

# FROM DATA TO KNOWLEDGE MANAGEMENT: What to Consider

By Melanie J. Adams, Paige E. Kane, PhD, CPIP, and Anne Greene, PhD

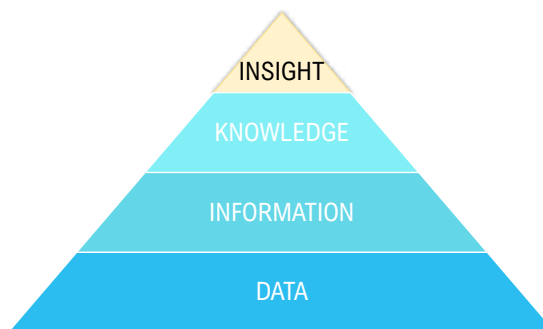
Although data and knowledge are both stand-alone disciplines that need to be systematically managed, they also must have a connection. Understanding the relationship between data and knowledge management processes and how people are leveraging advances like Pharma 4.0™ combined with these processes enables quality data transition to knowledge that can help pharmaceutical companies. The authors also want to generate understanding on how using the knowledge acquired by people through experience (tacit knowledge) can further connect both data and knowledge management systems, yield positive strategic results, and deliver more efficient processes within organizations.

Knowledge management (KM) is a stand-alone discipline; however, it has relationships with other disciplines. This article explores the relationship between data and knowledge, a deeper look that follows up on the Pharmaceutical Knowledge Ecosystem [1], which looks at how the pharmaceutical industry acquires data, transforms this data into tangible knowledge, and derives valuable insights throughout the process.

The origin of this ecosystem builds upon the data, information, knowledge, and wisdom (DIKW) hierarchy [2]. Over time, this theory has been developed, and was published in 2018 replacing *wisdom* with *insights*, as shown in Figure 1.

Kane reported that wisdom is widely agreed to be a “uniquely

Figure 1: DIKW hierarchy, as adapted by Kane [3].



human” characteristic, whereas insights take into account current technological advances and allow data transformation to lead to insights. Although insights may be derived by people with knowledge and experience, they may also be derived from computing or machine-learning models that identify trends and correlations previously not possible to see from experience alone.

Following on from that: Although it is useful to replace wisdom with insights in the DIKW hierarchy, on reflection, Lipa [4] proposed that the goal is to achieve understanding. Insights could be regarded as discrete events, whereas understanding represents a holistic comprehension: a state of mastery of a given domain or topic. This state of mastery could manifest, for example, as a mechanistic understanding of a complex chemical reaction or as an accurate predictive model for the relationship between process parameters and their impact on final product quality attributes. In each example, there is a progression from being naïve to developing

understanding (i.e., a state of mastery) based on accumulated data, information, knowledge, and insights, as depicted in Figure 2 [4].

Mastering the progression of data to information to knowledge to insights and understanding (DIKIU) presents the opportunity to be able to make informed and effective decisions based on accumulated evidence, as provided by the underlying structure.

## DATA VERSUS KNOWLEDGE

In everyday conversations, it is not unusual to hear the words *data* and *knowledge* used interchangeably. This section offers definitions and descriptions of these terms.

The Cambridge Dictionary defines data as “information, especially facts or numbers, collected to be examined and considered and used to help decision making, or information in an electronic form that can be stored and used by a computer” [5]. It defines knowledge as “understanding of or information about a subject that you get by experience or study, either known by one person or by people generally” [5].

The definition of data emphasizes information in its raw form, without context. It is context and understanding that increases data’s usefulness and transforms it into knowledge.

From the definitions of data and knowledge, it is clear that having information or understanding about a subject is gained through experience. It should be noted that experience is known or gained by people.

## MANAGING DATA AND KNOWLEDGE

### Managing Data

Transferring data to knowledge does not typically happen organically. Procedures that enable users to derive value (e.g., lead to decision-making or insights) from an organization’s data or knowledge base should be in place to ensure the information can be validated and trusted. To do this, there should be several procedures in place.

The ISPE *GAMP RDI Good Practice Guide: Data Integrity by Design* has described managing data as a life-cycle process with five phases [6]. The key points in the life cycle are:

- Creation
- Processing
- Review, reporting, and use
- Retention and retrieval
- Destruction

The authors of this article would like to highlight and include two further important activities and processes for managing data to this list within Table 1: data governance and data integrity.

Table 1 highlights examples of data-related processes and why they are important.

### Managing Knowledge

As with other management disciplines, definitions for KM are plentiful. In this article and in alignment with pharmaceutical industry related literature, two definitions are highlighted:

Figure 2: DIKW hierarchy, as adapted by Lipa [4].



