NDE 4.0: Smart NDE

Debejyo Chakraborty
General Motors Research & Development
30470 Harley Earl Blvd.
Warren, MI 48092-2031
Email: debejyo.chakraborty@gm.com

Megan E. McGovern
General Motors Research & Development
30470 Harley Earl Blvd.
Warren, MI 48092-2031
Email: megan.mcgovern@gm.com

***Abstract*—Current nondestructive evaluation (NDE) and testing practices are burdened by "conventional wisdom," which requires that the operator be involved in all aspects of the data collection, transfer, and analysis. Many of these shortcomings are rooted in inefficiency and can be addressed by updating standard practices to be more aligned with Industry 4.0 or Smart Manufacturing practices. In this document, we intend to delve deeper into the challenges for NDE and discuss how to take it to the next generation: "NDE 4.0" or "Smart NDE." Industry challenges, implementation aspects, and ethical considerations are discussed while recognizing the important influences by and on infrastructure, people, equipment, and data. This philosophy has a significant work-flow and behavioral impact in the laboratory and even a manufacturing environment, especially for research and advanced technology organizations.**

I. Introduction

In the current digital age, with ubiquitous networking and access to large stores of data, not all aspects of industry have kept pace. Technological advances often precede advances in corresponding best practices. Competing notions exist between new and contemporary processes and technologies due largely to humans' resistance to adopting change. Much of this resistance likely arises from risk-adverseness: change invites the unknown. The time does however eventually arrive when past decisions, which arose from risk-adverse intent to avoid change, pose a greater risk as existent technology and processes become obsolete. For example, the emergence of cybersecurity threats accompanied the rapid growth of the internet [1].

In the context of manufacturing, the challenges present in the standard statistical process control approach are outlined in [2] and summarized in Figure 1. Process monitoring for quality (PMQ) [3] proposes to address these problems. PMQ assumes NDE to be an evolved discipline, as indicated by (d). In reality, NDE is burdened by "conventional wisdom," which dictates that an operator acquires, transfers, and interprets data for NDE purposes. This is cumbersome, much of the time infeasible, and renders the desired 100% inspection impossible for many applications. Data acquisition challenges are non-trivial. For example, contact probe ultrasonic measurements are sensitive to probe location, coupling, and pressure. Measurement variation could be minimized by deploying well trained, dedicated personnel. The seemingly innocuous task of data transfer, especially for large volumes, involves judiciously written algorithms which compile adequate information, schedule for network load and latency, and manage downtime events.

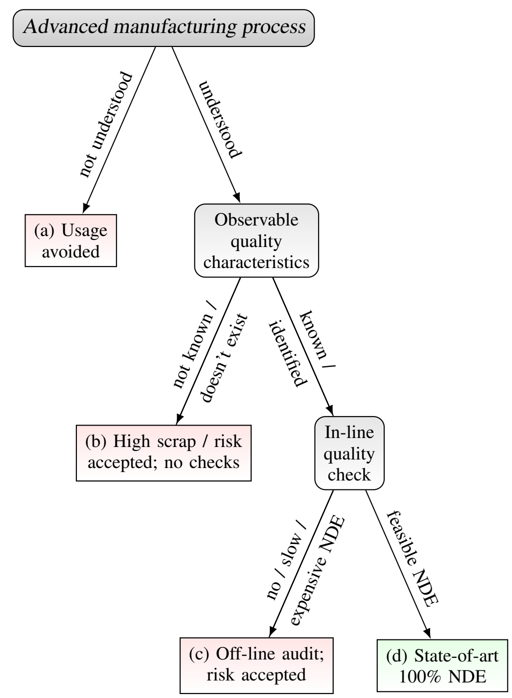

Fig. 1. Condensed view of statistical process control challenges from [2].

Then there is the aspect of communicating and archiving data in a widely understood, yet compressed, format. If data interpretation is solely left to the operator, the outcome risks being subjective based on operator experience, subject matter expertise, and condition (fatigue, ergonomics, etc.). It may therefore be prudent to employ a dedicated analysis team for data analysis and interpretation, provided adequate information regarding the NDE has been communicated.

In short, current NDE practices can hinder productivity by requiring the operator to be involved in all aspects of the data collection, transfer, and analysis. In this document, we intend to acknowledge such NDE challenges and provide an implementation framework that takes NDE to the next generation of NDE 4.0. While NDE 4.0, formally defined in Section II, is naturally integrated with the PMQ philosophy, it has a significant work-flow and behavioral impact in the laboratory environment, especially for research and advanced technology organizations. The challenges NDE 4.0 faces in the industry are discussed in Section III with potential

978-1-5386-8357-6/19/$31.00 © 2019 IEEE

978-1-5386-8357-6/19/$31.00 ©2019 IEEE

implementation framework in Section IV. Ethical concerns that may arise from this collaborative and co-depending environment are addressed in Section V, with conclusions in Section VI.

