ISD2K – Reference Model to Represent an Information System Oriented to Knowledge Production

M. Fidos Jr. and M. Ferreira

Abstract — **The generation of knowledge (K) is a continuous and necessary effort for human evolution. Electronic records have enabled data processing, storage, and transfer to extend and accelerate this evolutionary effort. Electronic systems have been created to automate and record data from a process, for governments, organizations, and even for individuals. The expansion of these systems and the gradual increase of their complexity demanded the creation of Information Systems (IS) that aggregate data from several other systems to perform analysis and reports to support decision making, i.e., knowledge production (KP). The concept of Big Data - large volumes of data, with diverse structures, produced at an uncommon speed - describes the exponential growth of data production, but the same did not happen with the production of K. The objective of this work was to contribute to reduce this gap through the presentation of a reference model of IS to represent comprehensive and qualitatively measurable architectures in the KP. The method proposed in this work was the literature review about K, IS and reference models for information systems architecture (ISA), followed by the proposition of an IS reference model oriented to the conversion of data into knowledge (ISD2K) through a static and dynamic conceptual model. This proposition was validated from the description of a use case using ISD2K and elaboration of adherence questionnaires and evaluation of two IS in use at INPE (Instituto Nacional de Pesquisas Espaciais). The model is based on a common usage language, describing six views or perspectives and nine representations to each view that allow to deliver an integrated representation from different points of view inside an organization of an IS where is possible to distinct data, information and knowledge and plan and qualify the KP data oriented. The aspects of quality, utilization and impact were highlighted in this model as an innovative contribution to an IS model.**

Index Terms—**Knowledge based systems, Information Systems, Systems Architecture.**

I. INTRODUCTION

Knowledge (K) is a continuous effort that characterizes human evolution. Its first efforts started in the creation of K by the human mind and in the oral dissemination. This effort evolved using external equipment that amplified the human capacity for K production (KP) and currently data in electronic format represents a major advance in the recording and processing of this data through electronic means of hardware and software.

KP can be characterized as an input-transformation-output process done by humans on a conscious or unconscious level, using data stored in their own resources (body, mind) and/or using external devices, and the transformation of these inputs is done through data interpretation and interaction, and its results to be considered K are dependent on the perception and goals of the user.

The production and recording of digital data have increased exponentially in recent decades. Charles Tilly [1], a social historian, first coined the term Big Data in an article to discuss the emergence of large volumes of data that could now confront through quantitative methods propositions and theories based solely on individual perceptions. Laney [2] characterized Big Data from the impacts generated by electronic commerce in the creation, dissemination, and use of digital data. Today the discussions observed by Tilly and Laney are perceived in dozens of other fields of science, in government management, and in companies seeking solutions to redefine the use of data based on the concept of Big Data.

Information systems (IS) are systems oriented toward the provision of electronic data. Electronic data are distinguished into three levels - data, information and/or knowledge - according to the usefulness, or impact provided to its user, which potentially allows IS to be KP systems, but the condition to distinguish data, information or knowledge has not been found in current information systems, nor success factors in the use of these systems that allow to assess how efficiently these systems potentially convert data into K. System designers, researchers, and customers use IS architectures (ISA) models to develop, understand, and acquire, respectively, information systems.

Problem: If electronic data can be considered as inputs to knowledge generation, and its production has increased exponentially, an increase, albeit not proportional, in KP would be expected. Why then is Big Knowledge not present in today's discussions? The search for this answer can help organizations like INPE to meet the demands of its strategic planning for the production and dissemination of K [3], in the evolution of systems engineering processes, and the design of a comprehensive and measurable ISA in KP and, mainly, allow the improvement of existing IS´s and data to produce more knowledge/ better decisions.

The main objective of this work is the proposition of an IS reference model to represent comprehensive and qualitatively measurable IS oriented to KP (ISD2K).

The main contributions of this research are: (1)A reference model, which aligns distinct views within an organization for KP, using common language, formalized by a static (6 views and 9 representations) and a dynamic (functioning) conceptual

M. Fidos Jr. is a master's student in the graduate program in Space Systems Engineering at INPE. e-mail: mfidosjr@gmail.com
M. Ferreira is a scientific advisor to FAPESP in the area of Software Engineering and works as a coordinator and researcher at the Center for Tracking and Control of Satellites at INPE. e-mail: mauricio.ferreira@inpe.br

model, which can be employed for new developments or evolution of existing IS; (2) The consolidation of pre-existing building blocks - K, IS, ISA, D-I-K as an original contribution; (3) Supporting INPE in the evaluation and evolution of two existing IS to become more KP oriented; (4) the inclusion of aspects such as quality, utilization and impact of use in a more knowledge-oriented IS development model as another original contribution.

