have led to the generation of large amounts of data in various formats in future wireless networks. In addition, the explosive growth of connections leads to the rapid expansion of the information space. For example, in scenarios where multiple nodes interact to accomplish a task, real-time sensing data exchange, information fusion, and collaborative decision-making among these nodes may increase the network complexity in terms of signaling costs and protocol overhead. Additionally, the interaction between human-machine-thing and its corresponding digital world in 6G services necessitates collaboration on multimodal information, putting an urgent demand for “Intellicise” transmission and processing. In such cases, semantic communication can reduce the amount of data transmission required by traditional communication methods by conveying concise and effective semantic information. This reduces the transmission overhead while maintaining the accuracy of information transmission. Besides, semantic communication can adapt to the diverse service demands of “Intellicise” wireless networks. Different applications may require different types and amounts of semantic information, and semantic communication can provide flexible compression and transmission of this information to meet these varying demands. Therefore, semantic representation schemes and transmission mechanisms that process source data, filter out redundant data, and retain only intent-related information, provide an evolutionary path for the development of “Intellicise” networks and cater to the higher demands of the future 6G network.

Semantic Communications for “Intellicise” networking The ubiquitous interconnection of the 6G network makes it highly complex. Traditional mobile communication networks that rely on precise mathematical models and ideal prior assumptions face challenges when applied to future 6G scenarios. To address

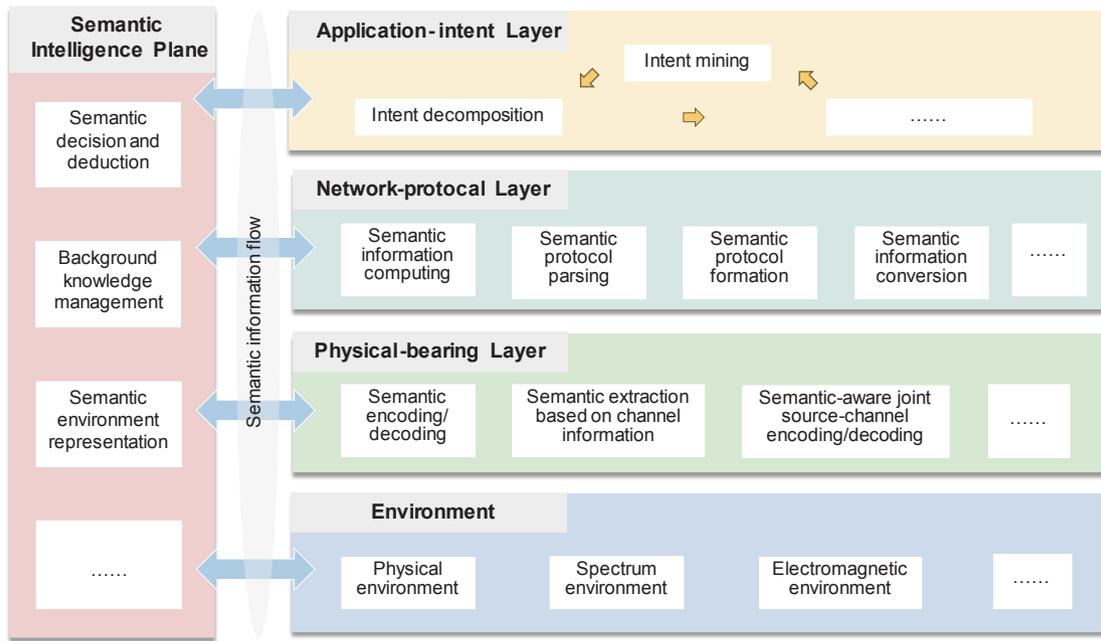


Figure 4 Semantic communication-empowered 6G architecture [13].

this, there is an urgent need to shift from the traditional “precise mathematical modeling” network construction concept to a new “Intellicise modeling” concept that constructs a 6G ubiquitous wireless network with native intelligent capabilities such as self-learning, self-adaptation, and self-evolution, to achieve an intelligence-endogenous and primitive-concise 6G networks. Moreover, by conveying semantic information that connects the physical and digital spaces, semantic communication can enhance the coordination between various nodes, leading to more efficient and intelligent network operations. As shown in Figure 4, by introducing the semantic intelligence (SI) plane, Zhang *et al.* [13] proposed an intelligent and efficient semantic communication (IE-SC) architecture, which is composed of an SI plane and three layers: the semantic-empowered physical-bearing layer (S-PB), network protocol layer (S-NP), and application-intent layer (S-AI). In detail, the S-AI layer extracts, understands, and decomposes systematic primitives into sub-intents, and sends sub-intents to the SI plane. Then, the SI plane synthesizes the sub-intents and develops networking policies based on them. Moreover, the SI plane continuously learns background knowledge from the environment and shares knowledge with three layers, driving the network towards a state of self-optimization and self-evolution.

Semantic communications for goal-oriented applications

6G is promised to extend the vertical applications already supported by 5G, aiming to provide more diverse and personalized services as well as support a variety of intelligent applications with capabilities such as perception, learning, decision-making, and evolution. As shown in Figure 5, semantic communication has a wide range of potential application scenarios that can empower various emerging industries, including intelligent healthcare, intelligent transportation, and intelligent factory [107].

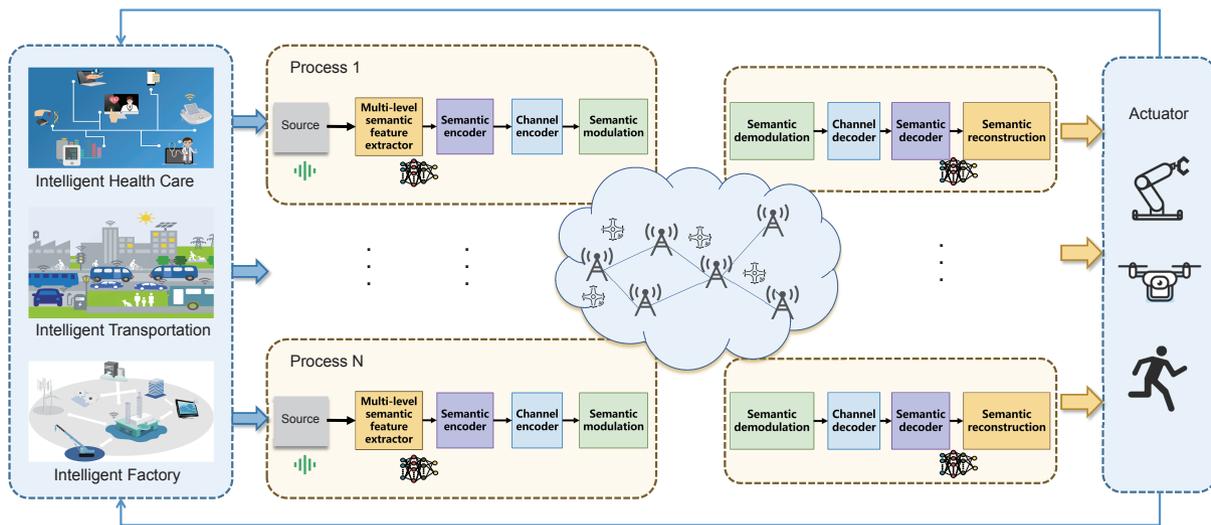


Figure 5 Semantic communications for goal-oriented applications modified from Ref. [107].