II. NDE 4.0: Smart NDE

Traditional NDE refers to any testing, inspection, or evaluation process which does not destroy the serviceability of the tested material, component, or system. This term is often used interchangeably with NDT, NDI, or NDT&E for testing, inspection, and testing and evaluation, respectively [4]. In this manuscript, NDE will be used to encompass all of the above.

A natural extension to traditional NDE is "NDE 4.0." This term attributes its origins to the combination of NDE and Industry 4.0 (to be described in more detail in II-A) and was first published in [5]. NDE 4.0 inherits facets of Industry 4.0 into NDE, and supports the Quality 4.0 [6] philosophy. The NDE 4.0 paradigm is still in its infancy; consequently, there is potential for it to acquire a name other than NDE 4.0. For example, "Smart NDE" (analogous to "Smart Manufacturing") may become the prevailing term. Here, we refer to NDE 4.0 and will describe the concept in the context of Industry 4.0.

A. Industry 4.0

To fully understand the concept of NDE 4.0, it is prudent to begin with a description of Industry 4.0 (aka smart manufacturing, manufacturing 4.0) and its origins. Industrial revolutions have occurred as groundbreaking technological innovations have triggered paradigm shifts in industry. Four such industrial revolutions are chronologically identified as follows: (1) mechanization, (2) mass production, (3) digitization and automation, and (4) cyber-physical systems and the Internet of Things (IoT) [5], [7]–[12].

The result of the fourth industrial revolution is termed Industry 4.0, or Smart Manufacturing. Presently, Industry 4.0 has not yet been fully realized and represents a future vision of industry. It addresses the need for manufacturing factories to cope with agile manufacturing [9]. Industry 4.0 addresses this need through IoT, interoperability, (semi-)autonomous decentralized decision making, industrial automation, and big data analytics.

B. NDE Revolutions

Shorter life cycles and smaller volumes necessitate 100% inspection be achieved by a new connected paradigm of NDE. The trajectory to NDE 4.0 draws many parallels with Industry 4.0 [5] as shown in Figure 2, and it has been argued that one cannot survive without the other [7]. Analogous to "industrial revolution," "NDE revolution" could be used to describe groundbreaking innovations which have triggered paradigm shifts in the context of NDE. The chronological NDE revolutions are: (1) human senses, (2) augmented human senses, (3) NDE, and (4) the connected lab.

Nondestructive evaluation began with the use of human senses to determine the quality of an object. One obvious example of this is visual inspection. Visual inspection, though

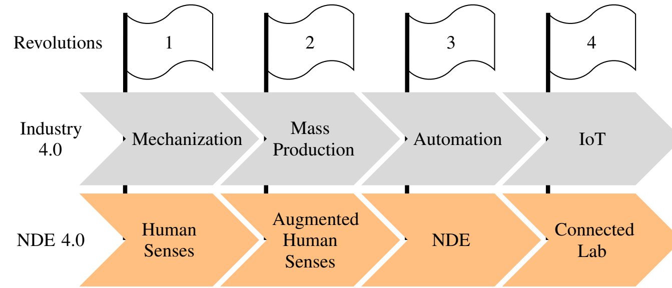

Fig. 2. Trajectory of the industrial revolutions and their conceptual parallels with NDE (derived from [5], [7]).

rudimentary, continues to be a powerful tool. For example, a highly corroded metal object visibly manifests its condition with a rust coating. While visual inspection may be the most obvious, other human senses have been used as well. For example, wooden telephone poles have historically been and continue to be inspected by impacting the pole with a hammer and listening to the resulting sound. A severely rotted pole will exhibit a different acoustic signature than a completely sound pole of the same wood type [13].

The second NDE revolution began with the use of tools to augment human senses. While it is tempting to assume that sensors fall within this period, it actually refers to basic tools like a stethoscope. The use of microscopes falls neatly in this category as they augment one's vision. X-ray machines would not fall within this period as they allow humans to perceive that which their senses, however augmented, would not.

The third NDE revolution began with the invention of the X-Ray technique in 1895 [14]. This period is suitably named NDE, as this is when the term was coined, and it aligns with the modern interpretation of NDE. Since its inception, NDE techniques have become prevalent across industries, academia, and government laboratories. Journals, textbooks, and societies are dedicated to advancing NDE technology and methods. The most commonly used NDE technologies are ultrasonic, phased array, x-ray, eddy current, and thermography. Although some techniques have been automated for deployment in a production environment [15], [16], many require experienced operators to collect and interpret the data [4].

