Research and Application of General Information Measures Based on a Unified Model

Jianfeng Xu

***Abstract*—To enable comparisons among various information systems, it is necessary to establish general measures based on a unified model. This paper sets out four concise postulates about information, from which a unified information model is derived using axiomatic methods, which provide a theoretical foundation for developing a series of general information measures. By using *volume*, the most commonly used information measure, a formula is derived for the relationship between information and mass, energy, and time in closed systems. This formula can be used to analyze and optimize the power consumption requirements of information systems. Through eleven corollaries, the author also demonstrates that the proposed theory can express classical or commonly used principles in information science and technology for each type of measure. This finding suggests that a unified theoretical system that integrates many classical information principles can be constructed. Finally, the output information from some information systems is tested and evaluated, and the results show that the proposed general information measures can be used to support the design and evaluation of various information systems.**

***Index Terms*—Evaluation of information systems, information measures, information model, objective information theory (OIT), postulates about information, relationships between information and matter, energy and time.**

I. Introduction

THE rapid progression of information science and technology has brought humankind into the information age. Information science is generally considered a discipline that studies the acquisition, processing, transmission, and application of information. It encompasses various fields, such as computer science, communication technology, artificial intelligence, data science, and sensor science. Some of its classical principles, such as Shannon information theory [1], the radar equation [2], the Nyquist sampling theorem [3], Metcalfe's law [4] and Kalman filtering [5], have played substantial roles in the design and development of information systems such as the telegraph, radar, television, the internet, AI systems, and electronic commerce systems in their respective fields. For example, Shannon information theory applies statistical methods to study the basic rules of information transmission and processing in communication systems and solves many technical problems, such as information encoding, data compression, data transmission and data storage in communication and computing systems. However, these classical information principles lack internal cohesion and have not yet formed a unified theoretical system to effectively guide the practice of complex information system engineering. Furthermore, to date, our comprehension of information has varied considerably. For example, *information for a system R is a capacity to cause changes in the system R*, *information is a difference that makes a difference*, *information is the values of characteristics in the processes' output*, *information can be viewed as a collection of symbols* [6], etc. These concepts are extracted from the information systems of certain specific situations. There is no consensus on a universal concept of information that is suitable for increasingly complex information systems of all types. With the increasing extension of information systems to pervade human society, there is an urgent need to employ a set of theoretical methods based on a unified information concept and model to guide and standardize the design, development, application and evaluation of various large-scale and complex information systems.

Our everyday reality is characterized by objective information [7]. Consistent with Winner's famous assertion that *information is information, not matter or energy* [8], a Harvard study noted that *matter, energy, and information are the three main kinds of societal resources: without matter, nothing exists; without energy, nothing happens; and without information, nothing makes sense* [9]. This author proposed the fundamental concept of objective information theory (OIT) [10], established a definition and model of information, analyzed its basic properties, and defined nine measures. In the context of aerospace control systems, the authors of [11] verified the applicability and reasonableness of OIT. In [12], conditions for information restorability were introduced, and eleven types of information measures were revised and expanded, including *volume, delay, scope, granularity, variety, duration, sampling _rate, aggregation, coverage, distortion, and mismatch*. Further efficacy analyses were conducted for these eleven types of measures that an information system can potentially achieve. The structure of information system dynamics (ISD) was also established to provide comprehensive index guidance and model support for the design and integration of large-scale information systems [13].

Axiomatization is a fundamental approach for constructing modern scientific theoretical systems. Can we establish a concise axiomatic hypothesis that derives the basic model and mathematical definition of information based on people's general perception of information? From such a model, can we derive

Manuscript received 24 August 2023; revised 27 December 2023; accepted 31 December 2023. Date of publication 5 January 2024; date of current version 12 February 2024. This work was supported by the National Key R&D Program of China under Grant 2016YFC0800801. Recommended for acceptance by R. Wang.

The author is with the Institute for Smart Courts, Shanghai Jiao Tong University, Shanghai 200030, China (e-mail: xjfcetc@163.com).

Digital Object Identifier 10.1109/TC.2024.3349650

formulas that express the relationships between the three elements of the objective world to support more accurate engineering calculations and designs? Additionally, can we prove the intrinsic relationship between OIT and a series of classical information principles using mathematically well-defined information measures, thus laying the foundation for a unified information science theory? Since information is the common output of all kinds of information systems and since measuring and evaluating information are the basic prerequisites for developing information systems, answering these questions is crucial to our endeavor to establish a set of advanced information science theories that support the engineering of large-scale, complex information systems.

The contribution of this paper is to demonstrate that a unified information model can establish general information measures that can reveal the quantitative relationships among information and matter, energy and time and can accommodate many traditional classical information principles while also having important practical significance. In the second part, four basic postulates on information are proposed, namely, binary attributes, existence duration, state representation, and an enabling mapping. Then, it is proved that the sextuple model defined by OIT satisfies these postulates as necessary and sufficient conditions. In the third part, the information volume defined by the sextuple model is analyzed, and the additivity of the atomic information volume is verified. The information volume formula for a single quantum carrier over a time period is also proven using the Margolus–Levitin theorem. Furthermore, a formula is derived that relates the information volume to mass, energy, and time based on Einstein's mass–energy conversion formula. The proposed theorem reveals more precise relationships among information, matter, energy, and time. Macroscopically, this approach enables us to calculate and analyze the volume of information that the universe can contain at this time. In engineering, this approach can help optimize the power consumption design of information systems. In the fourth section, the relationships between the eleven information measures defined by OIT and related classical or commonly used information principles are demonstrated. It is shown that the general information measures based on the sextuple model can also be used to express many classical principles. In the fifth part, we perform statistical experiments using the eleven types of information measures to test three well-known information systems. The results indicate that general information measures based on the sextuple model can be used to analyze and evaluate the key performance of various information systems and guide the development and construction of large-scale complex information systems.

II. Axiomatic Derivation of the Information Model

While people may have varying opinions on the nature of information, there are commonly accepted examples of information, which are collected, transmitted, processed, stored, and utilized in computer and internet systems. These examples include various types of numerical data; texts; and audio, video, and multimedia files. The author proposes that matter is the fundamental form of existence, energy represents the capability of motion, and information reflects the content of both the objective and subjective worlds. Matter, energy, and information are the three essential elements of the objective world [10]. Using an axiomatic approach, we propose four postulates to facilitate a potential consensus on the nature of information and develop rigorous information models:

Postulate 1: (Binary Attribute). Every piece of information comprises two components: the noumena, which represents the intrinsic essence of the information, and the carrier, which embodies the objective form of the information.

An important distinction between information and matter or energy is that while matter and energy can exist independently, information always has a "shadow" lurking behind it. Questions are often raised about the validity of a specific piece of information, such as "Is this information true or false?" and "How accurate is this information?". Indeed, the "surface" of information is always compared with the "shadow" behind it, and the fundamental significance of information application is related to the ability to recover and apply the truth hidden behind the apparent outer form of information. It is important to note that the "objective form" of information requires that the apparent outer form cannot be altered by people's subjective will. Thus, the carrier of information must be something that exists in the objective world. This serves as the basis for Postulate 1.

Postulate 2: (Existence Duration). The content and form of information have their own respective temporal existences.

Like matter and energy, a particular piece of information has a temporal existence, regardless of whether the noumena or the objective form in the carrier is considered. The duration of existence has significant implications for information. For instance, the study of information delay relies on the relationship between the durations of existence of the noumena and the carrier. In numerous applications of information, such as disaster weather forecasting, military intelligence, and document exchange, this delay can have a decisive impact. This forms the basis for Postulate 2.

Postulate 3: (State Representation). The states of the noumena and the carrier of information have their own respective representations.

One key distinction between information and matter or energy is that the essence of the content of matter is still matter, and the essence of the content of energy is still energy, while the essence of the content of information is the state of events or objects. The states of both the noumena and the carrier should be taken into account because of the binary attribute nature of information. Therefore, if we want to analyze the authenticity and accuracy of information, we must compare the states of the noumena and the carrier. This is the foundation for Postulate 3.

Postulate 4: (Enabling Mapping). The state of the noumena can be mapped onto the state of the carrier via an enabling mapping. This implies that a surjective mathematical mapping can be established from the state of the noumena to the state of the carrier, and the state of the carrier becomes objective reality due to the state of the noumena.

In the absence of any relation between the states of the noumena and the carrier, information cannot exist. Only when a surjective mapping relationship between the two states is established mathematically and a cause-and-effect relationship can be physically formed between the former and the latter can information come into existence. This is the fundamental principle underlying Postulate 4.

Postulates 1–4 establish the fundamental components and relationships of information, which serve as the basis for deriving the following sextuple model of information.

Lemma 1: The necessary and sufficient condition for the existence of an information model satisfying Postulates 1–4 is the existence of an enabling mapping I such that the set of states of noumena $f = f(o, T_h)$ is surjectively mapped to the set of carrier reflection states $g = g(c, T_m)$, where o is a nonempty set of information noumena, T_h is the set of times when the noumena in o occur, c is a set of objective carriers, and T_m is the set of reflection times of carriers.

We call the mapping I, denoted by $I = \langle o, T_h, f, c, T_m, g\rangle$, the sextuple model of information.

Proof: If Postulates 1–4 are satisfied, we have nonempty sets o and c by Postulate 1. Postulate 2 yields the sets T_h and T_m, the durations for the occurrences of the noumena and reflections in the carriers. Postulate 3 provides the set of states of the noumena, denoted by $f(o, T_h)$, which depends on the set of noumena o and the set of noumena occurrence times T_h, as well as the set of states of reflections in the carrier, denoted by $g(c, T_m)$, which further depends on c, the set of carriers, and T_m, the set of reflection times. According to Postulate 4, surjective enabling mappings exist from $f(o, T_h)$ to $g(c, T_m)$. We denote these mappings in I and obtain the sextuple model $I = \langle o, T_h, f, c, T_m, g\rangle$. The necessity is proven.