(1) Intelligent healthcare

The healthcare industry is being transformed by disruptive technologies, such as AI, IoT, 5G, cloud/edge computing, etc. In particular, the upcoming 6G networks are expected to support intelligent healthcare applications that require ultra-low latency, high throughput, ultra-high reliability, and energy efficiency, such as tactile/haptic Internet, intelligent Internet of medical things (IIoT), and hospital-to-home services. For example, remote surgeries require high-quality video and audio captured on the user side to be transmitted to the remote surgeon or team with minimal latency to ensure immediate feedback and guidance to the surgical team. The growing number of connected devices and applications in intelligent healthcare generates health-related data in various sizes and formats. Processing such a massive amount of health-related data has the potential to lead to improved human well-being, early diagnosis of diseases, and even help combat future pandemics.

Semantic communication has the potential to enhance communication efficiency and satisfy the stringent transmission demands of intelligent healthcare by extracting semantic features from data, and preserving its meaning while compressing and eliminating irrelevant data. The authors in Ref. [108] developed a deep semantic communication (DeepSC)-based network service framework that incorporated semantic cognition into the healthcare cyber-physical system. This framework optimized the transceiver in communication by incorporating physical layer blocks into the conventional communication system. The authors in Ref. [109] proposed a domain knowledge-driven semantic communication system (DKSC) that uses a dual-path framework that performs semantic extraction and reconstruction at both information and concept levels. Taking image transmission in the medical domain as an example, they designed and optimized an end-to-end wireless communication system based on DNNs for medical image transmission. Furthermore, semantic feature extraction from raw data can not only reduce the amount of transmitted information but also enhance data security. In the event of a semantic information leak, it can make it more difficult for attackers to obtain the original data from the extracted information.

To summarize, the implementation of semantic communication in healthcare has the potential to meet the

industry's diverse needs and accelerate its digital development. Although semantic communication and its applications are still in the early stages of development, the anticipated impact of semantic communications on existing wireless network infrastructures in healthcare is expected to revolutionize the industry.

(2) Intelligent transportation

Intelligent transportation systems (ITSs) are becoming increasingly vital in our daily lives. Vehicles can be considered as intelligent agents equipped with sensing, computing, caching, and wireless communication capabilities, enabling autonomous driving and cooperative vehicle networks without human involvement. The integration of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) networking plays a critical role in future ITS to support massive connectivity and satisfy diverse requirements. However, the rapid development of autonomous vehicles poses challenges in fulfilling ITS demands. On the one hand, efficient processing of a significant amount of data collected by vehicles is crucial, but the transmission of the original data requires a large bandwidth and results in significant communication overhead, which is not feasible for the current Internet of vehicles (IoV) due to limited spectrum resources. On the other hand, the rapid growth of different types of autonomous vehicles makes it challenging to accommodate the heterogeneous requirements of ITS.

The advancement in intelligence and edge computing enables vehicles to perform terminal data processes, which supports semantic communication by extracting the core semantic data and transmitting only the meaning. In autonomous driving scenarios, vehicles can efficiently interact with each other or exchange information with roadside units/edge nodes through semantic communications. At present, several works have proposed various semantic communication frameworks to tackle the challenges associated with ITS. The authors in Ref. [110] proposed a semantic communication framework that utilizes a convolutional auto-encoder (CAE) to extract semantic concepts from traffic signals, which are then transmitted to the macro base station (MBS). The MBS employs a proximal policy optimization (PPO) algorithm to make decisions for connected and autonomous vehicles (CAVs). The authors in Ref. [111] proposed a cooperative semantic-aware architecture to reduce data traffic by conveying critical semantics from collaborated users to the server. They also provided a case study on vehicle image retrieval tasks in ITSs. The authors in Ref. [112] proposed SemCom-enabled service supplying solution (S4), addressing two fundamental problems, namely knowledge base construction (KBC) and vehicle service provisioning (VSP). This solution provides superior performance in terms of queuing latency, semantic data packet throughput, user knowledge matching degree, and user knowledge preference satisfaction compared with the two benchmarks. The authors in Ref. [113] proposed a dynamic resource allocation method for D2D vehicular networks based on semantic communication to achieve efficient transmission while meeting user SINR requirements, improving transmission efficiency and resource utilization.

On the other hand, UAVs have become an essential component in future ITS thanks to their flexibility in deployment and fast construction. Semantic segmentation has been widely utilized in UAV image analysis to locate people and perform other specific tasks. The authors in Ref. [114] proposed a task-oriented semantic communication framework for image classification in UAVs and used a DRL-based algorithm to explore the semantic blocks with the most significant contribution to the back-end classifier under different channel conditions. The authors in Ref. [115] proposed a multi-agent deep reinforcement learning (MADRL) framework and established a graph attention exchange network (GAXNet) among UAVs to achieve ultra-reliable and low-latency communication (URLLC). The authors in Ref. [76] proposed an energy-efficient

task-oriented semantic communication framework that uses a triple-based scene graph and a personalized semantic encoder based on user interests. Additionally, they investigated the impact of wireless fading channels on semantic transmission and proposed a multi-user resource allocation scheme to optimize the system performance.

To summarize, semantic communication can be utilized to extract crucial semantic information from the raw sensor data, which includes information related to vehicle kinematics, road conditions, and traffic signals. For example, in task-specific applications, only the semantic information relevant to the task execution is extracted at the transmitter and used for decision-making at the receiver. This approach can improve communication efficiency, reduce transmission overhead, and enhance data security. However, the research and application of semantic communication in the ITS are still in their infancy, with many open issues that need to be addressed. First, ensuring URLLC is critical, especially when aiming for adequate semantic fidelity in large-scale V2V communications. This requires addressing the latency requirement related to varying processing efficiencies of different semantic representations and coding solutions based on DL models. Furthermore, the reliability requirement related to the strict SKB matching between the transmitter and receiver should also be well addressed. In addition, exploring the use of semantic communication for more advanced ITS applications beyond basic information exchange, such as intelligent route planning and decision-making, is a promising research direction. Overall, further research and development in the area of semantic communication for ITS are needed to fully realize its potential in improving safety, efficiency, and sustainability in transportation.