This paper is organized into seven chapters: 1. Introduction, 2. Concepts and literature review, Comparative study between works on ISA, 3. Proposed framework (ISD2K), 4. Use case, 5. Creation and application of questionnaires, 6. Results and discussion, and 7. Conclusion.

II. Concepts and Literatures Review

A. Knowledge - Creating and communicating K is a continuous effort that characterizes human evolution, whether this process is carried out consciously or unconsciously. In its origin there was no stage of permanent recording of K.

The first major evolution in the transmission of K after oral communication was the persistent registration, initially on cave walls, which has evolved to the present day with the use of electronic media. The act of accumulating and transmitting K through images and symbols began from the moment man reached cognitive and manual conditions for this, with its first evidence around 30,000 B.C. [4].

Currently the concepts involving K present different interpretations, with similarities and distinctions. Several authors distinguish K in an extrinsic and intrinsic approach, considering, with some similarity, that the extrinsic means are recorded externally to the human mind, and the intrinsic ones, internally.

Nonaka et al. [5] interpreted K by dividing the concept into two - explicit K (organized, structured, materialized form) and implicit K (intangible, internal human, unrecorded). Other authors do not refer directly to this segmentation but deal with the concept considering it internal [6] or external [7]. In an intrinsic way, we can highlight the evolutionary processes transmitted via DNA, which has an efficient process of recording, updating and dissemination [8].

The central aspect of all the above descriptions is the human beings involved in the KP, at different levels of consciousness, but even happening internally, K can also be recorded externally.

We can describe KP as: *"A process developed by human beings on a conscious or unconscious level, internally or through external interaction, organized in an autonomous, self-organized or collectively organized way, using their own resources (body, mind) and/or external devices, used to store and process the produced K, materialized internally (DNA, thought) or externally in a collectively interpreted way (based on language, social rules, procedure) or in a way that only makes sense to the one who creates it".*

Considering a hypothetical exercise to scale the K available in the world would require the sum of various materializations of K (DNA, Thoughts, Records, language-based records - digital or physical, explicit, and implicit social rules, products, procedures, literature - print and digital...). The dimension of time could be included in this exercise to increase, exponentially, the amount of potential D and K already developed. From this perspective, the concept of Big Data as cited by Tilly and Laney is only a small part of the potential data and knowledge produced or potentially produced in the world.

B. Concept of Measurable KP - For delimitation in this paper, it will be considered outputs of the KP process that are measurable, or those that use external devices. Based on this delimitation in this paper the following definition of measurable KP will be used:

"A process that occurs at a conscious level, individually or collectively elaborated, individually or collectively organized, using internal (mind, minds) AND external devices as production tools, being mandatorily recorded/stored in external equipment, e.g., registers, language-based records, recorded social rules, products and procedures."

Nonaka [9] presents K Management (KM) from the perspective of an organizational creation process with a cyclical approach between the internal and conscious (tacit) level and an external and conscious (explicit) level happening in individuals and that the organization helps to enhance this process by structuring and extending KP. Although the author presents an important reference on the production of K, it is not possible to identify a measurable procedure for converting D into K.

Spender [10] developed studies that resulted in a theory of the knowledge-based firm, placing the firm as a dynamic, evolving, and quasi-autonomous system of KP and application. His proposal considers the production function and resource base theories, and mentions that in addition to these theories, K should be considered, presenting the IT process as an example - its real contribution cannot be measured only by capital investment criteria, placing the transformation actors of the production-function, such as managers/strategists, as nodes of imaginative leadership and influence in the complex of emotionally and politically charged heterogeneous K systems that make up our socially constructed reality. In this approach, K can be interpreted as an input and output in a production function theory, and people can be considered as agents of transformation in this process.

Sharma [11] presents the origins of the hierarchical model between D, I, K, and wisdom (W) - DIKW - which compares the perceived value between each item, considering W as the most valuable item, and cites two domains where DIKW was initially discussed, KM and Information Science. In the KM domain authors such as Ackoff [12] provided the first references. Cleveland [13] was noted as the first author to explain the pyramid in detail. But the author also points out that it was the poet T. S. Eliot who first mentioned W, K, and I in a hierarchical value approach. Each stage of DIKW has a distinct value considering its hierarchical model. According to Ackoff [14] D are symbols that represent the properties of objects and events, and I consist of D processed to increase its usefulness, its function, and contains answers to (direct) questions such as who, what, when, where, and how many. K can be recorded as explanations, answers to how-to questions. W involves the exercise of judgment, that is, the interpretation of D, I, and K according to values, principles, and cannot be generated by computerized systems, only by individuals. For the purposes of this study, W will also be excluded, as a level that depends heavily on internal factors (personal values) that cannot be easily recorded on external devices.

The conversion of D2K, and consequently, KP, depends on a transformation process that changes the perceived value of D, I and/or K, for a group or individual, which may or may not be associated with a specific goal (answering direct questions or "how-to" questions), and the result of this process may be D, I and/or K.