On the other hand, if the sets o, c, T_h, T_m, $f(o, T_h)$, and $g(c, T_m)$ exist together with a surjective enabling mapping from $f(o, T_h)$ to $g(c, T_m)$, then the sextuple model $I = \langle o, T_h, f, c, T_m, g\rangle$ clearly satisfies all four postulates. Sufficiency is also proven.

Based on Postulates 1–4 and Lemma 1, we can now provide a precise and rigorous mathematical definition of information within the framework of OIT.

Definition 1: (Mathematical definition of information). Let O denote the set of all content of the objective world, let S denote the set of content of the subjective world, and let T be a set of times. The elements in these sets can be specified according to the requirements in different areas. A noumenon is thus an element of the power set $2^{O \cup S}$ (or a subset of $O \cup S$), the occurrence duration is $T_h \in 2^T$, and $f(o, T_h)$ is the set of states of o over T_h. The carrier $c \in 2^O$, the set of reflection times $T_m \in 2^T$, and the set of reflection states $g(c, T_m)$ are all nonempty sets. Information I is thus an enabling mapping from $f(o, T_h)$ to $g(c, T_m)$, which is expressed as $I : f(o, T_h) \to g(c, T_m)$ or $I(f(o, T_h)) = g(c, T_m)$. The collection of all information is referred to as the information space, denoted by $\mathfrak{I}$, which is one of the three essential elements of the objective world according to OIT.

Notably, the mapping I must be a surjective mapping from $f(o, T_h)$ to $g(c, T_m)$ in the mathematical sense. However, only when $g(c, T_m)$ is created through $f(o, T_h)$ in the objective world by self-excitation or an external force can I be considered information. This is why the concept of enabling mapping is introduced in Definition 1. Therefore, a purely mathematical surjective mapping from $f(o, T_h)$ to $g(c, T_m)$ is not information if $g(c, T_m)$ has no physical connection to $f(o, T_h)$. Surjectivity is a necessary but not sufficient condition for information. Information is both mathematical and physical [14], and this is its essential property.

Definition 2: (Restorable information). If the information $I = \langle o, T_h, f, c, T_m, g\rangle$ is also an injective mapping from $f(o, T_h)$ to $g(c, T_m)$, i.e., for any $o_\lambda, o_\mu \in o, T_{h\lambda}, T_{h\mu} \in T_h, f_\lambda, f_\mu \in f$, then when $f_\lambda(o_\lambda, T_{h\lambda}) \neq f_\mu(o_\mu, T_{h\mu})$, one must have $I(f_\lambda(o_\lambda, T_{h\lambda})) \neq I(f_\mu(o_\mu, T_{h\mu}))$. In this case, I is an invertible mapping, meaning that its inverse I^{-1} exists. For any $c_\lambda \in c, T_{m\lambda} \in T_m, g_\lambda \in g$, there is a unique $o_\lambda \in o, T_{h\lambda} \in T_h, f_\lambda \in f$ such that $I^{-1}(g_\lambda(c_\lambda, T_{m\lambda})) = f_\lambda(o_\lambda, T_{h\lambda})$. As a result, $I^{-1}(g(c, T_m)) = f(o, T_h)$. We refer to information I as restorable, and the restored state of information I is $f(o, T_h)$.

We observe that by utilizing $g(c, T_m)$ and I^{-1}, the state of o over $T_h, f(o, T_h)$, can be restored. This represents the restorability of information. In practical scenarios, people rely on information to search for the restored state, which embodies the essential property and significance of information.

III. Relationships Between Information and Matter, Energy and Time

As matter, energy, and information are the three fundamental elements of the objective world, and time is the essential condition for carrying them, their interrelationships have become a fundamental issue in scientific research. The sextuple model of information in OIT defines the volume of information as a measure of its reflection set [12]. Similar to mass, energy, and time, the volume is the most appropriate measure to express the relationships between these elements.

A. Combination of Atomic Information and Volume Additivity

The representation of information as a sextuple model $I = \langle o, T_h, f, c, T_m, g\rangle$ does not emphasize its nature as a set. We simply need to modify the notation slightly to see that $I = \{\langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda\rangle\}$, where $o_\lambda \in o$, $T_{h\lambda} \in T_h$, $f_\lambda(o_\lambda, T_{h\lambda}) \in f(o, T_h)$, $c_\lambda \in c$, $T_{m\lambda} \in T_m$, $g_\lambda(c_\lambda, T_{m\lambda}) \in g(c, T_m)$, and $\lambda \in \Lambda$ is an index set. We then see that I is in fact a set of elements with six components. The concepts of subinformation and proper subinformation can be defined similarly to subsets and proper subsets in set theory. Just as matter can be broken down into indivisible elementary particles and energy into indivisible quanta, any information can also be decomposed to the most basic level, at which it cannot be further divided. The fundamental unit of information is referred to as atomic information (Fig. 1). Atomic information does not imply that information is associated with an atomic carrier or noumenon; rather, it is the smallest and most basic component of the information space. Therefore, atomic information plays a crucial role in understanding the composition and nature of information. Any information can be considered a combination of atomic information, and there is no overlap among the atomic information. Building upon the principles of set additivity and measure theory [15]错误!未找到引用源。,

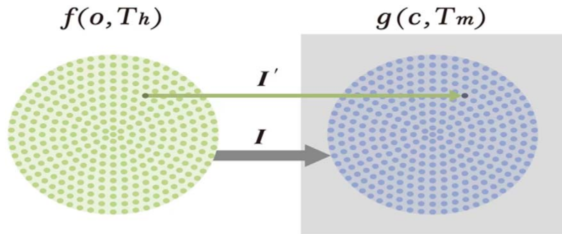

Fig. 1. Illustration of the concept of atomic information.

we define the additivity of a volume of atomic information and its combinations as follows:

When $g_\lambda(c_\lambda, T_{m\lambda})$ is σ-measurable for any atomic information $I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle$ of restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$, where $\lambda \in \Lambda$, which is a countable index set, then we have

$$\text{Volume}_\sigma(I) = \sum_{\lambda=1}^{\infty} \text{volume}_\sigma(I_\lambda). \quad (1)$$

When Λ is a finite set, we can simply change the superscript ∞ in the sum (1) to a finite number representing Λ's cardinality. However, if Λ is an uncountable set, Formula (1) cannot be simply changed into an integral representation since, in general, the measure does not have additivity with regard to an uncountable set.

B. Information Volume That an Individual Quantum Can Carry

Quantum information theory is an important achievement of contemporary information technology. It is commonly agreed that there is no radiation with energy less than a photon [16]. A single quantum such as a fermion (e.g., a quark or lepton) or boson (e.g., gluon, w-boson, photon, or graviton) [17] is a basic element that cannot be further divided in terms of energy; thus, it is the most efficient carrier of atomic information.

Corollary 1: (Information volume of a single quantum carrier). For restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$, when carrier c is a single quantum, I is called single quantum carrier information. We define the measure σ as the number of distinguishable states experienced by c over reflection duration T_m; let $t = \sup T_m - \inf T_m$ be the reflection duration of information I. Then, the upper bound of the volume of information I with respect to measure σ is

$$\max\{\text{volume}_\sigma(I)\} = \begin{cases} 4\Delta Et/h, & t \text{ is sufficiently large;} \\ 1, & t \text{ is or is approaching } 0. \end{cases} \quad (2)$$

where the unit of $\text{volume}_\sigma(I)$ is a qubit, ΔE is the average energy of quantum c, and $h \approx 6.6 \times 10^{-34}$ j/s is the Planck constant.

Proof: According to the definition of information volume, $\text{volume}_\sigma(I)$ is the number of distinguishable states experienced by quantum carrier c over reflection duration T_m. The state of a quantum is composed of two mutually orthogonal basis states $|0\rangle$ and $|1\rangle$ and their quantum superposition $a|0\rangle + b|1\rangle$, where a and b are complex numbers satisfying $|a|^2 + |b|^2 = 1$ [18]. However, the distinguishable states of a quantum must be orthogonal [19]. Thus, by the Margolus–Levitin theorem [20], the shortest time delay from one state to another orthogonal state depends on the associated average energy ΔE. The transition time is no shorter than $\Delta t = h/4\Delta E$. Since each distinguishable state of a single quantum can carry exactly one qubit of information [21], for carrier c, if the average energy is ΔE, then for a reflection duration t, if $t < \Delta t$, c can display only one state that is distinguishable from the others. This means that $\text{volume}_\sigma(I) = 1$ (qubit), i.e., c can carry only one piece of qubit information. This proves the second case of (2).

More generally, for any longer reflection duration t, we have $\max\{\text{volume}_\sigma(I)\} = [t/\Delta t] + 1 = [4\Delta Et/h] + 1$, where $[\cdot]$ denotes the function that takes the integer part of a rational number. Since $h \approx 6.6 \times 10^{-34}$, when t is sufficiently large, for example, $\Delta Et \geq 10^{-30}$, according to the preceding formula, the relative error between $[4\Delta Et/h] + 1$ and $4\Delta Et/h$ is smaller than the scale of 10^{-4} and is negligible. Thus, we obtain $\max\{\text{volume}_\sigma(I)\} \approx 4\Delta Et/h$ (qubit). This proves the first case of (2).