(3) Intelligent factory

ICT technology has been instrumental in driving the deep integration of the digital economy and the real economy, facilitating the process of industrial digitization. In the industrial Internet, data are often sparse and heterogeneous, making it difficult to extract relevant information. In this context, semantic communication offers a practical solution that is content-aware, task-oriented, and selective in transmitting only essential information to users and applications. By reducing data redundancy and enabling more efficient signal generation and transmission, semantic communication can support the processing and transmission of large volumes of real-time data in an efficient manner. In terms of signal generation, semantic sampling allows for adjusting the sampling rate based on available semantic information, reducing the communication burden by only sampling useful data. In terms of signal transmission, task-oriented semantic communication enables tailored data compression, resulting in higher compression ratios. This approach is particularly useful in scenarios where sensors generate large amounts of data, but only a small portion is relevant. Academics have conducted initial research on semantic communication architecture for the industrial Internet. The authors in Ref. [116] proposed a generic layered architecture for semantic-aware cyber-physical systems (SCPSs) by developing a communication layer and semantic layer on top of the existing CPS system architectures. The proposed SCPS architecture enables semantic machine-to-machine (M2M) communications in a network of manufacturing devices. The authors in Ref. [45] proposed a lightweight distributed semantic communication system designed for low-computing-capability devices by leveraging a model compression algorithm.

To summarize, semantic communication is a content-aware and task-oriented approach that focuses on transmitting only the most essential information to the users and applications, making it an effective solution for supporting the large volume of real-time data processed and transmitted in the industrial Internet. Recent research has demonstrated the promising potential of semantic communication in the industrial internet

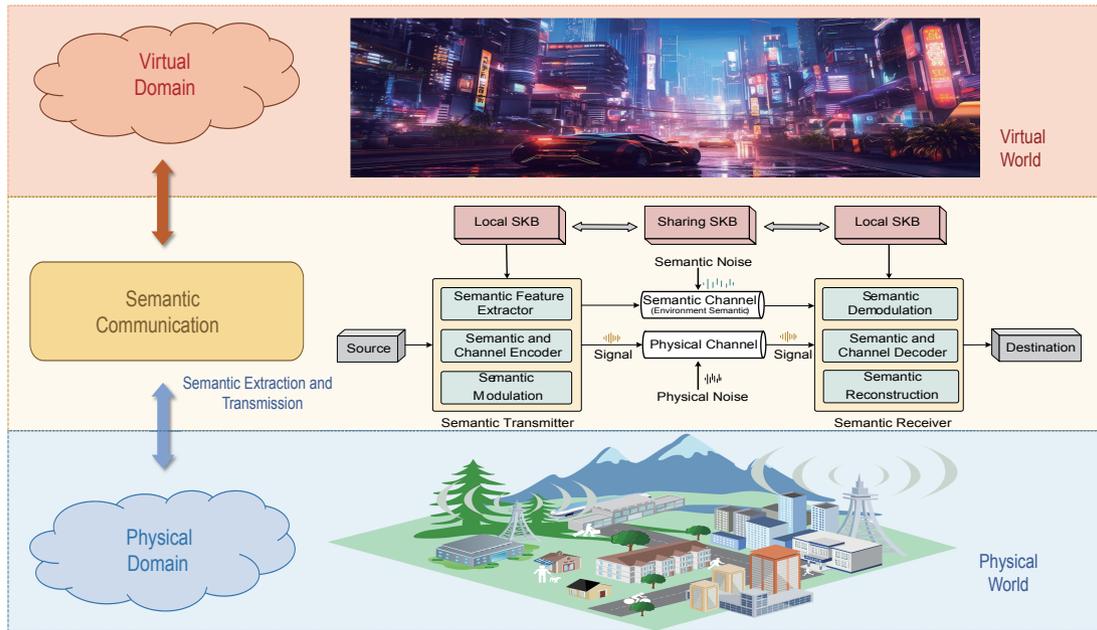


Figure 6 Semantic communications for Metaverse (Insert image credit: Ralph, CC BY 4.0.).

landscape, but further research and development are needed to realize its potential fully.

Semantic communications for Metaverse

The concept of the Metaverse, as first described in Neil Stevenson’s science fiction novel *Snow Crash* in 1992, refers to a world where the virtual interacts with reality and creates value through various social activities. With recent advances in technologies, including 5G/6G, digital twins, XR, and AI, the Metaverse has rapidly gained popularity. The Metaverse is a digital space that creates immersive virtual worlds that interact with the physical world, requiring high-quality 3D visual scenes that resemble the physical world closely, as well as high synchronization and low latency to provide users with an optimal experience. As shown in Figure 6, semantic communication is essential in this context as it enables end devices to transmit only relevant information to the Metaverse server for operation, reducing bandwidth and computing latency [117]. Additionally, the Metaverse server is capable of extracting semantic information based on the user’s preferences while disregarding irrelevant details, thereby reducing downlink pressure. The implementation of semantic communications in the Metaverse has transformative potential in enabling efficient communication, facilitating human-machine interaction, and promoting visual scene reconstruction.

Semantic communication reduces metaverse bandwidth consumption One unique aspect of the Metaverse is the real-time synchronization between the physical and virtual worlds, which requires a large amount of data to enhance the fidelity of the virtual world [118]. By prioritizing understanding before transmission, semantic communication reduces the transmission of redundant information and increases information transmission efficiency, which is particularly important for XR/AR services in the Metaverse. Semantic communication allows for the transmission of simple semantic information that describes specific scenes, which can be rendered at the receiver side, rather than the traditional method of rendering light that

involves sending a large amount of raw information. Moreover, joint semantic-channel coding and SKB-based semantic information extraction have been proposed to further enhance the efficiency of semantic transmission in Metaverse [105]. In addition, semantic communication systems can extract semantic information about the user's desired image, transmit only the region of interest [119], and enhance privacy protection by not transmitting raw data.

Semantic communication facilitates human-computer interaction The human-computer interaction (HCI) is a critical aspect of the Metaverse, requiring real-time feedback of user actions and interactions with other users to create a seamless virtual world. With the advancement of technologies such as motion capture, XR, and digital twins, user actions can be fed back to the computer and generated on the platform through a network connection. For example, XR technologies can enable users to experience multisensory feedback, such as haptic feedback, which enhances the immersive experience. Digital twins can enable the creation of highly accurate virtual representations of physical objects, environments, and processes, enabling users to interact with and manipulate them in a natural way. In addition, the use of brain-computer interface (BCI) technology in the Metaverse allows users to directly control their avatars or external devices using conscious brain activities, eliminating the need for physical devices and enhancing the efficiency and convenience of HCI. To support real-time HCI or BCI among massive users in the Metaverse, synchronized and low-delay data collection, transmission, and processing are essential. Semantic communication aims to process and transmit the meaning of multi-sensory data, such as images, audio, and text, making it more efficient for machines to understand and process. For instance, semantic communication involves extracting semantics from the data collected and tracked by end devices, such as head movements, arm swings, gestures, and speech. The relevant information is then transmitted from the end device to the Metaverse server, thereby saving bandwidth and reducing computing latency. Additionally, semantic communication networks can support the distributed interaction requirements of terminals and clouds in the Metaverse [105]. Furthermore, with the advancement of brain science, it has become evident that semantics are pervasive and can be characterized [120], making the implementation of semantic communications in Metaverse feasible.