Table 1: Data-related processes.

Process	Reason for Importance
Data governance	Governance refers to what decisions must be made to ensure effective management and use of IT (decision domains) and who makes the decisions (locus of accountability for decision-making) [7].
Creation: data creation and collection	Many different data sources exist; generally the use of spreadsheets is widespread, and some data is available in handwritten notes, lab notebooks, and printouts from stand-alone devices. These manual notes and printed data sheets are manually transcribed into electronic format.  There does exist a more sophisticated case where data is stored in commercially available databases such as laboratory information management systems (LIMS) or in-house systems set up by organizations themselves [8].
Processing: data analysis and processing	The main purpose of collecting and analyzing data in commercial manufacturing is to set up a product and process control environment. Raw data is given context by adding information and explaining what the data means, thus presenting information in a required format.
Retention and retrieval: data retention and retrieval	In routine manufacturing, manufacturing execution systems (MES) control and document the manufacturing processes.  For analytical measurement results, LIMS systems are often used along with Excel spreadsheets. In the case of Excel spreadsheets, GMP validation is possible.  Manual extraction of data from paper-based batch records is another option [8].
Review, reporting, and use: data storage, dissemination, reporting, and use	Once generated, the data and information require long-term storage and simple reuse options. KM tools organize the acquisition, storage, and dissemination of the product knowledge.
Destruction: data destruction	Ensure the correct original data is disposed of after the required retention period [6].
Data integrity	Product data should ensure end-to-end traceability and data integrity in order to release a batch. It is expected that the integrity of pharmaceutical data assets should be compliant with attributable, legible, contemporaneous, original, and accurate (ALCOA) principles [9].

KM processes can assist in ensuring knowledge is shared in the form it is required for the end user and it is communicated, consistent, and findable.

ICH Q10 defines KM as:

A systematic approach to acquiring, analysing, storing and disseminating information related to products, manufacturing processes and components. Sources of knowledge include but are not limited to prior knowledge (public domain or internally documented); pharmaceutical development studies; technology transfer activities; process validation studies over the product lifecycle; manufacturing experience; innovation; continual improvement; and change management activities [10].

American Productivity and Quality Center (APQC) defines KM as:

The application of a structured process to help information and knowledge flow to the right people at the right time so they can act more efficiently and effectively to find, understand, share, and use knowledge to create value [11].

The ICH definition describes KM with a more narrow perspective than the APQC definition; the APQC definition is more commonly used by KM practitioners because it embraces the two main aspects of KM: The needs of the knowledge user and the needs of managing knowledge within an organization.

Table 2 presents examples of KM processes and tools that enable a systematic approach to knowledge flow and indicating their importance. These KM are discussed in length in the ISPE *Good Practice Guide: Knowledge Management in the Pharmaceutical Industry* [12].

### RELATIONSHIP BETWEEN DATA AND KNOWLEDGE

Some challenges in assessing the relationship between data and knowledge include large volumes of information make it difficult to focus on the most important elements; multigenerational preferences in the workplace for consuming information; the concept of data privacy; and demonstration of the KM value proposition, which enables buy-in and sponsorship, embedding the concept of knowledge as an asset [13].

Table 2: KM processes.

Processes and/or Tools	Reason for Importance
KM plan KM maturity assessment	These are required for planning, understanding requirements of the organization, and defining the process [12].
Content management Searching platforms Product knowledge	These relate mostly to explicit-based knowledge: “a declarative type of knowledge that can be readily articulated (in words or images), coded, stored, and accessed” [12]. Explicit knowledge can be learned as facts.
Communities of Practice Lessons learned Tacit knowledge retention	These relate mostly to tacit knowledge: “a context-specific type of knowledge, acquired through personal experience or internalization and would reside within people’s minds rather than a physical media or information system. Often referred to as ‘know how.’” [12]. Tacit knowledge is gained through experience. It is rarely written down and is hard to capture and validate, but when applied, it increases right first time (RFT) and facilitates continual improvement.
KM roles KM training KM governance	Enablers to the KM process [12].

It is through data analysis and processing that the relationship between data and knowledge becomes evident. To manage the large volumes of information and extract the important elements, the analysis and processing of data has to add value. To focus on what that value is for an organization, define the objective that an organization or a team needs to achieve from the data, perhaps in the format of a problem statement. To solve the problem, one needs to understand what sources of data and information are needed, and in particular what type of analysis is to be carried out. For example:

- Descriptive analysis: Identifies what has already happened.
- Diagnostic analysis: Focuses on understanding why something happened.
- Predictive analysis: Allows one to identify future trends based on historical data.
- Prescriptive analysis: Allows one to make recommendations for the future.