Corollary 1 illustrates that the capacity of a quantum carrier to carry distinct information is bounded by its corresponding energy, thereby limiting the amount of information that can be carried over any given time interval. Furthermore, given that the number of quanta in the universe is finite and that the information carried by a quantum carrier constitutes the fundamental building block of information, it can be mathematically proven that even with an assumed infinite future, any information can be decomposed into, at most, a countable number of atomic information units. Therefore, the additivity of volume for information combinations holds true under all circumstances.

C. Information Volume That a General Carrier Can Carry

Since the discovery of cosmic expansion by Hubble [22], the standard assumption has been that all energy in the universe is in the form of radiation or ordinary matter [23]. Any matter or radiation in the objective world can be a carrier of information, and the information volume that a carrier can carry depends on its physical properties and the level of technology [24]. At present, silicon chips are commonly the carriers with the largest carrying volume in information systems. With the most advanced manufacturing technology, the mass of a silicon chip is approximately 1.6 grams, and its storage volume can reach 10^{12} bits. Therefore, the information volume that a kilogram of silicon chips can carry is 6.25×10^{14} bits.

According to the principles of thermodynamics and the mass–energy conversion formula [25], the minimum mass of matter required to store one bit is $m_{bit} = k_b T \ln(2)/C^2$, where $k_b \approx 1.38 \times 10^{-23}$ j/ kg is Boltzmann's constant, T is the absolute temperature of the information carrier, and $C \approx 3.0 \times 10^8$ m/s is the speed of light in a vacuum. Therefore, if the carrier of information I is matter with a mass of 1 kilogram and its energy does not dissipate, the upper bound of the information volume that can be carried is always $\text{volume}_\sigma(I) = 1/m_{bit} = C^2/k_b T \ln(2)$ bits, where the measure σ is in bits. When T is at a normal temperature of 300 K^o, $\text{volume}_\sigma(I)$ is approximately 10^{37} bits. Because this formula is derived from principles of classic thermodynamics, it is applicable only to the equilibrium state of classic digital memory and cannot be applied to the case of quantum carriers such as electrons and photons.

For information with a quantum as its carrier, Formula (2) provides the information volume that a single quantum carrier can hold at a specific time. Due to the disjoint property of quantum carriers, the additivity of the information volume shown in Formula (1) can help us estimate the information volume that a quantum carrier can hold.

Theorem 1: (The relationship of information, matter, energy, and time). Assume that c is the carrier of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ and that one part of c exists in the form of matter and the other part exists in the form of radiation, with mass m and energy E_r, respectively. Assume further that E is the total energy of c, measure σ is the number of states of c over the reflection time duration T_m, and t is the reflection duration of information. Then, when t is sufficiently large, $\max\{\mathrm{volume}_\sigma(I)\} = 4Et/h = 4(m\mathrm{C}^2 + E_r)t/h$ (qubit). To express this formula in a simpler way, we let I directly represent the upper bound of its information volume. Then, a very concise formula can be obtained:

$$I = \frac{4Et}{h} = \frac{4\left(m\mathrm{C}^2 + E_r\right)t}{h},\ t \text{ is sufficiently large} \tag{3}$$

where I has a unit of qubits, C is the speed of light and h is the Planck constant.

Proof: The total energy of carrier c is E for restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$. This total energy remains constant at any moment in the reflection duration T_m according to the basic principle of the conservation of energy [17]. Without loss of generality, we can assume that carrier c is composed of N quanta and that each individual quantum has an average energy ΔE; then, the number of quanta is $N = E/\Delta E$ within the reflection duration T_m. We can also assume that $c = \cup_{i=1}^{N} c_i$, where each $c_i (i = 1,,, N)$ is a single quantum that serves as a carrier in the reflection duration T_m. Let I_i denote the subinformation carried by c_i. Each I_i is a single quantum carrier's information, and t is the reflection duration between information I and subinformation I_i. From Equation (2), we have $\max\{\mathrm{volume}_\sigma(I_i)\} = 4\Delta Et/h$ when t is sufficiently large. Since we also have $I = \cup_{i=1}^{N} I_i$, each c_i is a distinct quantum, this subinformation does not overlap, and I_i is a combination of $[4\Delta E\mathrm{t}/\mathrm{h}] + 1$ pieces of atomic information. By the additivity of measure σ over mutually disjoint sets and Formula (2), we have the upper bound of the volume of information I:

$$\max\{\mathrm{volume}_\sigma(I)\} = \sum_{i=1}^{N}\max\{\mathrm{volume}_\sigma(I_i)\} \approx \sum_{i=1}^{E/\Delta E} 4\Delta Et/\mathrm{h}$$
$$= 4Et/h(\text{qubit}),\ t \text{ is sufficiently large}$$

Furthermore, the total energy E of c is the sum of the energy contained in its matter form and the energy in its radiation form. According to the matter-energy conversion formula [26], [27], the matter form has energy $m\mathrm{C}^2$, where C is the speed of light. Then, $E = m\mathrm{C}^2 + E_r$. Thus, when t is sufficiently large, we have $\max\{\mathrm{volume}_\sigma(I)\} = 4(m\mathrm{C}^2 + E_r)t/h$ (qubit). The theorem is proved.

D. Discussion

Since Formula (3) covers the basic measures of information, matter, energy, and time with a simple and clear expression, we call this the relationship formula of information, matter, energy, and time. Imagine that we have an information carrier that can be decomposed into pure quanta [17] and that the energy it possesses is equivalent to 1 kg of matter. By Formula (3), the upper bound of the information volume carried over 1 second is $4\mathrm{C}^2/\mathrm{h} \approx 5.3853\times 10^{50}$ (qubit).

The universe is full of all kinds of information and is thought to be a gigantic quantum computer. In recent decades, researchers have explored the question of how much information the whole universe may have possibly carried to date [28]. The formula for the relationship between information and matter, energy, and time from Theorem 1 and the basic principle of energy conservation enable us to answer this question quite easily.

Standard inflationary theory predicts that the universe is spatially flat. Einstein's general relativity theory determines that such a universe has a total energy density equal to the critical density $\rho_c = 3\mathrm{H_0}^2/8\pi\mathrm{G}$ [23], where $\mathrm{H}_0 \approx 2.1\times 10^{-18}$/s is the current Hubble parameter and $\mathrm{G} \approx 6.7\times 10^{-11}\ \mathrm{m}^3/\mathrm{kgs}^2$ is the gravitational constant, i.e., $\rho_c \approx 7.9\times 10^{-27}\ \mathrm{kg/m}^3$. Additionally, the radius of the observable universe is approximately $L = 4.56\times 10^{10}$ light years [29], from which the volume of the whole universe can be estimated as $V \approx (4/3)\,\pi L^3 \approx 3.35\times 10^{80}\ \mathrm{m}^3$. Therefore, the total mass of the universe can be estimated as $m = \rho_c V \approx 2.6\times 10^{54}$ kg. According to common understanding, the age of the universe is approximately 13.7 billion years [30]; that is, the information reflection duration is $t \approx 4.3\times 10^{17}$ s. According to Formula (3), the upper bound of the information volume carried by the universe to date is $I = 4m\mathrm{C}^2 t/h \approx 4\times 2.6\times 10^{54}\times \left(3.0\times 10^8\right)^2\times 4.3\times 10^{17}/(6.6\times 10^{-34}) \approx 6.1\times 10^{122}$ (qubit). This estimate is almost the same as the number of logical operations of the universe given in [28]: 10^{123}. According to OIT, any logical operation of the universe must create a particular state, and these states must contain objective information. Therefore, the number of logical operations of the universe at a particular time is the same as the number of states the universe has at that time. On the other hand, this estimate is not the same as the 10^{90} bits of information capacity of the universe estimated in [28]. This is because the information capacity is defined through entropy [31], which differs greatly from the definition of information volume in Theorem 1. By comparison, we believe that it is more universally applicable to define the information volume as "the measure of all states of the carrier over the reflection time". More importantly, Theorem 1 establishes the general relationships among the three major elements, namely, matter, energy, and information about the objective world, and time. For any information carrier, if the mass, energy, and time are known, we can apply Formula (3) to obtain the upper bound of the information volume.

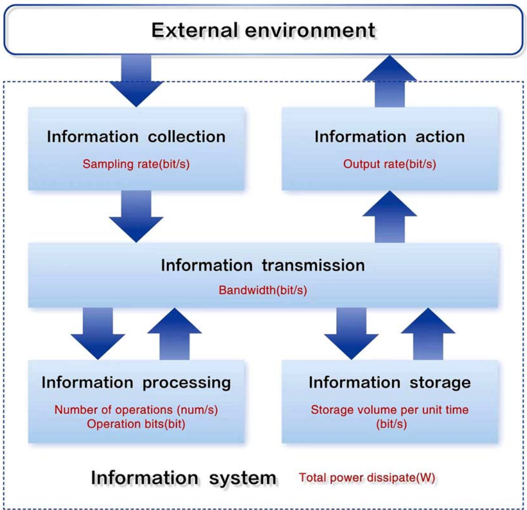

Fig. 2. Illustration of a general information system structure.

Theorem 1 has significant implications for the optimization of power dissipation and design allocation in information systems. As depicted in Fig. 2, an information system typically comprises five components: information collection, transmission, storage, processing, and action. The total information volume carried by the system is equal to the sum of the information volume of each component, while the total power dissipation of the information system is also equal to the sum of the power dissipation of each component.