Semantic communication promotes metaverse visual scene reconstruction The Metaverse requires high-quality 3D virtual visual scenes that resemble the physical world closely, as well as high synchronization and low latency. The achievement of this goal relies heavily on 3D reconstruction, which can be classified into three categories: object-level, human-level, and scene-level. This component involves various techniques, such as 3D semantic segmentation, 3D object detection and recognition, 3D instance segmentation, 3D pose estimation, and 3D reconstruction. Among them, the performance of 3D reconstruction greatly depends on the semantic representation of output data, which significantly impacts both computational efficiency and reconstructed quality. Particularly, semantic information can also be used to support simultaneous localization and mapping to acquire 3D structures of unknown environments and perceive their motion [121]. Unlike traditional purely model-based methods, semantic communication with the help of SKBs can learn prior knowledge from large amounts of data, allowing for more effective processing of 3D reconstruction tasks.

To summarize, semantic communication plays a vital role in improving transmission efficiency by enabling users to transmit only the relevant content to the receiver, thereby enhancing the HCI and BCI experience in the Metaverse. Additionally, semantic-aware networks have become increasingly important in supporting the distributed interaction requirements of terminals and clouds in the Metaverse, enabling real-time synchro-

nization. Further research and development on semantic communications are necessary to maximize the potential in enhancing the user experience and promoting the development of the Metaverse.

Summary and lessons learned

Lessons learned for “Intellicise” wireless networks Semantic communication is a key solution for enabling “Intellicise” wireless networks. By conveying the meaning of information, semantic communication can effectively reduce the data transmission volume of traditional communication methods and reduce communication overhead while maintaining the accuracy of information transmission. The concept of Sebs highlights the importance of embedding contextual meaning within transmitted information, enhancing the relevance and intelligibility of received data. Furthermore, ultra-concise network optimization strategies have the potential for self-optimizing and self-evolving networks, achieving resource maximization and autonomous optimization. While previous research has primarily focused on alleviating existing communication network burdens from a technical standpoint, this paradigm shift still leaves much room for further research.

Lessons learned for goal-oriented applications Semantic communications offer the potential for significant performance improvements in intelligent networks geared toward goal-oriented applications, including intelligent healthcare, intelligent transportation, and intelligent factories. As we discussed above, semantic communication has great potential in data processing and transmission, which, in turn, facilitates subsequent tasks. In the context of networked intelligent systems, where the expeditious and effective execution of tasks is paramount, semantics-empowered networked intelligent systems could enhance task execution efficiency while minimizing costs. However, despite the research progress, fundamental performance enhancements are still in their infancy and the semantic communication for goal-oriented application is still worth further exploration.

Lessons learned for Metaverse Metaverse over wireless networks pose unprecedented challenges in terms of its multi-modal data transmissions with stringent latency and reliability requirements. The performance of the Metaverse heavily relies on the collection and processing of data related to human movements and environmental changes. Effective tracking and accurate prediction are key to reducing transmission and computation latency and ensuring a smooth user experience. Semantic communications offer a feasible way of reducing end-to-end latency by ignoring irrelevant details and transmitting the information after understanding. In addition, semantic communication refers to the exchange of information between different systems and platforms, which enables interoperability and mutual understanding, resulting in a more seamless and user-friendly experience.

OPEN PROBLEMS AND KEY CHALLENGES

Despite the promising prospects of semantic communication, several significant research challenges need to be addressed. In this section, we present some of these open problems and research challenges, along with a broader discussion of their implications.

Semantic theory

Nowadays, there is still a lack of mathematical foundations for the theoretical analysis of semantic information and semantic transmission. There is a need to develop methods for measuring information and achieving the bound of the transmission rate. Current research has yet to make a breakthrough in these areas, and there is a need to further explore mathematical models for semantic communication to provide a theoretical basis for its practical applications.

Semantic information compression and transmission Efficient semantic compression and transmission are essential for the development of semantic communications. Traditional data compression methods rely on metrics such as entropy, whereas semantic information is subjective and fuzzy, making it challenging to measure the amount of semantic information contained. Additionally, semantic information can be carried through multiple modalities, which makes it challenging to measure the amount of information across different modal sources. Furthermore, semantic information is task-oriented, meaning that different tasks may require different types and amounts of semantic information. It is necessary to develop a metric for task-oriented semantic information. Therefore, to achieve efficient and effective semantic information compression and transmission, the theoretical framework for quantitatively analyzing and measuring semantic information must be refined.

Semantic knowledge base The implementation of semantic communication relies on a multi-level semantic knowledge representation of the source, task, and channel in typical scenarios. However, current research mainly focuses on the construction of SKB based on source characteristics under ideal interaction conditions, with insufficient attention given to the collaborative update of SKB in semantic transmission. Moreover, there is no theoretical analysis of SKB performance, including the semantic representation capabilities of SKB, and the theoretical bound of performance gain that SKB can bring for semantic communication. Therefore, there is an urgent need to develop theories and methods for constructing multi-level SKBs to provide unified methodological guidance in typical scenarios. Additionally, developing collaborative update methods for SKBs that can adapt to changes in the source, channel, and tasks as well as maintain the accuracy of the transmitted information is a critical area for further research.

Performance analysis Despite some preliminary explorations, there still lacks a unified evaluation method for semantic communications. To promote the generalization of semantic communication algorithms under various scenarios, it is essential to establish a unified framework for semantic evaluation, serving as the basic scale for diverse scenarios of semantic communications. However, the emergence of various applications introduces further challenges where more metrics are required, such as for multi-modal and cross-modal data interaction. It is challenging to unify semantic metrics with different evaluation models. Furthermore, in line with the vision of 6G networks, the evaluation framework should jointly consider communication efficiency, end-to-end latency, energy efficiency, and other factors. Thus, there is a need for a generalized evaluation framework that comprehensively considers the efficiency and complexity of semantic communications in various scenarios for future wireless networks.

By addressing these research gaps, we can further enhance the effectiveness of semantic communication and unlock its full potential in a wide range of applications. Furthermore, the exploration and innovation of new methods and theories for semantic communication will enable the development of more advanced and intelligent communication systems, which can further promote the integration of AI and communication

technology.

Semantic communication architectures and techniques

Semantic communication redefines the fundamental aspects of communication systems. Despite the advances made in the semantic communication systems, there are still several challenges and open issues that need to be addressed.