After the sources of data and information needed are identified and the type of analysis determined, the required data should be collected and aggregated. This includes quantitative (numerical) data or qualitative (descriptive) data. In the pharmaceutical sector, several types of data management platforms that automate data collection are used; some examples can be found in Table 1.

The data from these platforms can be considered “clean” (i.e., data that has had errors, duplicates, and unwanted data points removed) because they are validated systems. The data is reported in a structured manner.

It is through the analysis of data that information, knowledge, and insights are gained. These insights should be shared within the

organization with key members who need them. This flow of knowledge is important because raw data will yield no value without knowledge; thus, analysis is needed, which enables insights to be shared in a digestible manner by everyone who receives the information.

Often key decisions are made based on these insights, which have been communicated in the form of reports, dashboards, and interactive visualizations, so they must be clear and unambiguous. Ideally, all data should be shared so decisions are made based on a complete picture, and the final decision is scientifically sound and based on insightful facts. Insights that are open to interpretation should be flagged. Communication is key when sharing this information. KM processes can assist in ensuring knowledge is shared in the form it is required for the end user and it is communicated, consistent, and findable. This is the real function of the Knowledge Ecosystem.

### **FUTURE CONSIDERATIONS**

Pharma 4.0™ [14] proposes that the pharmaceutical industry adopt a standardized approach to the collection, storing, and analyzing of data. It suggests that the pharmaceutical industry needs a system that can span across one organization to remove silos and data isolation, is a user-friendly database, and can interact with other systems (interface). The purpose of this is to avoid data

inconsistency. Data itself cannot take any actions other than what it is programmed to do; however, it can be programmed to take actions that could lead to future problems due to inconsistency.

When maximizing the flow of knowledge in an organization, four key factors should be considered to enable a holistic KM program: people, process, content, and technology [13]. All of these factors are required to be successful; if one is missing, knowledge flow will not succeed. People are the primary consumers and generators of knowledge. Technology and content alone will not solve knowledge flow issues. If people are not using the Knowledge Ecosystem, knowledge flow will be poor. People manage processes and understand the content required, keeping in mind as well that people hold the organization's tacit knowledge.

Knowledge is a valuable asset, but often it is not treated that way. Approaches to KM and sometimes data management can vary. This can also result in poor flow of knowledge. Organizations should understand that in the current climate of increasingly complex information generation and large volumes of data, those who manage knowledge well can realize a competitive advantage [13].

### **CONCLUSION**

With the use of technology, a huge amount of data and information can be processed. This ability is growing exponentially;



## **Instructor-Led, On Demand, and Custom Training**


**The Trusted Resource for Pharmaceutical  
Manufacturing Professionals and Regulators  
Worldwide**

**Learn More at [ISPE.org/Training](https://www.ispe.org/training)**

## When maximizing the flow of knowledge in an organization, four key factors should be considered to enable a holistic KM program: people, process, content, and technology.

however, processing through technology solutions is limited to data and explicit knowledge. Although various technologies have been developed to store, organize, and reuse information, tacit knowledge (the human factor) is still needed to integrate and make sense of this information to create value. Through KM processes (capturing explicit knowledge) and communities of practice connecting people (capturing tacit knowledge), explicit and tacit knowledge become available for use. The more subject matter experts (SMEs) connect across the organization, the more powerful decision-making and the resulting actions will become.

Managing organizational data and knowledge should be a process-driven systematic approach with a life cycle so that data, information, and knowledge are proactively and continuously captured, analyzed, stored, and disseminated. A robust and reliable KM ecosystem integrates product and process information and supports the capture of explicit and tacit knowledge.

As pharmaceutical organizations adopt the Pharma 4.0™ philosophy and embrace the huge amount of data, data connections, structured information, and knowledge in repositories, opportunities for more effective decision-making emerge. This will have a profound effect on how business is managed in the future. 