For an information system, let its sampling rate be C, the internal communication bandwidth be T, the storage volume per unit time be S, the operation speed of information processing be P, the number of bits of operation be B, and the output rate be O. Then, for any given duration t, the information volumes carried by each part are Ct, Tt, St, PBt and Ot, respectively. According to Theorem 1, we have

$$Ct + Tt + St + PBt + Ot \leq 4Et/h \tag{4}$$

where E is the total energy consumed by the information system at time t and h is the Planck constant. The ratio factor of volume to energy (RFVE) of the information system is defined as ρ_{total}, and $\rho_{\text{c}} = Ch/(4E)$, $\rho_{\text{T}} = Th/(4E)$, $\rho_{\text{P}} = PBh/(4E)$, $\rho_{\text{S}} = Sh/(4E)$ and $\rho_{\text{O}} = Oh/(4E)$ are the five ratio factors of the volume to the energy of each part. These factors reflect the power consumed by the corresponding system in carrying information, which is an important index to be considered in system implementation.

According to Formula (4), we can obtain

$$\rho_{\text{total}} = \rho_{\text{c}} + \rho_{\text{T}} + \rho_{\text{S}} + \rho_{\text{P}} + \rho_{\text{O}} \leq 1 \tag{5}$$

Equations (4) and (5) give the constraint conditions between the volume of carrying information and the energy consumption of the general information system and its components. Although the formulas are derived from an information system with a qubit as the volume unit, they can also be used for studying traditional information systems with a bit as the volume unit. We studied the parameters related to the information volume and power consumption of laptops, servers, clouds, supercomputers and quantum computers and obtained the RFVEs, as shown in Table I.

TABLE I
THE RATIO FACTORS OF VOLUME TO ENERGY (RFVE) OF SOME TYPICAL COMPUTING SYSTEMS

Object Type	Laptop	Server	Cloud	Super-Computer	Quantum Computer
Sampling rate (bit/s, qubit/s)	2.7×10^9	6.0×10^{10}	8.0×10^{10}	5.0×10^8	1.0×10^6
Communication bandwidth (bit/s, qubit/s)	3.4×10^{10}	3.2×10^{10}	8.6×10^{10}	1.6×10^{11}	—
Storage volume per unit time (bit, qubit)	3.4×10^{10}	3.2×10^{10}	8.6×10^{10}	1.6×10^{11}	—
Number of operations (num/s)	3.0×10^{10}	4.5×10^{13}	5.3×10^{15}	1.4×10^{17}	1.0×10^6
Operation bits (bit, qubit)	64	64	64	64	20
Output rate (bit/s, qubit/s)	2.7×10^9	6.0×10^{10}	8.0×10^{10}	5.0×10^8	1.0×10^4
Total power dissipation (W)	65	526	51311	1212000	25000
Total RFVE	5.1×10^{-24}	9.0×10^{-22}	1.1×10^{-21}	1.2×10^{-22}	1.4×10^{-31}

Table I shows that the RFVEs of various computing systems at present are far less than the theoretical upper limit of 1.0, indicating that the information system has a large amount of other energy consumption in addition to carrying information and that there is still much space for energy savings and consumption reduction. The RFVE of a laptop is 2-3 orders of magnitude lower than those of servers, clouds and supercomputers, and there is a possibility of further energy savings. Because a supercomputer needs to provide intensive supercomputing power, its RFVE is lower than that of the cloud by 1 order of magnitude. Quantum computers do not involve internal communication or storage units and are still in the stage of experimental exploration. Their RFVE is far lower than that of traditional computing systems, and reducing energy consumption is a key problem that needs to be solved in the future.

IV. The Unification of Classical Information Principles Based on General Information Measures

The original intention of OIT was to enrich and improve general measures of information and support the efficacy analysis and research of information systems. Following the sextuple model of information, eleven types of general information measures can be concretely defined [12]. We demonstrate that for each measure, a corresponding example can be found from classical or commonly used information principles, which also shows that OIT can accommodate many existing information science theories and technical methods to form a unified theoretical framework.

A. The Volume of Information and Shannon Information Entropy

Volume is the most commonly used information measure [32]. The volume of restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$ can be defined as a measure of its reflection set; i.e., $\text{volume}_\sigma(I) = \sigma(g(c, T_m))$.

Corollary 2: (Minimum restorable volume of random event information). Let X be an event taking possible values x_i randomly with a probability of $X = x_i$ equal to p_i, $i = 1, ..., n$. Assume that events $\{X = x_i\}$ and $\{X = x_j\}$ are independent when $i \neq j$ and $\sum_{i=1}^{n} p_i = 1$, and information $I = \langle o, T_h, f, c, T_m, g\rangle$ represents the code of X's values transmitted over a communication channel. In this case, I is called random event information. In this information I, the noumenon o is a random event X, the occurrence duration T_h is the occurrence duration of event X, and the state set $f(o, T_h)$ consists of values x_i and $i = 1, ..., n$. The carrier c is the channel for transmitting the value of X, the reflection duration T_m is the time needed for channel c to transmit the value of X, and the reflection set $g(c, T_m)$ is the set of specific codes for channel c to transmit the value X. If measure σ represents the number of bits of $g(c, T_m)$, then the minimum volume for which I can be restored is $\text{volume}_\sigma(I) = -\sum_{i=1}^{n} p_i \log_2 p_i$.

Proof: As in [1], we assume that the semantics of communication are independent of engineering problems. Therefore, to minimize the required channel bandwidth, the communication process for transmitting the value of X does not need to directly transmit the specific value of x_i but only needs to use different binary codes to represent the event that X takes the value of x_i, $i = 1, ..., n$, and then transmit the corresponding codes during the communication process. That is, as long as an appropriate coding method for $g(c, T_m)$ is selected, the required channel bandwidth can be reduced.

We now prove the corollary using a proof by contradiction. Assume, on the contrary, that there is also some encoding of $g(c, T_m)$ such that $H' = \sigma(g(c, T_m)) < -\sum_{i=1}^{n} p_i \log_2 p_i = H$ and I remains as information that can be restored (i.e., the communication channel can transmit the source information completely), where H is the information entropy of event X. Assume that the bandwidth of channel c is W; then, we have $W/H' > W/H$. That is, channel c can completely transmit the source information at a rate greater than W/H. This obviously contradicts the conclusion of Theorem 9 of [1] that the channel transmission rate W/H cannot be exceeded.

Therefore, there is no encoding method such that $\sigma(g(c, T_m)) < -\sum_{i=1}^{n} p_i \log_2 p_i$ while keeping I as restorable information. Additionally, $\text{volume}_\sigma(I) = -\sum_{i=1}^{n} p_i \log_2 p_i$ is the minimum information volume of I. The corollary is proved.

Corollary 1 shows that the information volume defined by OIT can also express Shannon's information entropy principle, and the information volume defined on the basis of information entropy is simply a special case of information volume. The definition of information volume requires only that I be a one-to-one surjective mapping from $f(o, T_h)$ to $g(c, T_m)$ and that $g(c, T_m)$ be a measurable set with respect to measure σ. Compared with information entropy, this definition has many fewer mathematical constraints, so it should have broader applications.

B. The Delay of the Information and Serial Information Transmission Chain

Delay is also a commonly used information measure [33]. The delay of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ can be defined as the difference between the supremum of its reflection time and the supremum of its occurrence time; i.e., $\text{delay}(I) = \sup T_m - \sup T_h$.

According to the transitivity of information, we can let a set $\{I_i = \langle o_i, T_{hi}, f_i, c_i, T_{mi}, g_i\rangle | i = 1, n\}$ be a set of restorable information [12]. If, for any $i < n$, we obtain $c_i = o_{i+1}$, $T_{mi} = T_{h(i+1)}, g_i = f_{i+1}$, then $\{I_i = \langle o_i, T_{hi}, f_i, c_i, T_{mi}, g_i\rangle | i = 1, n\}$ is the serial information transmission chain between information I_1 and I_n, and there exist $\{I_i' = \langle o_1, T_{h1}, f_1, c_i, T_{mi}, g_i\rangle | \ i = 1, n\}$, which are pieces of restorable information and have the same restored state $f_1(o_1, T_{h1})$.

Corollary 3: (Serial information transfer delay) Let $\{I_i = \langle o_i, T_{hi}, f_i, c_i, T_{mi}, g_i\rangle | i = 1, , n\}$ be the serial information transmission chain between information I_1 and I_n. Then, $I = \langle o_1, T_{h1}, f_1, o_n, T_{mn}, g_n\rangle$ is also a piece of restorable information, and $\text{delay}(I)$ is the sum of the delays of all the information I_i.

Proof: For the serial information transmission chain $\{I_i = \langle o_i, T_{hi}, f_i, c_i, T_{mi}, g_i\rangle | i = 1, , n\}$, by its definition, for any $i < n$, we have $T_{mi} = T_{h(i+1)}$. Therefore, we have $\sum_{i=1}^{n} \text{delay}(I_i) = \sum_{i=1}^{n} (\sup T_{mi} - \sup T_{hi}) = \sup T_{mn} - \sup T_{hn} + \sum_{i=1}^{n-1} (\sup T_{h(i+1)} - \sup T_{hi}) = \sup T_{mn} - \sup T_{h1} = \text{delay}(I)$. The corollary is proved.

C. The Scope of Information and the Radar Equation

Scope is a measure often used for information systems [34]. The scope of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ can be defined as the measure of its noumena; i.e., $\text{scope}_\sigma(I) = \sigma(o)$.

Corollary 4: (Extent of radar detection information). Let restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ be the radar detection information, let noumena o be the object detected by the radar, let the state occurrence time T_h be the time when the radar beam shines on the detected object, and let the state set $f(o, T_h)$ be the state of the detected object itself and its motion. In this case, carrier c is the radar; the reflection time T_m is the time required for the radar to receive, process, store, and display the echo signal and/or the data of the detected object; and the reflection set $g(c, T_m)$ is the echo signal and/or data of the detected object for the radar to receive, process, store, and display. In addition, we define the measure σ of the noumena o as its reflection area. According to the radar equation, when the radar transmitting power, antenna gain, antenna effective aperture, and minimum detectable signal are determined, the maximum detection range of the radar depends on the $\text{scope}_\sigma(I)$ of information I and is proportional to its quartic root.