C. Distinction between D, I and K - D is the only direct measurable input in KP. Relationships in D, or the association between two or more raw data, give meaning and change their initial status from D to I. When a user relates two raw data electronically and logs this successful attempt, it is possible to measure the amount of raw data with status I.

The same applies to K, but the distinction here is made by the level of usefulness provided to a specific user.

D. Information Systems - From the delineation proposed in the last section - a measurable KP process in which inputs and outputs can be represented by digital data (words, numbers, sounds, and images) - it is possible to relate this process to IS.

IS at the enterprise level play a dominant role in today's industrial automation [15]. ISO/IEC/IEEE 15288 [16] defines a system as *"a combination of interactive elements organized to achieve one or more stated purposes"* and INCOSE [17] as *"An integrated set of elements, subsystems, or assemblies that accomplish a defined purpose. These elements include products (hardware, software, firmware), processes, people, information, techniques, facilities, services, and other supporting elements."*

The definition of IS has evolved according to new technologies and/or use/purpose, according to several authors. Boell et al. [18] conducted a literature review using 34 IS definitions, presenting that IS involves information technologies (IT) - computers, software, databases, communication systems, internet, mobile device... - to perform specific tasks, interact and inform various actors - individuals, groups, or organizations - in different organizational or social contexts, to meet the actors' information needs and requirements about specific goals and practices.

Four thematic views were used to classify these 34 definitions - *technology* (data processing, storage, and transformation using computer hardware and software), *social* (IS as a social system, human and social systems as main agents), *sociotechnical* (interrelationship between social and technological components), and *process* (IS activities performing and supporting and processes such as capturing, transmitting, storing, retrieving, manipulating, and displaying information).

The IS meets the main system definitions and can be considered an expert and information-oriented system, and its process can be presented as an input-transformation-output flow. D is the key element for representation for input and output of an IS, considering that the representation of D, I and K can be done by means of digital data.

E. IS Success Factors - Communication systems theory was used to support understanding of IS success factors, considering that IS provides information as output and communication systems provide messages [19] and six levels of IS success were proposed: system, information, service, intention to use (usage), user satisfaction, and net benefits.

Nguyen et al. [20] proposed 9 elements of success to associate 45 related items for IS success: system quality, information quality, service quality, usage intention, usage, user satisfaction, individual impact, organizational impact and net benefits.

F. IS Architectures Comparison - A methodology to describe an architecture is useful for research, development, and system acquisition, because architectures document *"the structure of the components, their relationships, and the principles and guidelines that govern their design and evolution over time."* [21].

The definition of system architecture presented by ISO [22] has a similar approach: *"the fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and the principles of its design and evolution."* This definition is broad in the sense of the life cycle stages (Concept/ Development/ Production/ Use/ Support/ Deactivation - [23]).

In contrast, a more flexible answer to what is an ISA was provided by Zachman [24]: "there is no such thing as an information system architecture, but a set of them!" - a set of representations, from different points of view, representing the same architecture and are additive and complementary. Later, Sowa and Zachman [25] presented an extension of this framework (SZF).

This paper considered 47 papers on IS - 34 about their definitions and goals [18] and 13 directly associated with ISA and deployment. As a comparison the following aspects were evaluated: Approach (A), description types (DE), key representative elements (KRE), and contributions.

TABLE I
COMPARISON OF RELATED ISA

Author	A	DE	KRE	Contribution to ISA
1.Zachman [24]	D	3	VR	Initial benchmark
2.Sowa e Zachman [25]	D	6	VR	Evolution of Types
3.Evernden [26]	D	6	PC (IBM)	IS ready to use
4.Jamuna e Ashok [27]	M	-	OM	IS interoperability
5.Covington et al [28]	D	6	PC (Oracle)	IS ready to use
6.Wout et. Al [29]	D	6	PC (Capgemini)	IS ready to use

7.Zachman [30]	D	6	VR	Ontologies
8. Yoo et al. [31]	M	-	OM	Modularity
9. Cugola e Margara [32]	M	-	OM	Complex event processing
10. Majd et al [33]	D	-	OM	Service Oriented Architecture
11. Christoph et. Al [34]	D	6	PC (SAP)	IS ready to use
12. T.O. Group [35]	DP	6	OM	Evolutionary deployment method
13. Chang et al. [36]	DP	6	OM	Employment of big data technologies

Table. I. A – descriptive (D), method (M), deployment (DP); DE – 3 (What, How, Where), 6 (3+ Who, When, Why); KRE – (VR) Views/Representations, (PC) private components, (OM) own methods.

(A) is descriptive in most cases (8) and was also used in ISD2K, the other options (M and DP) are more practical implementations discussions. Six (DE) are the common sense (8), and the others do not define or has smaller numbers of types of view description. In this work, six views were used supported by 9 representations in each view for a broader description. Although only 3 related ISAs use Views and Representations as key representative elements, this approach is more model-oriented, providing a better understanding of how to design an IS, and for this reason it was considered in ISD2K.