SKB modelling and updating As the foundation of semantic communications, the SKB provides multi-level semantic knowledge descriptions for source information, transmission environments, and task requirements. It is a database that expresses the meaning, logic, and relationships between various sources, enabling the representation of more complex semantics and advanced querying and reasoning. Building an SKB requires numerous publicly available large-scale datasets, and the process of updating and maintaining it is time-consuming and resource-intensive due to the diverse sources, varying tasks, and intelligent environment. To simplify the complexity of semantics, SKB should minimize high dimensional semantics, fuse similar semantic knowledge, and reduce redundant semantic knowledge, by learning more characteristics based on the feedback from the intelligent transmission environment and the receiver tasks. SKB should also correct corrupted semantic knowledge based on a self-detection and self-update mechanism to ensure accuracy and consistency. In addition, the SKB should update with new semantics in the source information, transmission environments, and task requirements, enriching the diversity of semantics and improving the representation capability of the SKB. It is worth noting that the local SKBs of different individual users can differ significantly, and addressing the collaborative updates between the public SKB stored in the cloud servers and the local SKBs at the end devices is crucial to accelerate the updating of the SKB. Future research should focus on developing more efficient methods for constructing and updating the SKB, as well as enhancing its scalability and adaptability, to improve the performance and experience of semantic communication systems.

End-to-end semantic transmission The use of semantics allows for concise and efficient representations, resulting in highly effective end-to-end semantic transmission. However, developing a general semantic transmission framework for different sources and varied channel conditions is not available yet. Achieving flexible compression of semantic information for different semantic tasks and channel capabilities is also challenging. During semantic transmission, semantic tasks have different requirements and need to be compressed in a targeted manner. Heterogeneous sources also need flexible compression with their different DL-based codec schemes. The dynamic and evolving nature of the SKB can also affect the robust transmission of semantics. Furthermore, differences in background knowledge can lead to semantic distortions. Designing a robust semantic communication system incorporating ML techniques is a significant challenge. Finally, adapting to varying wireless channels can be achieved by transmitting essential semantic features in subchannels with high SNRs, taking advantage of the unique features of wireless channels in semantic communication systems.

Semantic resource orchestration Efficient scheduling and resource allocation policies are crucial for semantic communication and networking. While semantic communication reduces transmission overhead, it requires more computing resources dedicated to complex DNNs. Thus, exploring the fundamental trade-off between communication efficiency and computing resources is vital to ensure the sustainability of semantic

communication systems. Different modal information sources have differentiated semantic features, which have different requirements for computing resources. It is also important to address the level of compression and reasoning that can be achieved with limited computing resources. Additionally, considering devices with different communication and computing capabilities, investigating how a semantic communication network can be built and the performance gains that can be achieved for semantic communication in such heterogeneous systems are essential. Furthermore, the semantic-aware resource orchestration should explore the trade-off relationship between semantic accuracy, semantic processing latency, and energy consumption to achieve the maximum performance gain.

Co-design of semantic sensing, communication, computing, and control The life cycle of information is closely related to the processes of sensing (sampling), communication, computation, caching, and control. By integrating the modules of sensing, communication, computing, caching, and control, the system can be optimized to perform information processing tasks more efficiently. In particular, multi-modal sensors continuously perceive high-resolution raw data, computation produces precise semantic extractions and effective intelligent decisions, communication serves as a priority-aware information pipeline for disseminating information, and actuators execute control commands accurately to transform the information into practical usage. Co-designing these components can result in a seamless and robust system that is capable of handling diverse and complex information processing tasks, providing a promising foundation for developing intelligent systems and leading to practical applications in healthcare, transportation, manufacturing, and other fields. However, challenges remain in integrating the different modules, particularly when dealing with large amounts of data and information. The scalability of the system can also be a challenge, and integrating different modules can lead to security and privacy concerns. As such, further research is necessary to develop effective co-design frameworks that are efficient, accurate, and secure while being adaptable to meet different requirements across various fields.

The advent of semantic communication has ushered in a new era in the design of communication systems. Besides the above open issues, there are still some aspects that need to be addressed in semantic communication systems. For example, the development of standard protocols for semantic communication is also an open issue that needs to be addressed to ensure interoperability and compatibility between different systems. Overall, addressing these open issues is critical for the development of efficient and reliable semantic communication systems and unlocking their full potential for improving communication in various domains.

CONCLUSIONS AND PERSPECTIVES

Semantic communication is a promising paradigm for the development of communication systems in the 6G era. With traditional communication technologies facing challenges in terms of system complexity and sustainability, semantic communication offers a promising new paradigm. In this paper, we provide a systematic review of semantic communication. First, we discuss the SIT and the overall development of SeB-based semantic communication architecture. Then, we explore the goal-oriented applications of semantic communication in vertical industries and highlight the irreplaceable nature of semantic communication for Metaverse. Finally, we outline and identify open issues and key challenges with a view to enhancing research in semantic communication. It is clear that semantic communication offers new possibilities for the future of

mobile communication and could play a significant role in shaping the next generation of communication systems.

Funding

This work was supported by the National Natural Science Foundation of China (62293480, 62293481), and the National Natural Science Foundation for Young Scientists of China (62001050).

Author contributions

P.Z. and Y.L. designed the research and wrote the initial manuscript. Y.S. and J.Z. provided the source supports and supervision. All authors have contributed to the discussion of the content and have revised and edited the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

- 1 Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* 1948; **27**: 379–423.
- 2 Wang Y, Gao Z, Zheng D, *et al.* Transformer-empowered 6G intelligent networks: From massive MIMO processing to semantic communication. *IEEE Wireless Commun* 2022; **30**: 127–135.
- 3 Uysal E, Kaya O, Ephremides A, *et al.* Semantic communications in networked systems: A data significance perspective. *IEEE Network* 2022; **36**: 233–240.
- 4 Shannon CE, Weaver W. *The Mathematical Theory of Communication*. Champaign: University of Illinois Press, 1949.
- 5 Zhong Y. A theory of semantic information. *China Commun* 2017; **14**: 1–17.
- 6 Morris CW. Foundations of the theory of signs. In: *International Encyclopedia of Unified Science*. Chicago: Chicago University Press, 1938. 1-59.
- 7 Carnap R, Bar-Hillel Y. An outline of a theory of semantic information. RLE Technical Reports 247, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, 1952.
- 8 Bar-hillel Y, Carnap R. Semantic information. *Br J Philosophy Sci* 1953; **4**: 147–157.
- 9 Barwise J, Perry J. Situations and attitudes. *J Philosophy* 1981; **78**: 668–691.
- 10 D’Alfonso S. On quantifying semantic information. *Information* 2011; **2**: 61–101.
- 11 De Luca A, Termini S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Readings in Fuzzy Sets for Intelligent Systems*. Morgan Kaufmann, 1993. 197–202.
- 12 Floridi L. Outline of a theory of strongly semantic information. *Minds Machines* 2004; **14**: 197–221.
- 13 Zhang P, Xu W, Gao H, *et al.* Toward wisdom-evolutionary and primitive-concise 6G: A new paradigm of semantic communication networks. *Engineering* 2022; **8**: 60–73.
- 14 Kim Y. Convolutional neural networks for sentence classification. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, 2014. 1746-1751.
- 15 Bengio Y, Ducharme R, Vincent P. A neural probabilistic language model. In: *Proceedings of the 13th International Conference on Neural Information Processing Systems*. Denver, 2000. 932–938.
- 16 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, 2017. 5998–6008 .
- 17 Radford A, Narasimhan K, Salimans T, *et al.* Improving language understanding by generative pre-training. OpenAI blog, 2018. <https://openai.com/index/language-unsupervised/>.
- 18 Devlin J, Chang MW, Lee K, *et al.* Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv: [1810.04805](https://arxiv.org/abs/1810.04805).