Proof: According to the definition of radar detection information $I = \langle o, T_h, f, c, T_m, g\rangle$, noumenon o is the object detected by radar, and the $\text{scope}_\sigma(I)$ of I is the reflection area σ of o. In the radar equation $R_{max}^4 = P_t G_t A_e \sigma / (4\pi)^2 S_{min}$ [2], R_{max} is the maximum detection range of the radar, P_t is the radar transmitting power, G_t is the radar antenna gain, A_e is the effective aperture of the radar antenna, S_{min} is the minimum detectable signal of the radar, and σ is the reflection area of the detected object. When the important parameters P_t, G_t, A_e, and S_{min} of the radar itself are determined, the maximum detection range of the radar is determined entirely by σ and is proportional to the fourth root of the scope of information I. The corollary is proved.

D. The Granularity of Information and the Rayleigh Criterion for Optical Imaging

Granularity is also a measure often used for information systems [35]. The granularity of restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$ can be defined as the average of all noumenon measures of its atomic information; i.e., $\text{granularity}_\sigma(I) = \int_\Lambda \sigma(o_\lambda) \mathrm{d}\mu / \mu(\Lambda)$, where $o_\lambda (\lambda \in \Lambda)$ is a noumenon of the atomic information, Λ is the index set, and μ is the measure of Λ.

Corollary 5: (Resolution of optical imaging information). Let the restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$ be the optical imaging information, the noumenon o be the object to be photographed or filmed, the state occurrence time T_h be the time when the shutter is opened or the camera acquires the video, and the state set $f(o, T_h)$ be the state of the photographed object itself and its motion. In this case, carrier c is an image or video in the camera; the reflection time T_m is the time required by the camera to shoot, process, store, and display the image of the target object; and the reflection set $g(c, T_m)$ is the image or video of the target object. Here, we define the measure σ of noumenon o as the minimum distinguishable angle at the time when the target object is photographed; then, the resolution of the optical imaging information I, i.e., $\text{granularity}_\sigma(I)$, is proportional to the wavelength of light and inversely proportional to the width of the photosensitive unit.

Proof: For an object o to be shot, each frame in the optical imaging information $I = \langle o, T_h, f, c, T_m, g \rangle$ contains a large number of pixel points, and each pixel point is a local image of o that cannot be subdivided. Thus, the pixel represents the atomic information $I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle$, where $\lambda \in \Lambda$ and Λ is the index set. By definition, the $\text{granularity}_\sigma(I)$ of information I is the average of the measure $\sigma(o_\lambda)$ of the noumena o_λ of all the atomic information I_λ. The principle of optical imaging indicates that $\sigma(o_\lambda)$ is the same for all $\lambda \in \Lambda$, and according to the Rayleigh criterion [36], $\sigma(o_\lambda) = l/a$, where l is the wavelength of light and a is the width of the photosensitive unit. Thus, the $\text{granularity}_\sigma(I) = \sigma(o_\lambda)$ (for any $\lambda \in \Lambda$) of information I is proportional to the wavelength of light and inversely proportional to the width of the photosensitive unit. The corollary is proved.

E. Invariance Principle of Restorable Information Variety

The type of information contained is a frequently used measure for information systems [37], and it can be expressed in terms of variety. The variety of restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$ can be defined as the number of equivalence classes on its set of states; i.e., $\text{variety}_R(I) = \overline{[f(o,T_h)]_R}$, where R is an equivalence relation on the set of states $f(o, T_h)$.

Corollary 6: (Invariance of restorable information variety). For restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$, let R be the equivalence relation on the state set $f(o, T_h)$; the set of equivalence classes of the elements in $f(o, T_h)$ relative to R is $[f(o,T_h)]_R$. Then, there must exist an equivalence relation Q on the reflection set $g(c, T_m)$ such that set $[g(c,T_m)]_Q$, the set of equivalence classes of the elements in $g(c, T_m)$ relative to Q, and set $[f(o,T_h)]_R$ form a one-to-one surjective relation with information I so that the cardinalities of the two equivalence classes are equal. This is expressed as $vriety_R(I) = \overline{[f(o,T_h)]_R} = \overline{[g(c,T_m)]_Q}$.

Proof: It is easy to prove that the equivalence relation Q on reflection set $g(c, T_m)$ can be established according to the equivalence relation R on the state set $f(o, T_h)$. For any two pieces of subinformation of I, $I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle$ and $I_\mu = \langle o_\mu, T_{h\mu}, f_\mu, c_\mu, T_{m\mu}, g_\mu \rangle$, if $f_\lambda(o_\lambda, T_{h\lambda}) R f_\mu(o_\mu, T_{h\mu})$, there must be $g_\lambda(c_\lambda, T_{m\lambda}) Q g_\mu(c_\mu, T_{m\mu})$. Thus, there is also an equivalence class $[g(c,T_m)]_Q$ on $g(c, T_m)$. Note that the equivalence relation Q is entirely established based on the mapping relation of information I. Thus, the one-to-one surjective relation between the two equivalence classes $[f(o,T_h)]_R$ and $[g(c,T_m)]_Q$ can also be established entirely based on information I. Hence, we have $\overline{[f(o,T_h)]_R} = \overline{[g(c,T_m)]_Q}$. The corollary is proved.

F. Average Duration of Continuous Monitoring Information

The duration of information is an important measure for some systems [38]. The duration of restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$ can be defined as the difference between the supremum and infimum of T_h; i.e., $\text{duration}(I) = \sup T_h - \inf T_h$.

Corollary 7: (Average duration of continuous monitoring information). The average duration of continuous monitoring information is equal to the MTBF of the information collection device.

Proof: Let restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$ be continuous monitoring information, where the noumena o can be considered the monitored object, the state occurrence time T_h can be considered the time period when o is in the monitored state, the state set $f(o, T_h)$ can be considered the state when o is in the monitored time period, and carrier c can be considered the information collection device. In addition, the reflection time T_m can be considered the working period of c, and the reflection set $g(c, T_m)$ can be considered the information collected and presented by c. A continuous monitoring information system is generally required to maintain continuous and uninterrupted monitoring of the target object; i.e., its duration is often equal to the working period of the information collection equipment. Thus, we have $\text{duration}(I) = \sup T_h - \inf T_h = \sup T_m - \inf T_m$.

However, any piece of equipment may fail; therefore, it is necessary to specify the mean time between failures (MTBF) [39] of systems in engineering practice, which indicates the duration of the normal operation of the system without failure over the entire life cycle. In continuous monitoring systems, since the MTBF of information acquisition equipment c is the average of all the $\sup T_m - \inf T_m$ values in the full life cycle, the average of the continuous monitoring information $\text{duration}(I)$ is the same. The corollary is proved.

G. Sampling Rate of Information and Nyquist's Sampling Theorem

The sampling rate of information is also an important measure for some systems [40]. The sampling rate of restorable information $I = \langle o, T_h, f, c, T_m, g \rangle$ can be defined as the ratio of the number of interruptions of its occurrence to the length of the same occurrence; i.e., sampling rate$(I) = \overline{\Lambda}/|U|$, where $U = \cup_{\lambda\in\Lambda} U_\lambda$, $U_\lambda \subseteq [\inf T_h, \sup T_h]$ and $T_h \cap U_\lambda = \emptyset$, and Λ is an index set.

Corollary 8: (Minimum restorable sampling rate of periodic information). For restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$, if $f(o, T_h)$ is a set of values, and there exists a minimum value T such that, for $\forall x \epsilon o, t, t+T \epsilon T_h$, $f(x,t) = f(x, t+T)$. Moreover, let $\{U_\lambda\}_{\lambda\in\Lambda}$ be a family of pairwise disjoint connected sets such that for $\forall\lambda \in \Lambda$, all have $U_\lambda \subseteq [\inf T_h, \sup T_h]$, $T_h \cap U_\lambda = \emptyset$, and the Lebesgue measures of all U_λ are equal. Then, I is referred to as the periodic information, and the lowest restorable sampling rate of information I is equal to $1/(2T)$.

Proof: For periodic information $I = \langle o, T_h, f, c, T_m, g\rangle$, $\inf T_h \neq \sup T_h$. Here, $f(x,t) = f(x, t+T)$ for $\forall x\epsilon o, t, t+T\epsilon T_h$, and T is the smallest value satisfying this condition. Thus, for $\forall x \in o, f(x,t)$ has no frequency greater than $1/T$ relative to time t. According to Nyquist's sampling theorem [3], $f(x,t)$ is completely determined relative to time t by a series of values that are no farther than $T/2$ apart. In the definition of information I, $\{U_\lambda\}_{\lambda\in\Lambda}$ is a series of sampling intervals with equal measures, and the cardinal number $\overline{\Lambda}$ is the number of sampling intervals. Therefore, the Lebesgue measure $|U| = |U_\lambda|\overline{\Lambda}$ of $U = \cup_{\lambda\in\Lambda} U_\lambda$ holds for $\forall\lambda \in \Lambda$, and the sampling_rate of information I sampling_rate$(I) = \overline{\Lambda}/|U| = \overline{\Lambda}/|U_\lambda|\overline{\Lambda} = 1/|U_\lambda|$ holds for $\forall\lambda \in \Lambda$.

Note that $|U_\lambda| \le T/2$ holds for $\forall\lambda \in \Lambda$ if and only if sampling_rate$(I) \ge 1/(2T)$. Thus, for $\forall x \in o$, the values of $f(x,t)$ are completely determined, and I has a definite restored state. The corollary is proved.