III. Solution Proposed-ISD2K

The ISD2K - Information System Data to Knowledge Oriented - is a reference model that enables architects, managers, and users to design, realize, and evaluate an (IS) oriented to transform D into K. It uses a descriptive approach do provide a static and dynamic reference using natural language to describe an IS in an interactive process between distinct actors in an organization. The descriptive approach and the natural language add low complexity to this model making it easy to be implemented. Conceptual models are expressed by descriptions about the topic they represent and can be defined and presented through examples [37]. IS2DK uses 6 views representing distinct perspectives about an IS expected to KP to the organization (Table II).

TABLE II
ISD2K - Views

Views	Description
V1. Initial View	What top management and/or owners wants
V2. Organization View	What the organization wants
V3. Systemic View	Integration of the previous visions, systemic vision - SUM(V1 + V2)
V4.Technology View	Technological possibilities and limitations
V5. Detailed View	How to do it, based on the previous views – SUM (V3+V4)
V6. IS	The system itself

Each View is described per 9 representations (Table III).

TABLE III
ISD2K - Representations

Representation	Questions to be answered
R1. Motivation/ Strategy	Why to produce K?
R2. Process/Functions	What processes/functions will produce K?
R3. Data	What data will be used to produce K?
R4. Who/Organization	Who will be involved in producing K?
R5. Connections/ Network	What connections are needed to produce K?
R6. Time – Cycles/ Scheduling	When will it be produced, with what recurrence?
R7. Usage	How will the IS be used to produce K?
R8. Acceptance/ Quality	What levels of acceptance must the IS has to produce K?
R9. Impact	What impact, internal and external, will the produced K have?

This conceptual, static model can also be represented in dynamic form, describing how it works - how it operates. Fig. 1 represents this dynamic model.

A. Detailed Description - Views

V1.Initial View - Corresponds to an executive summary for an administrator, investor, or project manager, who wants to estimate the scope of an IS and the objectives to be achieved.

V2.Organization View – It´s a deepening that goes beyond the desirable, generalized initial vision, through the adaptation to the existing structure of the organization, and establishes specific definitions for its representations, considering aspects identified previously.

V3.Systemic View - It is the conversion of the previous views into a systemic approach, aggregating the perspectives of V1 and V2 into a single vision.

V4.Technology View - Match the systemic vision with the technological availability - the tools, skills, materials (hardware, software, operating system, network infrastructure), programming languages, storage and transmission technologies, devices that will access the IS, i.e.

V5.Detailed View - Elaborated after the scope definitions, organizational contextualization, and technological definitions of the IS. It evolves from what the IS should do to how it should be done.

V6.IS – The system itself.

IV. Use Case

Is it possible to represent an IS from ISD2K? A study was developed from a synthetic use case, to evaluate, in a controlled environment, how the model would behave in a small organization, with few technological resources, which corresponds to the vast majority of Brazilian companies (Small Medium Business/SMB). Here, a hypothetical company resells only a product and needs, on a daily basis, to define sales prices based on costs, sales history and profit, to optimize the results (profit) to be validated, later, based on the results achieved.