- 19 Russakovsky O, Deng J, Su H, *et al.* ImageNet large scale visual recognition challenge. *Int J Comput Vis* 2015; **115**: 211–252.
- 20 Lee H, Kwon H. Going deeper with contextual CNN for hyperspectral image classification. *IEEE Trans Image Process* 2017; **26**: 4843–4855.
- 21 Yang W, Du H, Liew ZQ, *et al.* Semantic communications for future internet: fundamentals, applications, and challenges. *IEEE Commun Surv Tutor* 2023; **25**: 213–250.
- 22 Lan Q, Wen D, Zhang Z, *et al.* What is semantic communication? A view on conveying meaning in the era of machine intelligence. *J Commun Inf Netw* 2021; **6**: 336–371.
- 23 Qin Z, Tao X, Lu J, *et al.* Semantic communications: Principles and challenges. arXiv: [2201.01389](https://arxiv.org/abs/2201.01389).
- 24 Iyer S, Khanai R, Torse D, *et al.* A survey on semantic communications for intelligent wireless networks. *Wireless Pers Commun* 2023; **129**: 569–611.
- 25 Lu Z, Li R, Lu K, *et al.* Semantics-empowered communication: A tutorial-cum-survey. *IEEE Commun Surv Tutor* 2024; **129**: 41–79.
- 26 Liu Y, Wang X, Ning Z, *et al.* A survey on semantic communications: Technologies, solutions, applications and challenges. *Digital Commun Networks* 2023; doi:[10.1016/j.dcan.2023.05.010](https://doi.org/10.1016/j.dcan.2023.05.010).
- 27 Getu TM, Kaddoum G, Bennis M. Tutorial-cum-survey on semantic and goal-oriented communication: Research landscape, challenges, and future directions. arXiv: [2308.01913](https://arxiv.org/abs/2308.01913).
- 28 Li K, Lau BPL, Yuan X, *et al.* Toward ubiquitous semantic metaverse: challenges, approaches, and opportunities. *IEEE Internet Things J* 2023; **10**: 21855–21872.
- 29 Shi G, Xiao Y, Li Y, *et al.* From semantic communication to semantic-aware networking: model, architecture, and open problems. *IEEE Commun Mag* 2021; **59**: 44–50.
- 30 Choi J, Loke SW, Park J. A unified view on semantic information and communication: A probabilistic logic approach. In: Proceedings of the 2022 IEEE International Conference on Communications Workshops (ICC Workshops). Seoul: IEEE, 2022. 705–710.
- 31 Liu X, Jia W, Liu W, *et al.* AFSSE: An interpretable classifier with axiomatic fuzzy set and semantic entropy. *IEEE Trans Fuzzy Syst* 2019; **28**: 2825–2840.
- 32 Kountouris M, Pappas N. Semantics-empowered communication for networked intelligent systems. *IEEE Commun Mag* 2021; **59**: 96–102.
- 33 Rényi A. On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Volume 1: Contributions to the Theory of Statistics. Oakland: University of California Press, 1961. 4: 547–562.
- 34 Liu J, Zhang W, Poor HV. A rate-distortion framework for characterizing semantic information. In: Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT). Melbourne: IEEE, 2021. 2894–2899.
- 35 Sun Y, Chen H, Xu X, *et al.* Semantic knowledge base-enabled zero-shot multi-level feature transmission optimization. *IEEE Trans Wireless Commun* 2023; **23**: 4904–4917.
- 36 Sun Y, Zhang J, Yu L, *et al.* How to define the propagation environment semantics and its application in scatterer-based beam prediction. *IEEE Wireless Commun Lett* 2023; **12**: 649–653.
- 37 Zheng Y, Wang F, Xu W, *et al.* Semantic communications with explicit semantic base for image transmission. In: Proceedings of the IEEE Global Communications Conference. Kuala Lumpur, 2023. 4497–4502.
- 38 Griffiths TL, Steyvers M. A probabilistic approach to semantic representation. In: Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society. Fairfax, 2002. 381–386.
- 39 Hinton GE, Roweis S. Stochastic neighbor embedding. In: Proceedings of the 15th International Conference on Neural Information Processing Systems. Vancouver, 2002. 833–864.
- 40 Mikolov T, Sutskever I, Chen K, *et al.* Distributed representations of words and phrases and their compositionality. In: Proceedings of the 26th International Conference on Neural Information Processing Systems. Vol. 2. Lake Tahoe, 2013. 3111–3119.