H. Invariance Principle of the Aggregation Degree of Restorable Information

The degree of information aggregation is a widely used measure [41]. The aggregation of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ can be defined as the ratio of the number of relations between all elements in the state set $f(o, T_h)$ to the number in the set $f(o, T_h)$; i.e., aggregation$(I) = \overline{\Re}/\overline{f(o, T_h)}$, where $\Re$ is the set of relations between all elements in the state set $f(o, T_h)$.

Corollary 9: (Invariance of the aggregation degree of restorable information). For restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$, let the cardinality of the state set $f(o, T_h)$ be $\overline{f(o, T_h)} \neq 0$, and let $\Re$ be the set of all relations between the elements of $f(o, T_h)$. Then, for the reflection set $g(c, T_m)$, whose cardinality is $\overline{g(c, T_m)} = \overline{f(o, T_h)}$, there exists a set of relations $\mathfrak{Q}$ on it with cardinality $\overline{\mathfrak{Q}} = \overline{\Re}$. Thus, we have aggregation$(I) = \overline{\Re}/\overline{f(o, T_h)} = \overline{\mathfrak{Q}}/\overline{g(c, T_m)}$.

Proof: Here, $I = \langle o, T_h, f, c, T_m, g\rangle$ is restorable information; thus, there is a surjective mapping from the state set $f(o, T_h)$ to the reflection set $g(c, T_m)$. Therefore, the cardinalities of the two sets are necessarily equal; i.e., $\overline{g(c, T_m)} = \overline{f(o, T_h)}$. In addition, we can define the set of relations $\mathfrak{Q} = \{Q_R | R \in \Re\}$ on $g(c, T_m)$ such that, for any $R \in \Re$, if $o_\lambda, o_\mu \in o, T_{h\lambda}, T_{h\mu} \in T_h$, $f_\lambda, f_\mu \in f$ and $f_\lambda(o_\lambda, T_{h\lambda}) R f_\mu(o_\mu, T_{h\mu})$, then $I(f_\lambda(o_\lambda, T_{h\lambda}))$, $I(f_\mu(o_\mu, T_{h\mu})) \in g(c, T_m)$, and we define $f_\lambda(o_\lambda, T_{h\lambda}) Q_R f_\mu(o_\mu, T_{h\mu})$ such that Q_R is a relation on $g(c, T_m)$. The set $\mathfrak{Q} = \{Q_R | R \in \Re\}$ has a one-to-one surjective relationship with the set $\Re$; thus, the cardinalities are exactly equal, i.e., $\overline{\mathfrak{Q}} = \overline{\Re}$. From this result, we obtain aggregation$(I) = \overline{\Re}/\overline{f(o, T_h)} = \overline{\mathfrak{Q}}/\overline{g(c, T_m)}$. The corollary is proved.

I. Scope and Coverage of Information and Metcalfe's Law

User reach coverage is an important measure of some information systems [42], and it is also a typical consideration in information coverage. The coverage of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ can be defined as the integral of all the measures of the carriers of its copies and itself; i.e., coverage$_\sigma(I) = \int_\Lambda \sigma(c_\lambda) d\mu$, where $\{I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda\rangle\}_{\lambda\in\Lambda}$ is a set containing I and all its copies, Λ is an index set, μ is a measure on index set Λ, and σ is a measure on measurable set c.

Corollary 10: (The value of a network system is equal to the product of the maximum scope and the maximum coverage of the information it carries). For restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$, if its carrier c is a network system comprising finite nodes, its value is equal to the product of the maximum possible values of the scope and coverage of I.

Proof: Given that carrier c of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ is a network system comprising a finite number of nodes, if the number of nodes in c is n, according to Metcalfe's law [4], the value of the network system c is equal to the square of the number of nodes n^2. In addition, we can consider I as the information from all nodes in a network; thus, the noumenon o is the network system c, and its measure σ is the number of nodes. In this case, the maximum scope of information I is scope$_\sigma(I) = n$. The measure of carrier c is also the number of nodes, and the maximum value of the coverage of information I is coverage$_\sigma(I) = n$. Therefore, the value of this network system is equal to the product of the maximum possible values of the scope and coverage of information I. The corollary is proved.

J. Distortion of Information and the Kalman Filtering Principle

The distortion of information is an important measure of error and plays an important role in the development of information systems [43]. The distortion of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ can be defined as the distance between its reflected state and its restored state in distance space; i.e., distortion$_J(I) = d(f, \tilde{f})$, where $\tilde{f}(\tilde{o}, \tilde{T}_h)$ is its reflected state.

Corollary 11: (Minimum distortion estimation method for discrete linear stochastic systems). Let $I = \langle o, T_h, f, c, T_m, g\rangle$ be the state information of a discrete linear stochastic system whose motion and measure are both affected by Gaussian white noise. Then, the minimum distortion estimation of I can be obtained based on the reflection J of the Kalman filter [5].

Proof: For the state information $I = \langle o, T_h, f, c, T_m, g\rangle$ of a discrete linear stochastic system in which the motion and measure are both affected by Gaussian white noise, the noumenon o is the system itself; the set of occurrence times T_h is a series of time sequences with equal intervals, which can be written as $1, 2, , , k$; the state set $f(o, T_h)$ can be written as $x(k)$, $k = 1, 2, \cdots$; and $x(k) = Ax(k-1) + BU(k) + W(k)$ holds. Here, $x(k)$ is the system state at time k, and $U(k)$ is the system input at time k. A and B are system parameters, which are matrices for a multimodel system; $W(k)$ is the motion noise of the system; and Q is its covariance.

In this case, carrier c is a measuring system, and the reflection time set T_m is the same as the occurrence time set T_h (denoted $1, 2, \cdots k, \cdots$). The reflection set $g(c, T_m)$ is a series of measured values on T_m, which is denoted by $z(k), k = 1, 2, \ldots$, and $z(k) = Hx(k) + V(k)$ holds, where $z(k)$ is the measured value at time k. In addition, H is the measuring system parameter, and for a multiple-measurement system, H is a matrix, $V(k)$ is the measuring noise at time k, and R is its covariance. Let J consist of the following five formulas:

$$x(k|k-1) = Ax(k-1|k-1) + BU(k),$$

where $x(k|k-1)$ is the result predicted by the previous state, $x(k-1|k-1)$ is the optimal result of the previous state, and $U(k)$ is the current system input state;

$$P(k|k-1) = AP(k|k-1)A^T + Q,$$

where $P(k|k-1)$ is the covariance corresponding to $x(k|k-1)$, $P(k-1|k-1)$ is the covariance corresponding to $x(k-1|k-1)$, A^T is the transpose of matrix A, and Q is the covariance of the motion of the system;

$$x(k|k) = x(k|k-1) + G(k)(z(k) - Hx(k|k-1))$$

where $G(k)$ is the Kalman gain; and

$$G(k) = P(k|k-1)H^T/(HP(k|k-1)H^T + R)$$

$$P(k|k) = (\mathrm{I} - G(k)H)\mathrm{P}(k|k-1)$$

where I is the identity matrix. Then, it is obvious that J can be inferred recursively to be the surjective mapping from $z(k)$ to $x(k)$, i.e., from $g(c, T_m)$ to $f(o, T_h)$. Thus, J is a reflection of I. Consequently, according to the Kalman filtering principle [5], $x(k|k)$ is the optimal estimation of $x(k)$; i.e., the minimum distortion estimation of I can be obtained based on the reflection J of the Kalman filter. The corollary is proved.

K. Mismatch of Information and Average Lookup Length of a Search Algorithm

The mismatch of information is a key measure of many systems, especially artificial intelligence (AI) systems [44]. The mismatch of restorable information $I = \langle o, T_h, f, c, T_m, g\rangle$ can be defined as the distance between itself and the target information in distance space; i.e., $mismatch_{I_0}(I) = d(I, I_0)$, where I_0 is the target information.

Corollary 12: (Average search length of minimum mismatch information for search algorithms). Let the target information $I_0 = \langle o_0, T_{h0}, f_0, c_0, T_{m0}, g_0\rangle$ and set $\{I_i = \langle o_i, T_{hi}, f_i, c_i, T_{mi}, g_i\rangle | i = 1, , , n\}$ be restorable information. Here, o_0 and o_i, T_{h0} and T_{hi}, f_0 and f_i, c_0 and c_i, T_{m0} and T_{mi}, and g_0 and g_i are elements of sets $\mathcal{P}_o, \mathcal{P}_{T_h}, \mathcal{P}_f, \mathcal{P}_c, \mathcal{P}_{T_m},$ *and* $\mathcal{P}_g$, respectively, and I_0 and $I_i(i = 1, , , n)$ are elements of the distance space $\langle(\mathcal{P}_o, \mathcal{P}_{T_h}, \mathcal{P}_f, \mathcal{P}_c, \mathcal{P}_{T_m}, \mathcal{P}_g), d\rangle$. Let $1 \leq m \leq n$ so that the following holds.

$$mismatch_{I_0}(I_m) < mismatch_{I_0}(I_i), \quad 1 \leq i \leq n \text{ and } i \neq m$$

Then, the ASL from $\{I_i | i = 1, , , n\}$ to I_m is related to both the mismatch $mismatch_{I_0}(I_m)$ and the search algorithm.