C. The Fourth NDE Revolution: NDE 4.0

Though the parallels between Industry 4.0 and NDE 4.0 are obvious, their respective revolutions have been purely conceptual and heretofore not necessarily been aligned in time. The respective fourth revolutions, i.e. Industry 4.0 and NDE 4.0, is the first point at which the two generally align in time[1].

As shown in Figure 3, NDE 4.0 comprises four essential aspects: (1) NDE, (2) connected equipment, (3) network storage, and (4) big data analytics. The first aspect, NDE, refers to the modern sense of the word as described in Section II. The other

[1]Of course the start of Industry 4.0 slightly precedes NDE 4.0; however, since they are inextricably linked and neither has been fully realized, they are being said to generally align in time.

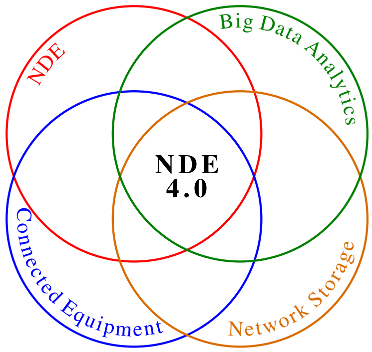

Fig. 3. The four essential aspects of NDE 4.0.

three aspects combine with the first to yield the "4.0" portion. Implementation of NDE 4.0 requires interconnectivity and networked storage. This IoT approach enables large amounts of data (big data) to be collected and stored. Big data analytics therefore encompass the fourth aspect of NDE 4.0.

Just as in Industry 4.0, data can be stored in a networked repository and accessed as needed by various applications. The same network enables machines to communicate among themselves and interface with humans, eliminating the burden of data transfer from the traditional operator. A networked data repository also enables task delegation among specialists for increased efficiency, experimental fidelity, and accuracy. For example, a trained technician focuses on the experiment while the software system logs all the observations which is then consumed by an analyst. In this case, comprehensive, structured experimental data logging, and sharing renders involvement of the analyst in the real experimentation redundant.

The fourth aspect is the incorporation of big data analytic techniques. Artificial intelligence (AI) methods have taken off in recent years due to algorithmic advances, increased compute power, access to large amounts of data, and IoT. These mathematical approaches can consume vasts amount of data and provide insight into otherwise non-intuitive solutions.

III. Industry Challenges

Similar to Industry 4.0, NDE 4.0 faces its adversaries.

A. Psychological Reservations

NDE 4.0 requires levels of understanding and cooperation from various entities. Everyone who has a stake in the experiment, including suppliers, plant workers[2], engineers, scientists, and management, should feel included in the project and able to share information comfortably amongst its participants. Unfortunately, this level of cooperation can be circumvented by the competitive notion that exists, arising possibly from a history of inadequate acknowledgment and credit sharing.

Psychological reservations naturally exist in regards to change. In the context of NDE 4.0, reservations include:

- fear of change — "Things work the way they are.";
- fear of downtime during implementation;
- misconception of intranet as internet;
- fear that network failure will lead to inoperable equipment;
- fear that the lab will become entangled in a mess of wires;
- problems that accompany networked computers, such as forced software updates;
- inability of data analyst to understand everything that occurred during data collection[3].

Notably, much of the reservations arise from negative past experience. It is our job to demonstrate that these fears are largely unfounded if NDE 4.0 is appropriately implemented.

B. Infrastructural Limitations

Connectivity, or lack thereof, is a challenge. Not all equipment is connected, including legacy equipment that was not designed to fit in the connected ecosystem. Also, connecting equipment translates to the presence of additional wires in the laboratory, which need to be properly managed.

In a manufacturing facility equipment connectivity is commonplace, and so is paper based operations logging. Paper logging, a traditional and psychologically deep-seated approach, inhibits NDE 4.0 from rendering its benefits.

C. Business Impediments

While the science of this initiative is very exciting, the business implications cannot be ignored. Requiring suppliers to provide more information might require a change in purchasing agreements. It can also impact the cost to cover expenses the suppliers might incur. To comply with the need for completely digital information, plant workers would need to abandon the convenient paper trail and take extra steps to log all activities digitally. Engineers and scientists will have to take extra steps to digitally create copious notes on the experiments. In other words, NDE 4.0 has a resource impact for which business units do not provision.

IV. Implementation

In the course of describing the implementation, we would be overcoming the challenges discussed in III.