- 41 Pennington J, Socher R, Manning CD. Glove: Global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, 2014. 1532–1543.
- 42 Bojanowski P, Grave E, Joulin A, *et al.* Enriching word vectors with subword information. *TACL* 2017; **5**: 135–146.
- 43 Peters ME, Neumann M, Iyyer M, *et al.* Deep contextualized word representations. arXiv: [1802.05365](https://arxiv.org/abs/1802.05365).
- 44 Farsad N, Rao M, Goldsmith A. Deep learning for joint source-channel coding of text. In: Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary: IEEE, 2018. 2326–2330.
- 45 Xie H, Qin Z. A lite distributed semantic communication system for internet of things. *IEEE J Sel Areas Commun* 2020; **39**: 142–153.
- 46 Zhou Q, Li R, Zhao Z, *et al.* Semantic communication with adaptive universal transformer. *IEEE Wireless Commun Lett* 2021; **11**: 453–457.
- 47 Liu Y, Jiang S, Zhang Y, *et al.* Extended context-based semantic communication system for text transmission. *Digital Commun Networks* 2022; doi:[10.1016/j.dcan.2022.09.023](https://doi.org/10.1016/j.dcan.2022.09.023).
- 48 Weng Z, Qin Z. Semantic communication systems for speech transmission. *IEEE J Sel Areas Commun* 2021; **39**: 2434–2444.
- 49 Tong H, Yang Z, Wang S, *et al.* Federated learning for audio semantic communication. *Front Comms Net* 2021; **2**: 734402.
- 50 Han T, Yang Q, Shi Z, *et al.* Semantic-preserved communication system for highly efficient speech transmission. *IEEE J Sel Areas Commun* 2022; **41**: 245–259.
- 51 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Vol. 1. Lake Tahoe, 2012.
- 52 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: [1409.1556](https://arxiv.org/abs/1409.1556).
- 53 He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016. 770–778.
- 54 Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: Proceedings of the International Conference on Machine Learning. Long Beach, 2019. 6105–6114.
- 55 Kingma DP, Welling M. Auto-encoding variational bayes. arXiv: [1312.6114](https://arxiv.org/abs/1312.6114).
- 56 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial networks. *Commun ACM* 2020; **63**: 139–144.
- 57 Boursoulatze E, Burth Kurka D, Gunduz D. Deep joint source-channel coding for wireless image transmission. *IEEE Trans Cogn Commun Netw* 2019; **5**: 567–579.
- 58 Yang M, Kim HS. Deep joint source-channel coding for wireless image transmission with adaptive rate control. In: Proceedings of the ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Singapore: IEEE, 2022. 5193–5197.
- 59 Zhang Z, Yang Q, He S, *et al.* Wireless transmission of images with the assistance of multi-level semantic information. In: Proceedings of the 2022 International Symposium on Wireless Communication Systems (ISWCS). Hangzhou: IEEE, 2022. 1–6.
- 60 Dai J, Wang S, Tan K, *et al.* Nonlinear transform source-channel coding for semantic communications. *IEEE J Sel Areas Commun* 2022; **40**: 2300–2316.
- 61 Xie H, Qin Z, Li GY. Task-oriented multi-user semantic communications for VQA. *IEEE Wireless Commun Lett* 2021; **11**: 553–557.
- 62 Xie H, Qin Z, Tao X, *et al.* Task-oriented multi-user semantic communications. *IEEE J Sel Areas Commun* 2022; **40**: 2584–2597.
- 63 Zhang G, Hu Q, Qin Z, *et al.* A unified multi-task semantic communication system for multimodal data. *IEEE Trans Commun* 2024; doi: [10.1109/TCOMM.2024.3364990](https://doi.org/10.1109/TCOMM.2024.3364990).
- 64 Wang S, Yang Q, Shi Z, *et al.* Cooperative task-oriented communication for multi-modal data with transmission

- control. In: Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops). Rome, 2023. 1635–1640.
- 65 Bo Y, Duan Y, Shao S, *et al.* Learning based joint coding-modulation for digital semantic communication systems. In: Proceedings of the 2022 14th International Conference on Wireless Communications and Signal Processing (WCSP). Nanjing, 2022.
- 66 Jiang P, Wen CK, Jin S, *et al.* Deep source-channel coding for sentence semantic transmission with HARQ. *IEEE Trans Commun* 2022; **70**: 5225–5240.
- 67 Zhou Q, Li R, Zhao Z, *et al.* Adaptive bit rate control in semantic communication with incremental knowledge-based HARQ. *IEEE Open J Commun Soc* 2022; **3**: 1076–1089.
- 68 Guo J, Wen CK, Jin S. Eliminating CSI feedback overhead via deep learning-based data hiding. *IEEE J Sel Areas Commun* 2022; **40**: 2267–2281.
- 69 Xu J, Ai B, Wang N, *et al.* Deep joint source-channel coding for CSI feedback: an end-to-end approach. *IEEE J Sel Areas Commun* 2022; **41**: 260–273.
- 70 Zhang G, Hu Q, Cai Y, *et al.* Adaptive CSI feedback for deep learning-enabled image transmission. In: Proceedings of the IEEE International Conference on Communications. Rome, 2023. 4163–4168.
- 71 Yan L, Qin Z, Zhang R, *et al.* Resource allocation for text semantic communications. *IEEE Wireless Commun Lett* 2022; **11**: 1394–1398.
- 72 Yan L, Qin Z, Zhang R, *et al.* QoE-aware resource allocation for semantic communication networks. In: Proceedings of the GLOBECOM 2022-2022 IEEE Global Communications Conference. Rio de Janeiro: IEEE, 2022. 3272–3277.
- 73 Wang Y, Chen M, Saad W, *et al.* Performance optimization for semantic communications: an attention-based learning approach. In: Proceedings of the 2021 IEEE Global Communications Conference (GLOBECOM). Madrid: IEEE, 2021. 1–6.
- 74 Liu C, Guo C, Yang Y, *et al.* Adaptable semantic compression and resource allocation for task-oriented communications. *IEEE Trans Cogn Commun Netw* 2024; **10**: 769–782.
- 75 Liu C, Guo C, Yang Y, *et al.* Bandwidth and power allocation for task-oriented semantic communication. arXiv: [2201.10795](https://arxiv.org/abs/2201.10795).
- 76 Kang J, Du H, Li Z, *et al.* Personalized saliency in task-oriented semantic communications: Image transmission and performance analysis. *IEEE J Sel Areas Commun* 2022; **41**: 186–201.
- 77 Xia L, Sun Y, Li X, *et al.* Wireless resource management in intelligent semantic communication networks. In: Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). New York, 2022.
- 78 Zhang W, Wang Y, Chen M, *et al.* Optimization of image transmission in cooperative semantic communication networks. *IEEE Trans Wireless Commun* 2024; **23**: 861–873.
- 79 Yang Z, Chen M, Zhang Z, *et al.* Energy efficient semantic communication over wireless networks with rate splitting. *IEEE J Sel Areas Commun* 2023; **41**: 1484–1495.
- 80 Zhou F, Li Y, Zhang X, *et al.* Cognitive semantic communication systems driven by knowledge graph. In: Proceedings of the ICC 2022-IEEE International Conference on Communications. Seoul: IEEE, 2022. 4860–4865.
- 81 Jiang S, Liu Y, Zhang Y, *et al.* Reliable semantic communication system enabled by knowledge graph. *Entropy* 2022; **24**: 846.
- 82 Shi G, Gao D, Song X, *et al.* A new communication paradigm: From bit accuracy to semantic fidelity. arXiv: [2101.12649](https://arxiv.org/abs/2101.12649).
- 83 Xiao Y, Sun Z, Shi G, *et al.* Imitation learning-based implicit semantic-aware communication networks: Multi-layer representation and collaborative reasoning. *IEEE J Sel Areas Commun* 2023; **41**: 639–658.
- 84 Klakow D, Peters J. Testing the correlation of word error rate and perplexity. *Speech Communication* 2002; **38**: 19–28.
- 85 Papineni K, Roukos S, Ward T, *et al.* Bleu: A method for automatic evaluation of machine translation. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, 2002. 311–318.