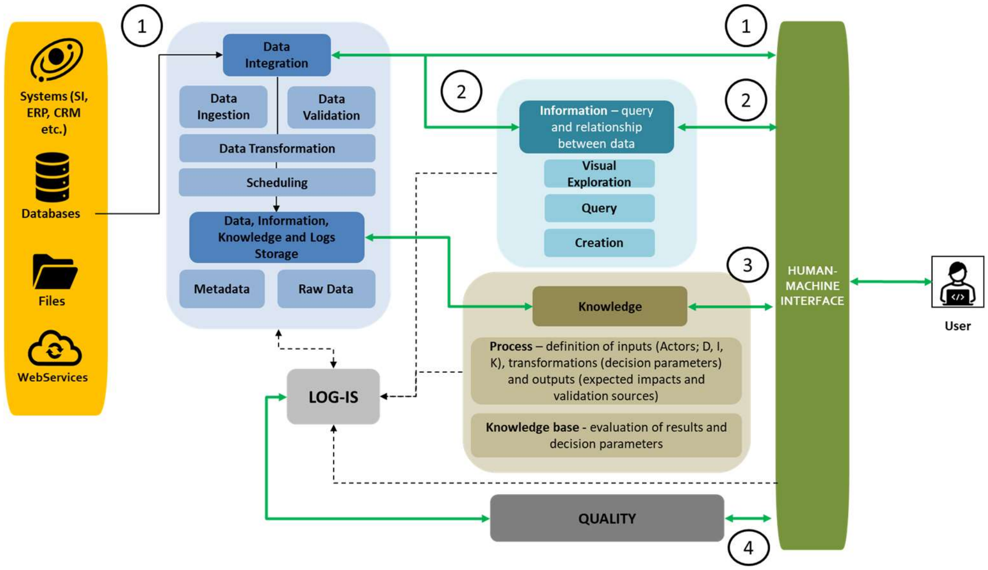

Fig.1. (1) Consuming data/metadata to enable validation of recurring data, transforming it if needed; scheduling definitions; data/metadata storage. (2) I - different D and their relationships classified as I; (3) K - definition of inputs/inputs, actors and D, I and K needed for decision making; parameters that will transform D, I and/or K into a decision; Expected results and data sources to validate this result; knowledge base of decision making, evolutionary aspect of K. (4) Log (LOG-SI) - data, user interactions, quality characteristics of the IS (response time and availability); QUALITY of the IS in KP; Effectiveness with which KP occurred (IMPACT).

The case was carried out through the first 4 views in each of the 9 representations since it was not the objective of this evaluation to detail the development (V5) or production of a software program (V6).

V1 and V2 were described by each of the 9 representations. V3 summarized V1 and V2 and was described by the following concept of operation: *The desired IS has the purpose of implementing a* KP *architecture to support decision-making. The IS will support the pricing decision making process performed on a daily basis. The IS will allow the construction of a knowledge base with decisions taken, so that these decisions can be compared with data from other sources to evaluate if the decision reached the desired goal and allow future decisions to be taken more assertively. The IS will be responsible for collecting data from other systems and sources in the organization, allowing decision makers to visualize and compare data relevant to pricing. Through the IS, it will be possible to register the price definition, which will be distributed automatically so that the salespeople can use the new definition immediately. The IS will also have features that allow the evaluation of the quality of the data that will support the decision-making process, the quality of the interaction (selection and consultation) with the data proposed by the user with the system, and the quality of the IS availability. The IS will have a graphic interface that allows performing the tasks of data extraction, validation, loading and distribution, as well as the scheduling of these activities, with alerts to the user in case something happens out of the pre-established deadlines and standards. All IS interactions with data (ingestion, processing, and distribution), with users and in the interaction between data and users, will be recorded for quality monitoring purposes.*

The dynamic representation of this use case was done in V3-R1, which expresses the functioning of a KP cycle having as its materialization the decision making, in this case, daily prices that need to be defined based on the costs of the day (suppliers' prices), on the sales history, and on the profit history (Fig. 2).

V4 for this use case considered the simplest as possible environment - the IS was described to be implemented in spreadsheet on-line, updated and accessed by required actors, with daily updates in a schedule agreed between the actors.

V. Creation and Application of Questionnaires

Is it possible to evaluate existing IS in terms of their ability to produce knowledge? In order to analyze the existing IS about their propensity to KP, according to ISD2K model, questionnaires were developed with 9 questions (representations), for each of the first 4 views, totaling 36 questions. This set allows an adherence analysis of total values to be performed, per vision and representation of an existing IS (Table IV).

Adherence does not mean that the IS produces K from D, but that it has the necessary conditions. The higher the adherence (total, by vision and/or representations), the higher its potential. To structure answers, the Lickert scale [38], was used (Table V).

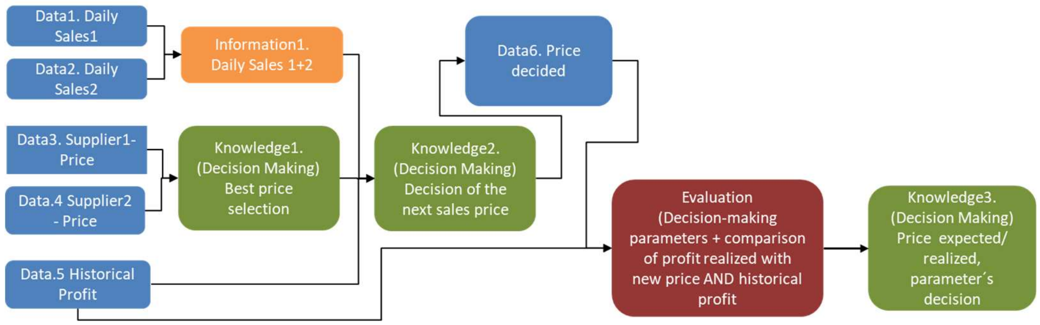

Fig.2. D1 and D2 become I1; D3 and D4 become K1; I1, K1, D5 become K2; K2 is stored as D6. An evaluation of K2, considering historical profit validates if it was a good decision based on previous similar decisions. All D,I, K and decision´s parameters are stored in a Knowledge Base (K3).