Proof: With the increasingly abundant information content available on the internet and the increasing number of information query tools in use, a large number of complex information queries have arisen. Note that users have requirements for the noumena, occurrence time, state of information, carrier, reflection time, and mode of information. However, it is difficult to clearly describe all these requirements. Thus, target information that completely matches the users' requirements is difficult to find. Advanced retrieval and intelligent recommendation systems often analyze and estimate the target information I_0 that satisfies a user's needs for a given application scenario, and such systems search and calculate information I_m with the minimum degree of mismatch $mismatch_{I_0}(I_i)(i = 1, , , n)$ from the limited information set $\{I_i | i = 1, , , n\}$, which is ultimately pushed to the end users.

According to the ASL principle [45], the definition of ASL is given as $ASL = \sum_{i=1}^{n} p_i c_i$, where p_i is the probability of finding information I_i. Generally, we assume that the probability of finding each piece of information is the same, i.e., $p_i = 1/n$. Here, c_i is the number of comparisons required to find the information I_i. Then, the following two situations must be considered.

The first situation is $mismatch_{I_0}(I_m) = 0$. Here, if the sequential search method is used, the mismatch degree $mismatch_{I_0}(I_i)$ is calculated incrementally from information I_1 until I_m is found. Thus, we obtain $ASL = \sum_{i=1}^{n} p_i c_i = (1/n) \sum_{i=1}^{n} i = (n+1)/2$.

If the bisection search method is adopted, the middle serial number information is always used as the root to divide the left and right subtrees, and each subtree is used as the root of the middle serial number information; division proceeds in a progressive manner until the subtrees cannot be divided any further. For each subtree, information I_i is searched from the root, and the mismatch $mismatch_{I_0}(I_i)$ is calculated until I_m is found. Thus, we obtain $ASL = \sum_{i=1}^{n} p_i c_i = (1/n) \sum_{i=1}^{n} 2^{i-1} i = ((n+1)/n)$ $log_2(n+1) - 1$. Here, $h = log_2(n+1)$ is the height of the n information discrimination trees.

The second situation is $mismatch_{I_0}(I_m) \neq 0$. In this case, the ASL is always n because we must compare the mismatch degree $mismatch_{I_0}(I_i)(i = 1, , , n)$ of all the information and select the minimum mismatch degree to obtain I_m. When n is very large, the number of search computations is also very large due to the ASL. Thus, an appropriate threshold can be set, and the search is completed when the mismatch degree $mismatch_{I_0}(I_i)$ is less than or equal to the threshold. The corollary is proved.

V. Application Experiments for the General Information Measures

The preceding section demonstrates that the general information measures based on the sextuple model can integrate many classical and widely used principles in information science into

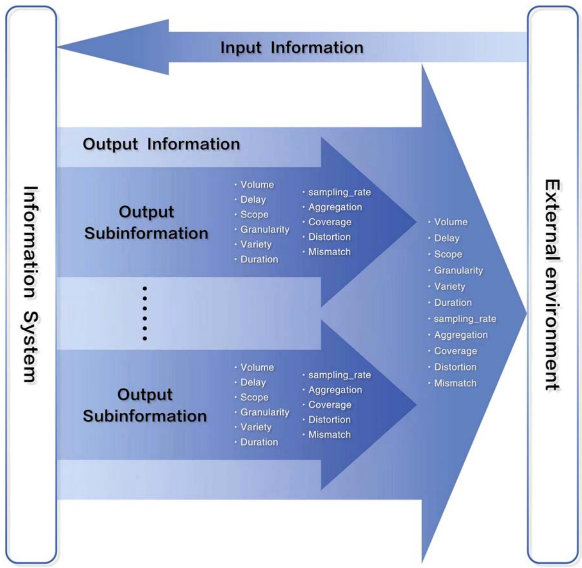

Fig. 3. Illustration of the general evaluation indices for the output of information systems.

a unified theoretical framework. This system provides a solid foundation for the research and analysis of large-scale complex information systems from a more comprehensive and systematic perspective by leveraging numerous existing scientific achievements. Indeed, any information system can be regarded as an input–output system that interacts with the external environment, as illustrated in Fig. 3. The information system receives input information from the external environment and simultaneously outputs information to it. The performance of the system is entirely dependent on the indicators of the output information. Researchers often study and analyze various forms of information systems based on different indicators, resulting in a general lack of comparability between systems, even for different stages of development within the same system. The eleven general measures based on the sextuple model can effectively address this important issue. In Fig. 3, all the output information from the information system can be measured by the eleven types of measures. All the measures can be used to evaluate the complete output information, and certain measures of some subinformation can be assessed based on diverse application requirements. Subsequently, we can apply corresponding scientific principles and methods to enhance system performance according to the evaluation results. This highlights the essential role of general measures in satisfying the research and analysis needs of complex information systems while achieving the complete unification of universality and flexibility.

Using scientific language, we conducted a test and analysis of the output performance of three well-known question-answering and retrieval systems, ChatGPT, Baidu, and Bing, on world population information. We used annual population data from March 2023 obtained from the worldometers.info website as the benchmark. A total of 194 countries and regions were sampled for a period of 13 years (1995, 2000, 2005, 2010, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, and 2023). For ChatGPT, we directly assessed the content of its answer, while for Baidu and Bing, we evaluated their push data results or the information from the first web page. Here, *Delay* refers to the delay of the latest population information contained in the tested system relative to 2023, *Scope* refers to the ratio of the number of countries and regions covered by the tested system with population information to the total number of countries and regions in the world, *Granularity* refers to the smallest regional level at which the population information contained in the tested system can be refined, *Variety* refers to the number of population information types contained in the tested system, *Duration* measures the coverage time of the population information contained in the system under test, *Sampling_rate* measures the update frequency of the population information of the tested system, *Aggregation* measures the correlation degree of the output information of the tested system that can be tested, *Distortion* measures the deviation between the output data of the tested system and the reference data, and *Mismatch* measures the deviation between the output information of the system under test and the expectation of the tester. The *Volume* and *Coverage*, which are difficult to obtain from the user side, were not used as indicators for this test. The results of this test show that the information measures can be used to evaluate and guide the development of information systems effectively.

As the baseline data of worldometers.info cover the period from 1950 to 2020, the data for the tested system beyond 2020 were excluded from the calculation of the Distortion and Mismatch measures. Additionally, due to historical events such as wars and the division and merging of countries, the statistical standards for some national population data are inconsistent, resulting in substantial differences in the data. To ensure accuracy, such national data, including from Serbia, the Congo, Eritrea, and Micronesia, among others, were excluded from the calculation of the Distortion and Mismatch measures. A total of 188 countries and regions were included in the comprehensive measure assessment.

Table II shows that both ChatGPT and Baidu could output the latest population information in 2023, resulting in a Delay of 0. In contrast, Bing's latest information output only reaches 2022, giving it a Delay of one year. ChatGPT and Baidu can export information from all countries and regions, while Bing can only export information from less than 0.5% of countries and regions. Additionally, ChatGPT and Baidu's information can be refined to the street level, while Bing's information can be refined only to the city level. ChatGPT can output the most comprehensive set of 9 types of information, whereas Baidu and Bing can only output 8 types of information. Both ChatGPT and Baidu can output 123 years of information since 1900, whereas Bing can only output 66 years of information from 1956 to 2022. The Sampling_rate for all three systems is once a year. The output of ChatGPT has a relevance score of 2/9, while the outputs of Baidu and Bing have no relevance. Baidu's output information has the lowest Distortion, followed by Bing, and ChatGPT has the highest Distortion. Conversely, ChatGPT's output information has the lowest Mismatch, while Baidu and Bing have much higher Mismatch values. The overall test results indicate that although ChatGPT has several indicators comparable to those of Baidu and Bing for world population

TABLE II
RESULTS OF THE TESTS OF POPULATION INFORMATION OUTPUT BY THE THREE SYSTEMS

Measure	Indicator Content	The test results of ChatGPT	The test results of Baidu	The test results of Bing
Volume	The total volume of its world population information	——	——	——
Delay	The delay of its latest population information relative to 2023	0 years	0 years	1 year
Scope	The ratio of its number of countries with population information to the total number of countries in the world	100%	100%	99.468%
Granularity	The smallest regional level of its population information that can be refined	Street	Street	City
Variety	The number of its population information types	9	8	8
Duration	The coverage time of the population information contained in the system under test	123 years	123 years	66 years
Sampling rate	The update frequency of its population information	1 time/year	1 time/year	1 time/year
Aggregation	The correlation degree of its output information that can be tested	2/9	0	0
Coverage	The number of users of its population information	——	——	——
Distortion	The deviation between its output data and the reference data	0.04008	0.02264	0.02993
Mismatch	The deviation between its output information and the expectation of the tester	0.00266	0.01223	0.21064

information, its Variety and Aggregation measures are better than those of the latter two systems. Despite the fact that ChatGPT's output information has the largest error, it meets users' needs significantly better than the other two information systems. This is why ChatGPT is widely regarded as having great potential.

VI. CONCLUSION

We can now answer the three key questions raised at the beginning:

First, four concise postulates about information are established in this paper according to the general perception of information by the public, and a unified model and the mathematical definition of information are derived using the axiomatic method. On this basis, general information measures, which include eleven types of measures, can be defined.

Second, the formula for the relationship between information and mass, energy, and time in closed systems is derived by utilizing the volume, which is the most commonly used information measure. The information volume that the universe may have carried thus far and the ratio of information volume to power consumption for several typical computing systems are analyzed using this formula. The results indicate that, from the perspective of carrying information, existing information systems have significant potential to conserve energy.

Third, through the eleven corollaries, it is demonstrated that each type of measure can express a classical or commonly used information science and technology principle through the proposed general information measures. This implies that a unified theoretical system of information science can be constructed by integrating many classical information principles.