- 86 Vincent E, Gribonval R, Fevotte C. Performance measurement in blind audio source separation. *IEEE Trans Audio Speech Lang Process* 2006; **14**: 1462–1469.
- 87 Rix AW, Beerends JG, Hollier MP, *et al.* Perceptual evaluation of speech quality (PESQ)—A new method for speech quality assessment of telephone networks and codecs. In: Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Salt Lake City, 2001. 749–752.
- 88 Heusel M, Ramsauer H, Unterthiner T, *et al.* Gans trained by a two time-scale update rule converge to a local nash equilibrium. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, 2017.
- 89 Bińkowski M, Sutherland DJ, Arbel M, *et al.* Demystifying mmd gans. arXiv: [1801.01401](https://arxiv.org/abs/1801.01401).
- 90 ITU-R BS.1534-3. Method for the subjective assessment of intermediate quality level of audio systems. Int Telecommun Union Radiocommun Assembly, Geneva, 2015. <https://www.itu.int/rec/R-REC-BS.1534-3-201510-I/en>.
- 91 MeTriX MuX Visual Quality Assessment Package README. Available: http://foulard.ece.cornell.edu/gaubatz/matrix_mux/.
- 92 Wang Z, Bovik AC, Sheikh HR, *et al.* Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process* 2004; **13**: 600–612.
- 93 Zhang R, Isola P, Efros AA, *et al.* The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 586–595.
- 94 Chen LC, Papandreou G, Schroff F, *et al.* Rethinking atrous convolution for semantic image segmentation. arXiv: [1706.05587](https://arxiv.org/abs/1706.05587).
- 95 Seshadrinathan K, Bovik AC. Motion-based perceptual quality assessment of video. In: Proceedings of the Human Vision and Electronic Imaging XIV. San Jose, 2009. 7240: 283–294.
- 96 Lin JY, Liu TJ, Wu ECH, *et al.* A fusion-based video quality assessment (FVQA) index. In: Proceedings of the Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific. Siem Reap: IEEE, 2014. 1–5.
- 97 Pinson MH, Wolf S. A new standardized method for objectively measuring video quality. *IEEE Trans Broadcast* 2004; **50**: 312–322.
- 98 Wolf S, Pinson MH. Video quality model for variable frame delay (VQM-VFD). US Dept Commer, Nat Telecommun Inf Admin, Boulder, CO, USA, Tech Memo TM-11-482, 2011.
- 99 Liu TJ, Lin YC, Lin W, *et al.* Visual quality assessment: Recent developments, coding applications and future trends. *SIP* 2013; **2**: e4.
- 100 Kosta A, Pappas N, Angelakis V. Age of information: A new concept, metric, and tool. *FNT Networking* 2017; **12**: 162–259.
- 101 Soleymani T. Value of Information Analysis in Feedback Control. Dissertation for Doctoral Degree. München: Technische Universität München, 2019.
- 102 Shisher MKC, Qin H, Yang L, *et al.* The age of correlated features in supervised learning based forecasting. In: Proceedings of the IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Vancouver: IEEE, 2021. 1–8.
- 103 Chen X, Zhao Z, Mao S, *et al.* Age of semantics in cooperative communications: To expedite simulation towards real via offline reinforcement learning. arXiv: [2209.08947](https://arxiv.org/abs/2209.08947).
- 104 Xie H, Qin Z, Li GY, *et al.* Deep learning enabled semantic communication systems. *IEEE Trans Signal Process* 2021; **69**: 2663–2675.
- 105 Park J, Choi J, Kim SL, *et al.* Enabling the wireless metaverse via semantic multiverse communication. In: Proceedings of the 2023 20th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). Madrid, 2023. 85–90.
- 106 Zhang P, Xu X, Dong C, *et al.* Intellicise communication system: Model-driven semantic communications. *J China*

- Univ Posts Telecommun* 2022; **29**: 2–12.
- 107 Li A, Wu S, Meng S, *et al.* Towards goal-oriented semantic communications: New metrics, open challenges, and future research directions. arXiv: [2304.00848](https://arxiv.org/abs/2304.00848).
- 108 Xiao W, Tang Y, Liu J, *et al.* Semantic-driven efficient service network towards smart healthcare system in intelligent fabric. *IEEE Trans Netw Sci Eng* 2023; **10**: 2480–2489.
- 109 Li A, Liu X, Wang G, *et al.* Domain knowledge driven semantic communication for image transmission over wireless channels. *IEEE Wireless Commun Lett* 2023; **12**: 55–59.
- 110 Raha AD, Munir MS, Adhikary A, *et al.* An artificial intelligent-driven semantic communication framework for connected autonomous vehicular network. In: Proceedings of the 2023 International Conference on Information Networking (ICOIN). Bangkok: IEEE, 2023. 352–357.
- 111 Xu W, Zhang Y, Wang F, *et al.* Semantic communication for the internet of vehicles: A multiuser cooperative approach. *IEEE Veh Technol Mag* 2023; **18**: 100–109.
- 112 Xia L, Sun Y, Niyato D, *et al.* xURLLC-aware service provisioning in vehicular networks: A semantic communication perspective. *IEEE Trans Wireless Commun* 2024; **23**: 4475–4488.
- 113 Su J, Liu Z, Xie Y, *et al.* Semantic communication-based dynamic resource allocation in D2D vehicular networks. *IEEE Trans Veh Technol* 2023; **72**: 10784–10796.
- 114 Kang X, Song B, Guo J, *et al.* Task-oriented image transmission for scene classification in unmanned aerial systems. *IEEE Trans Commun* 2022; **70**: 5181–5192.
- 115 Yun WJ, Lim B, Jung S, *et al.* Attention-based reinforcement learning for real-time UAV semantic communication. In: Proceedings of the 2021 17th International Symposium on Wireless Communication Systems (ISWCS). Berlin: IEEE, 2021. 1–6.
- 116 Lu Y, Asghar MR. Semantic communications between distributed cyber-physical systems towards collaborative automation for smart manufacturing. *J Manufacturing Syst* 2020; **55**: 348–359.
- 117 Wang J, Du H, Tian Z, *et al.* Semantic-aware sensing information transmission for metaverse: A contest theoretic approach. *IEEE Trans Wireless Commun* 2023; **22**: 5214–5228.
- 118 Liew ZQ, Du H, Lim W YB, *et al.* Economics of Semantic Communication in Metaverse: An Auction Approach. In: Proceedings of the 2023 IEEE 20th Consumer Communications & Networking Conference (CCNC). Las Vegas: IEEE, 2023. 398–403.
- 119 Huang Y, Bai B, Zhu Y, *et al.* ISCom: Interest-aware semantic communication scheme for point cloud video streaming. *IEEE J Sel Areas Commun* 2024; **42**: 1003–1021.
- 120 Huth AG, de Heer WA, Griffiths TL, *et al.* Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 2016; **532**: 453–458.
- 121 Schönberger JL, Pollefeys M, Geiger A, *et al.* Semantic visual localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018. 6896–6906.