TABLE IV
ISD2K – MAXIMUM SCORE - QUESTIONNAIRES

	V1. Initial View	**V2. Organization View**	**V3. Systemic View**	**V4. Technology View**	**Maximum Score**
R1. Motivation	5	5	5	5	20
R2. Process	5	5	5	5	20
R3. Data	5	5	5	5	20
R4. Who	5	5	5	5	20
R5. Connections	5	5	5	5	20
R.6 Time - Cycles	5	5	5	5	20
R7. Usage	5	5	5	5	20
R8. Acceptance	5	5	5	5	20
R9. Impact	5	5	5	5	20
Maximum Score	45	45	45	45	

TABLE V
ISD2K – SCORE SCALE - QUESTIONNAIRES

5-Complety agree	4-Partially agree	3-Neither agree nor disagree	2-Partially disagree	1-Strongly disagree
5	4	3	2	1

Questions to understand the alignment of the IS with the purpose, goals, business process, data used, availability, ease and effectiveness in using the IS, stakeholders, the relationship with the data and the decision-making process in use by the end users, and also whether the IS records the results of the decisions made as data in the system to understand their impacts were included in these questionnaires.

Two systems were evaluated - Mission Planning System and Burnt Areas DB System, both in use in INPE. Researchers who worked on the conception, development, and use of these IS answered this questionnaire. Then an analysis of each system was prepared, and this analysis was given to each researcher. Afterwards, to validate their effectiveness, a new questionnaire about the analysis and adherence questionnaires was presented to the researchers. Both researchers answered yes to all questions in this new questionnaire. Table VI.

Both systems did not show full adherence to the ISD2K (average of 74.20%), showing that both have conditions for producing information and knowledge, but are currently focused on data dissemination. The analysis by view was particularly interesting, because in most of the views, (6 out of

TABLE VI
ISD2K – SCORE SCALE - QUESTIONNAIRES

	Mission Planning System	**Burning DB System**
1. Was the presentation about the ISD2K framework and its purpose of proposing a reference model for ISs oriented to the conversion of data into knowledge, in other words, knowledge production, clear?	Yes	Yes
2. Were the questionnaires clear and succinct, about their questions and answers?	Yes	Yes
3. Was is understood, in each of the views, the issues for each representation?	Yes	Yes
4. Do the results of the total adherence assessment reflect your perceptions about the system?	Yes	Yes
5. Do the results of the vision adherence assessment reflect your perceptions about the system?	Yes	Yes
6. Do the results of the representation-based adherence assessment reflect your perceptions about the system?	Yes	Yes

9) the average value of the sum of visions 1 and 2 were not matched in vision 3, that is, in the systemic integration of these views there was a loss, either in the owner's or the organization's demand, for the definition of V3. In the thematic analysis (representations) it was observed that most of the functional aspects, common to data dissemination systems, are present, but

aspects related to use, quality and impact had the worst scores (average 12 out of 20).

In summary, both systems reflect their orientation towards data dissemination, with some functionalities for interaction and possible identification of information, and no characteristics associated with recording the KP process, being considered data distribution systems. Still, both have most of the necessary conditions (74.20%) to become IS that convert D into K.

VI. Results and Discussions

A. Comparison of the Proposed Method Against other Authors:

Comparison of the proposed method with those of other authors: ISD2K uses a descriptive model based on different insights, as in 9 of the 13 models evaluated (Table I); it uses 6 revised initial representations (based on SZF and others) and adds 3 additional representations (Use, Quality and Impact); The focus is on the planned and cyclical production of K from D instead of the other 13 models; Guiding aspect - sequential use of views and representations (only 2 models use this approach); Technical approach – generic instead of employing commercial technologies like 4 related papers. None of the related works label data based on its purpose (D, I, and K), and there is no mention to QUALITY as an aspect of IS to be considered during the modelling, either by the quality of its data, the interaction with the IS and the IS itself (availability, response time, etc.); And no evidence about the aspect of transformation of D into I, which is the contextual relationship between two or more data.

B. Critical Analysis - Adherence Questionnaires: Easy to use, to interpret and answer. The delimitation of views - 4 to 6 - fulfilled its objective of assessing IS orientation in KP.

C. Critical Analysis - Use Case: the use in sequence of visualizations and representations defined in the model was appropriate, other models do not advocate this approach; the proposed model was applied without software coding; demonstrated that it is possible to establish CP in an IS based on ISD2K.

D. Limitations – missing definitions of systems engineering aspects; need to test in other scenarios than aerospace systems; need to evaluate more real cases; assumption that decision-making processes already exist.

VII. Conclusion and Future Work

The studies in the literature presented the distinction and the relationship between D, I and K and the KP as a production function, containing D in its inputs, user interaction as the transformation step and K as the output of the process, allowing to measure this process and using IS as KP platforms. The success factors of IS allowed identifying characteristics to qualitatively evaluate KP in IS.

An comprehensive IS model was described, measurable in KP from a production function perspective, relating inputs, transformation processes and outputs to D, I and K, involving the perception of various actors in the use of data in an organization, and with representations of how the user produces I and K (USE), expected patterns of D, the interaction with the user (I) and the IS itself (QUALITY) and evaluating results achieved by the use of the K produced (IMPACT).