By utilizing the proposed general information measures, we conduct a comprehensive evaluation of the output information from the three widely recognized information systems. The performance features and limitations of each system are analyzed, and the directions for optimization of each system are identified. These results align with the conventional understanding of these systems, indicating that the general information measures presented in this study can aid in the design and assessment of diverse information systems, thereby meeting the pressing demands of large-scale information system engineering.

Further research could explore the properties and relationships among the components of the sextuple model with respect to other attributes and measures of information. Additionally, by comparing and testing its compatibility with other existing principles and algorithms of information science, we can facilitate the construction of a more complete and applicable theoretical system of information science and more comprehensive information system dynamics (ISD). This would provide a more robust theoretical foundation for the development and application of information technology.

DECLARATION OF COMPETING INTERESTS

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGMENT

Dr. Yingfei Wang from the Information Technology Service Center of the Supreme People's Court, senior engineer Wu Yue from the China Judicial Big Data Research Institute, Professors Zhenyu Liu and Yashi Wang from China University of Political Science and Law and Professors Shuliang Wang and Tao Zheng and Mr. Yingxu Dang from the Beijing Institute of Technology have offered indispensable assistance to the author with regard to the research experiments and literature collection. Sincere appreciation is expressed for all their assistance.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request, and all the other data generated or analyzed during the study are included in the present manuscript.

REFERENCES

[1] C. E. Shannon, "The mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[2] M. I. Skolnik, *Radar Handbook*, 2nd ed. New York, NY, USA: McGraw-Hill, 1990.

[3] H. Nyquist, "Certain topics in telegraph transmission theory," *Proc. IEEE*, vol. 90, no. 2, pp. 280–305, Feb. 2002.

[4] C. Shapiro and H. R. Varian, *Information Rules: A Strategic Guide to the Network Economy.* Boston, MA, USA: Harvard Business School Press, 1998.
[5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
[6] M. Burgin, *Theory of Information: Fundamentality, Diversity and Unification.* Singapore: World Scientific, 2010.
[7] M. Zwolak, "Amplification, inference, and the manifestation of objective classical information," *Entropy*, vol. 24, no. 6, 2022, Art. no. 781, doi: 10.3390/e24060781.
[8] N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine.* New York, NY, USA: MIT Press, 1961.
[9] Program on Information Resources Policy, "*Program projects, annual report volume 2 (1975–1976)*," Computation Laboratory, Harvard University, Cambridge, MA, USA, Rep. R-76-2, 1976.
[10] J. F. Xu et al., "Research on the model and measurement of objective information," *Sci. China Inf. Sci.*, vol. 45, no. 3, pp. 336–353, 2015.
[11] J. F. Xu, S. L. Wang, Z. Y. Liu, and Y. F. Wang, "Objective information theory exemplified in air traffic control system," *Chin. J. Electron.*, vol. 30, no. 4, pp. 743–751, 2021.
[12] J. F. Xu et al., "Foundations and applications of information systems dynamics," *Engineering*, vol. 27, no. 8, pp. 254–265, 2023, doi: 10.1016/j.eng.2022.04.018.
[13] J. F. Xu, F. H. Sun, and Q. W. Chen, *Introduction to the Smart Court System-of-Systems Engineering Project of China.* Singapore: Springer-Verlag, 2022.
[14] R. Landauer, "The physical nature of information," *Phys. Lett. A*, vol. 217, nos. 4–5, pp. 188–193, 1996.
[15] M. Zhou, *Real Analysis*, 2nd ed. Beijing, China: Peking University Press, 2008.
[16] N. Wiener, *The Human Use of Human Beings: Cybernetics and Society*, New ed. Cambridge, MA, USA: Da Capo Press, 1988.
[17] M. Taube, *Evolution of Matter and Energy on a Cosmic and Planetary Scale.* New York, NY, USA: Springer-Verlag, 1985.
[18] P. Zizzi, "Holography, quantum geometry, and quantum information theory," *Entropy*, vol. 2, no. 1, pp. 36–39, 2000.
[19] S. Lloyd, "Ultimate physical limits to computation," *Nature*, vol. 406, no. 6799, pp. 1047–1054, 2000.
[20] N. Margolus and L. B. Levitin, "The maximum speed of dynamical evolution," *Physica D*, vol. 120, nos. 1–2, pp. 188–195, 1998.
[21] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* New York, NY, USA: Cambridge Univ. Press, 2010.
[22] E. Hubble, "A relation between distance and radial velocity among extra-galactic nebulae," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 15, no. 3, pp. 168–173, 1929.
[23] N. A. Bahcall, J. P. Ostriker, S. Perlmutter, and P. J. Steinhardt, "The cosmic triangle: Revealing the state of the universe," *Science*, vol. 284, no. 5419, pp. 1481–1488, 1999.
[24] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM J. Res. Dev.*, vol. 5, no. 3, pp. 183–191, 1961.
[25] M. M. Vopson, "The mass-energy-information equivalence principle," *AIP Adv.*, vol. 9, 2019, Art. no. 095206.
[26] A. Einstein, "On the electrodynamics of moving bodies," *Ann. Phys.*, vol. 17, no. 10, pp. 891–921, 1905.
[27] F. Rohrlich, "An elementary derivation of $E = mc^2$," *Am. J. Phys.*, vol. 58, no. 4, pp. 348–349, 1990.
[28] S. Lloyd, "Computational capacity of the universe," *Phys. Rev. Lett.*, vol. 88, no. 23, 2002, Art. no. 237901.
[29] J. R. Gott III et al., "A map of the universe," *Astrophys. J.*, vol. 624, no. 2, pp. 463–484, 2005.
[30] B. Bryson, *A Short History of Nearly Everything.* New York, NY, USA: Broadway Books, 2003.
[31] L. E. Reichl, *A Modern Course in Statistical Physics.* Austin, TX, USA: Univ. of Texas Press, 1980.
[32] D. Kang, D. Kang, and S. Ha, "Multi-bank on-chip memory management techniques for CNN accelerators," *IEEE Trans. Comput.*, vol. 71, no. 5, pp. 1181–1193, May 2022, doi: 10.1109/TC.2021.3076987.
[33] Y. Li et al., "Reinforcement learning-based resource partitioning for improving responsiveness in cloud gaming," *IEEE Trans. Comput.*, vol. 71, no. 5, pp. 1049–1062, May 2022, doi: 10.1109/TC.2021.3070879.
[34] A. Bahmani et al., "A scalable, secure, and interoperable platform for deep data-driven health management," *Nat. Commun.*, vol. 12, 2021, Art. no. 5757, doi: 10.1038/s41467-021-26040-1.
[35] H.-M. Chen, S.-Y. Lee, T. Mudge, C.-J. Wu, and C. Chakrabarti, "Configurable-ECC: Architecting a flexible ECC scheme to support different sized accesses in high bandwidth memory systems," *IEEE Trans. Comput.*, vol. 68, no. 5, pp. 646–659, May 2019, doi: 10.1109/TC.2018.2886884.
[36] L. Rayleigh, "LVI investigations in optics, with special reference to the spectroscope," *London Edinburgh Dublin Philos. Mag. J. Sci. Ser. 5*, vol. 8, no. 51, pp. 477–486, 1879.
[37] N. K. Jha and S. Mittal, "Modeling data reuse in deep neural networks by taking data-types into cognizance," *IEEE Trans. Comput.*, vol. 70, no. 9, pp. 1526–1538, Sep. 2021, doi: 10.1109/TC.2020.3015531.
[38] D. K.-H. Yu and J.-W. Hsieh, "A management scheme of multi-level retention-time queues for improving the endurance of flash-memory storage devices," *IEEE Trans. Comput.*, vol. 69, no. 4, pp. 549–562, Apr. 2020, doi: 10.1109/TC.2019.2954398.
[39] J. Lienig and H. Bruemmer, "Reliability analysis," in *Fundamentals of Electronic Systems Design.* Cham, Switzerland: Springer-Verlag, 2017, pp. 45–73.
[40] L. Kong, S. Li, and R. Liu, "High-performance constant-time discrete Gaussian sampling," *IEEE Trans. Comput.*, vol. 70, no. 7, pp. 1019–1033, Jul. 2021, doi: 10.1109/TC.2020.3001170.
[41] J. Zhang et al., "Adaptive federated learning on non-IID data with resource constraint," *IEEE Trans. Comput.*, vol. 71, no. 7, pp. 1655–1667, Jul. 2022, doi: 10.1109/TC.2021.3099723.
[42] Y. Bai, L. Chen, S. Ren, and J. Xu, "Automated customization of on-device inference for quality-of-experience enhancement," *IEEE Trans. Comput.*, vol. 72, no. 5, pp. 1329–1342, May 2023, doi: 10.1109/TC.2022.3208207.
[43] B. A. Bilgin and P. Stanley-Marbell, "Probabilistic value-deviation-bounded source-dependent bit-level channel adaptation for approximate communication," *IEEE Trans. Comput.*, vol. 70, no. 11, pp. 1949–1961, Nov. 2021, doi: 10.1109/TC.2020.3031494.
[44] M.-C. Yang, C.-F. Wu, S.-H. Chen, Y.-L. Lin, C.-W. Chang, and Y.-H. Chang, "On minimizing internal data migrations of flash devices via lifetime-retention harmonization," *IEEE Trans. Comput.*, vol. 70, no. 3, pp. 428–439, Mar. 2021, doi: 10.1109/TC.2020.2989554.
[45] I. Flores and G. Madpis, "Average binary search length for dense ordered lists," *Commun. ACM*, vol. 14, no. 9, pp. 602–603, 1971.

Jianfeng Xu received the Ph.D. degree in computer science from Nanjing University, in 2001. His research interests mainly include the modeling and measurement of information and its application in computer and information system engineering.