### References

- Adams, M. J., P. E. Kane, A. Greene, and M. J. Lipa. "Exploring Pathways from Data to Knowledge to Insights in the Pharmaceutical Industry: 'Introducing the Pharmaceutical Knowledge Ecosystem.'" *Level 3*, no. 2 (2022). Accessed 28 March 2022. <https://arrow.tudublin.ie/level3/vol16/iss2/1/>
- Rowley, J. "The Wisdom Hierarchy: Representations of the DIKW Hierarchy." *Journal of Information Science* 33, no. 2 (2007). doi: 10.1177/0165551506070706
- Kane, P. E. "Leveraging Knowledge Management in the BioPharma Industry." Presentation at 2016 ISPE Europe Conference (2016).
- Lipa, M., K. O'Donnell, and A. Greene. "Knowledge as the Currency of Managing Risk: A Novel Framework to Unite Quality Risk Management & Knowledge Management." Institute of Validation Technology (2020). <https://www.ivtnetwork.com/article/knowledge-currency-managing-risk-novel-framework-unite-quality-risk-management-and-knowledge>
- Cambridge English Dictionary. Data and Knowledge Definition. Cambridge University Press, 2022. <https://dictionary.cambridge.org/dictionary/english/data>
- International Society for Pharmaceutical Engineering. *GAMP® RDI Good Practice Guide: Data Integrity by Design*. North Bethesda, MD: International Society for Pharmaceutical Engineering, 2020.
- Khatri, V. and C. V. Brown. "Designing Data Governance." *Communications of the ACM* 53, no. 1 (2010): 148–152. doi: 10.1145/1629175.1629210
- Steinwandter, V., D. Borchert, and C. Herwig. "Data Science Tools and Applications on the Way to Pharma 4.0." *Drug Discovery Today* 24, no. 9 (2019): 1795–1805. doi: 10.1016/J.DRUDIS.2019.06.005
- Leal, F., A. E. Chis, S. Caton, H. Gonzalez-Velez, J. M. Garcia-Gomez, M. Dura, A. Sanchez-Garcia, et al. "Smart Pharmaceutical Manufacturing: Ensuring End-to-End Traceability and Data Integrity in Medicine Production" *Big Data Research* 24: 100172 (15 May 2021). doi: 10.1016/J.BDR.2020.100172
- International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. "ICH Harmonised Tripartite Guideline Q10: Pharmaceutical Quality System." Published June 2008. <https://database.ich.org/sites/default/files/Q10%20Guideline.pdf>
- American Productivity & Quality Center. "Knowledge Management." (1993). <https://www.apqc.org/expertise/knowledge-management>
- International Society for Pharmaceutical Engineering. *Good Practice Guide: Knowledge Management in the Pharmaceutical Industry*. North Bethesda, MD: International Society for Pharmaceutical Engineering, 2021. <https://ispe.org/publications/guidance-documents/good-practice-guide-knowledge-management-pharmaceutical-industry>
- Biophorum. "BPOG Biomanufacturing Technology Roadmap: 7. Knowledge Management." <https://www.biophorum.com/download/biomanufacturing-technology-roadmap-7-knowledge-management/>
- International Society for Pharmaceutical Engineering. "Pharma 4.0 Operating Model." <https://ispe.org/initiatives/pharma-4.0>

### About the authors

**Melanie J. Adams** is an Associate Director in the MSD (Merck & Co., Inc. Carlow, Ireland) Knowledge Management Center of Excellence. She has a BSc from Maynooth University and a Hons HDip from UCC. Melanie is an industry leader with over 17 years of experience, including SOD manufacturing operations, quality control, technical services, technical transfer, and the deployment of KM strategies in SOD and biopharma sites in Wyeth and Pfizer pharmaceutical organizations. She was also a member of the ISPE *Good Practice Guide: Knowledge Management in the Pharmaceutical Industry* team, leveraging her practical experience deploying KM within industry when writing the guide. Melanie has been an ISPE member since 2019.

**Paige E. Kane, PhD, CPIP**, is a member of the Technical University Dublin Pharmaceutical Research Team. She is an industry leader with over 29 years of experience including Merck & Co., Genetics Institute, Wyeth, Pfizer, Monsanto, and US government, spending many years leading knowledge management programs and approaches for the pharmaceutical industry. Paige previously led quality systems groups focusing on automation compliance, data integrity, computer validation, and change control during startup and operations for biologics facilities in the US and Ireland. She has been a member of multiple industry guidance author teams and global industry Communities of Practice (GAMP® and Biotechnology) as well as leading other strategic teams for industry, including the ISPE *Good Practice Guide: Knowledge Management in the Pharmaceutical Industry* team. She is a co-editor/contributor for "A Lifecycle Approach to Knowledge Excellence in the Biopharmaceutical Industry" (2017) and holds a PhD in pharmaceutical and regulatory science from Technical University Dublin. Paige has been an ISPE member since 2000.

**Anne Greene, PhD**, heads the Pharmaceutical Regulatory Science Team (PRST) in Technological University Dublin, where she is a senior lecturer and has spearheaded the development of several MSc and BSc pharmaceutical programs. She has supervised industry-based executive students to PhD awards in areas of quality risk management, knowledge management, operational excellence, and PQS. Prior to embarking on an academic career, Anne worked at a senior level for several years in the pharmaceutical sector in validation and technical management roles. Anne has a BSc and PhD in chemistry from University College Dublin. She has been an ISPE member since 2011.