The use case and the evaluation of existing IS validated ISD2K in both situations: design and evaluation of existing systems. The general objective of this work was reached, through the presentation of a model that contemplates a static and dynamic description.

The related ISA models presented are focused on the availability of D and its accessibility. The distinction between D, I and K is understood but not defined in the IS features. I and D are sometimes treated as synonyms. It is evident that new propositions are needed on model references that address this distinction and the production of K from D.

Existing IS have already been implemented to handle KP, albeit indirectly and/or implicitly. Designing, assessing and developing their KP capacity makes it possible to reduce the gap between D and K volume.

The results obtained allow future work on D2K conversion through SI, such as, but not limited to: a) extending the model to the use of modeling languages or model-based systems engineering techniques; b) employ the proposed model in real use cases and applying adherence analysis questionnaires in other sectors than aerospace; c) developing specific components, attachable to existing IS, for integration, production of I, K and Quality; d) automating I production through unsupervised learning algorithms and K through supervised learning algorithms.

References

[1] C. TILLY, “The Old New Social History and the New Old Social History”. *Review* (Fernand Braudel Center), vol. 7, no. 3, pp. 363–406,1984.

[2] D. LANEY, “3D data management: Controlling data volume, velocity, and Variety”. *MetaGroup*. 2001. Available: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-VolumeVelocity-and-Variety.pdf.

[3] PDU - PLANO DIRETOR DA UNIDADE - INPE - 2022-2026. 2022. Available: https://www.gov.br/inpe/pt-br/central-de-conteudo/publicacoes/repositorio-de-arquivos/plano-diretor-2022-2026.pdf.

[4] H.; VALLADAS, J. CLOTTES, J.M GENESTE, M. A GARCIA, H. CACHIER and N. TISNÉRAT-LABORDE, “Paleolithic paintings: Evolution of prehistoric cave art”. *Nature*, Vol. 413, no.6855, pp.479-479, 2001.

[5] I. NONAKA and H. TAKEUCHI, “The Knowledge Creating Company”. Oxford University Press, New York, USA. 1995, pp. 50-51.

[6] P. SCHUBERT, D. LINCKE, B. SCHMID, “A global knowledge medium as a virtual community: the NetAcademy concept”, *in Proc. 4th Americas Conference on Information Systems*, Baltimore, MD, USA, 1998.

[7] M. ZACK, “An architecture for managing explicated Knowledge”. *Sloan Management Review*, Sep, 1998.

[8] J. Shendure *et al.*, “DNA sequencing at 40: past, present and future,” *Nature*, vol. 550, no. 7676, pp. 345–353, Oct. 2017, doi: 10.1038/nature24286.

[9] I. NONAKA, “A Dynamic Theory of Organizational Knowledge Creation”, *Organization Science*, vol. 5, pp.14-37, 1994.

[10] J.C. SPENDER and R.M. GRANT, "Knowledge, and the firm: Overview". *Strategy Management Journal*, vol. 17, pp. 5-9, 1996. Available: https://doi.org/10.1002/smj.4250171103.
[11] N. SHARMA, "The Origin of Data Information Knowledge Wisdom (DIKW) Hierarchy", 2008.
[12] R. ACKOFF, "From data to wisdom". *Journal of Applied Systems Analysis*, vol.16, pp. 3-9, 1989.
[13] H. CLEVELAND, "Information as a Resource". *Futurist*, vol.16, pp.34-39, Dec 1982.
[14] R. L. ACKOFF, "Re-Creating the Corporation: A Design of Organizations for the 21st Century", Oxford University Press, New York, USA, 1999.
[15] N. NIU, L. XU, and Z. BI, "Enterprise Information Systems Architecture—Analysis and Evaluation", *IEEE Trans. on Industrial Informatics*, Ed. 9, pp. 2147-2154, 2013.
[16] Systems and Software Engineering — System Life Cycle Processes. ISO/IEC/IEEE 15288, 2015.
[17] INCOSE. "Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities", version 4.0. Hoboken, NJ, USA: John Wiley and Sons, 2015.
[18] S.K. BOELL, D. CECEZ-KECMANOVIC, "What is an Information System?", *in Proc. 48th Hawaii International Conference on System Sciences*, Hawaii, vol.8, no.238, pp. 4959-4968, 2015.
[19] W.H. DELONE and E.R. MCLEAN, "Information systems success revisited", *in 35th HICSS Proc*., pp. 2966-2976, 2015.
[20] T.D. NGUYEN, T. NGUYEN, T. CAO, "Information Systems Success: A Literature Review", *in Proc*. *International Conference on Future Data and Security Engineering*, vol. 9446, pp. 242-256, 2015.
[21] C. E. DICKERSON, S. M. SOULES, M. R. SABINS, P. H. CHARLES, "Using Architectures for Research, Development, and Acquisition", 2004, Available: https://apps.dtic.mil/sti/pdfs/ADA427961.pdf.
[22] Systems and Software Engineering—Recommended Practice for Architectural Descriptions of Software Intensive Systems, ISO/IEC/IEEE 42010, 2011.
[23] System and Software Engineering—Life Cycle Management—Part 1: Guide for Life Cycle Management, ISO/IEC TR 24748-1, 2010. Available: http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html.
[24] J.A. ZACHMAN, "A Framework for Information Systems Architecture". IBM Syst. J., vol. 26, pp.276-292, 1987
[25] J.F. SOWA and J.A. ZACHMAN, "Extending and Formalizing the Framework for Information Systems Architecture". *IBM Syst. J*., vol.31, pp.590-616, 1992
[26] R. EVERNDEN, "The information framework". *IBM Syst. J.* vol.35, pp.37–68, Jan. 1996.
[27] R. S. JAMUNA and M. S. ASHOK, "A survey on service-oriented architecture for e-learning system in Intelligent Agent & Multi-Agent Systems", *in Proc. International Conference on Intelligent Agent & Multi-Agent Systems*, Chennai, India, pp.1-3, 2009, 10.1109/IAMA.2009.5228029.
[28] R. COVINGTON, H. JAHANGIR, G. WRIGHT, P. SILVERSTEIN, H. DIA and B. RASMUSSEN, "The Oracle enterprise architecture framework". *White Paper Oracle*, Oct. 2009.
[29] J.V. WOUT, M. WAAGE, H. HARTMAN, M. STAHLECKER and A. HOFMAN, "The Integrated Architecture Framework Explained: Why, What, How". Ed. *Springer Science+Business Media*, Berlin/Heidelberg, Germany, 2010, pp. 10-14.
[30] J.P. ZACHMAN, "The Zachman Framework Evolution", Apr, 2011. [Online]. Available: https://www.zachman.com/ea-articles-reference/54-the-zachman-framework-evolution.
[31] Y. YOO, O. HENFRIDSSON, and K. LYYTINEN, "The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research", *Information Systems Research,* vol. 21, pp. 724-735, 2010.
[32] G. CUGOLA and A. MARGARA, "Processing flows of information: From data stream to complex event processing". *ACM Comput. Surv*., vol. 44, 15:1-15:62, 2012.
[33] S. MAJD, A. MARIE-HÉLÈNE, and M. ALOK, "An Architectural Model for System of Information Systems," in *Proc.* ISDE15 Sixth International Workshop on Information Systems in Distributed Environment, Rhodes, Greece, 2015, pp.411-420.
[34] B. CHRISTOPH, M. VALENTIN, C. PLUDER and V. LEHNERT, "Data Information Framework". U.S. Patent Application 20170270163, 2017.
[35] *The TOGAF Standard, Version 9.2*, T. O. GROUP – TOGAF, Hertogenbosch, The Netherlands, 2018, [Online], Available: https://publications.opengroup.org/c182.
[36] W.L. CHANG, D. BOYD and O. LEVIN, "NIST Big Data Interoperability Framework: Vol. 6, Reference Architecture", 2019. [OnLine]. Available: https://doi.org/10.6028/NIST.SP.1500-6r1.
[37] M. ALAVI and D. LEIDNER, "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues". *MIS quarterly*. Vol. 25, no.1, pp. 107-136, Mar. 2001.
[38] A. JOSHI, S. KALE, S. CHANDEL, and D. PAL, "Likert Scale: Explored and Explained". *British Journal of Applied Science & Technology*, Vol.7, pp.396-403, 2015.

Marco Fidos holds a bachelor's degree in military sciences from *Academia Militar das Agulhas Negras* (1997), a postgraduate degree in Marketing from *Universidade Veiga de Almeida* (2001) and an MBA from *Fundação Getúlio Vargas* (2002). He is a master's student in the graduate program in Space Systems Engineering at INPE. He currently holds the position of technical manager of big data and artificial intelligence solutions at ATECH, defining architectures and implementing solutions for big data volumes and the use of various analytical approaches.

Mauricio G. V. Ferreira was graduated in Technology in Data Processing from Faculdade de Administração e Informatica (1987), graduated in Business Administration from Faculdade Maria Augusta (1993), master's degree in Applied Computing from Instituto Nacional de Pesquisas Espaciais (1996) and PhD in Applied Computing from Instituto National Institute for Space Research (2001). He currently is a full Professor at INPE's Graduate Course in Space Engineering and Technology (ETE); Member of the organizing committee of the international space congress - SPACEOPS. He produced more than 180 articles, supervised 17 doctors, 27 masters and 2 post doctorates. He´s also works as a coordinator and researcher at the Center for Tracking and Control of Satellites at INPE.