A. The People Aspect

Arguably the biggest challenge posed to the adoption of the NDE 4.0 practices outlined in this document will arise from peoples' natural resistance to change. There are generally cogent, solid reasons in support of conventional practices; however, conventional practices are often rendered obsolete and thus incompatible with new technologies and processes. It is therefore important for said practices to be updated accordingly. Of course, the practices of NDE 4.0 are only implemented insofar as people adopt them. To implement NDE 4.0 on a large scale requires influencing the culture, which involves overcoming psychological reservations (see III-A). It therefore falls on us, the proponents of NDE 4.0, to convince others of its value. A strong motivator for this document was to serve

[2]Manufacturing processes are the richest data source for NDE.

[3]This is only true in the case that the data collector and analyst are different people.

as such a tool to advocate for and promote NDE 4.0 and to provide a concise, easy to follow, "quick-start" guide.

One tangible way to motivate people for NDE 4.0 is to locally pioneer only relevant aspects of its practices, i.e. not all at once or on a large scale. Incremental changes are generally more palatable than broad, sweeping changes. Successful cases will serve as a strong motivator to adopting NDE 4.0 practices on a larger scale. For example, setting up one networked data repository for a lab-space with a few connected test equipments and computers can illustrate to lab technicians and analysts how such a data storage method enables efficient data transfer.

B. Infrastructure

NDE 4.0 requires a fully connected lab/production environment to address the limitations addressed in III-B. Every equipment and computer participating in data acquisition and analysis must be connected by a private data intranet. The private data intranet ensures dedicated bandwidth essential to acquiring large volumes of data in real time. Wires can be concealed or appropriately routed. If a programmable logic controller network is needed, it is wise to keep that separate from the data network. Additional electronics and host computers may be necessary to provide connectivity to legacy equipment without ethernet options.

Business and corporation contributions are also very essential. Successful implementation hinges on cooperative information technology (IT) infrastructure, where IT serves in a role to support and enable this paradigm shift, while implementing adequate security and access policies. The IT needs are application specific and an unified policy might be inadequate. In addition, the business must perceive the value and make the investment in the infrastructure.

C. Information Organization

Archiving and organizing digital content has become ever more important. In [17], Ng recommended organizational changes in modern company that introduces new leadership roles around data organization and data warehousing. While we neither expect nor desire the oversight of chief data and analytics officer (CDAO) / chief artificial intelligence officer (CAIO) in day-to-day laboratory experimentation operations, we must train and implement adequate data management strategies to inhibit experimental data loss and improve the access and value of such data.

The first step to a successful data management strategy is to arrive at a common data organizational strategy that would work for all the participating entities. For the current discussion, we would refer to each NDE initiative as a project, that would have (a) project code, (b) version controlled repository for codes and documents, (c) project folder, (d) data folder, and (e) schema in a database.

1) Project Code: Each project should be associated with a project code, which hereon would be simply referred to as the code. For example, a project on leak test that started on the first quarter of 2018, could have the code `2018Q1_LKT`, or an ultrasonic welding project that started at the 4th quarter of 2011, could have the code `2011Q4_USW`. The following nomenclature template has worked for us:

```
<code> = <yyyy>Q<quarter#>_<3-5 letters>
```

Note that the code should not have any spaces or special characters besides underscore.

2) Version Controlled Repository: A version controlled repository could be created, one per project named with the code, for the codes and documentation folder discussed in IV-C3. GIT is a popular version controlling environment that is free to install and use.

3) Project Folder: Subsequently, when creating a folder for a project, the folder name should start with the code, followed by a descriptive name, `<code>_<descriptive name>`. It is bad practice to name folders with spaces or special characters, especially if those folders would participate in software development/execution. Alluding to the project examples discussed above, the folder names could be `2018Q1_LKT_leakTest` or `2011Q4_USW_ultrasonicWelding`, and so on. These names should be relatively short, because there is most often a 255 character limit on the file path. Here are some examples of folder names and their appropriateness:

- `2018Q1_LKT_leakTest` — most preferred, camel case (https://en.wikipedia.org/wiki/Camel_case).
- `2018Q1_LKT_LeakTest` — acceptable, Pascal case.
- `2018Q1_LKT_leak_Test` — discouraged, as it contains a redundant underscore
- `2018Q1_LKT_leak-Test.` — forbidden, must never be used as it contains special characters and spaces.

While the project owners can try to keep the folder names short, there is some amount of deadweight, like the server name, that is included in the file path. For instance, using a Windows® server, one would use a path like `\\<server>.<domain>\<folder>\`. The server name (along with the domain name) is the necessary inevitable deadweight. While we do not have control over the domain name, because it needs to be meaningful to humans [18], it is wise to keep the server names as short as possible. If a four digit number in base 62 is used[4] to name the server, there could be $62^4 = 14,776,336$ unique server names. One might argue that the long server name could be circumvented by mapping the folder as a drive letter (Windows®) or mounting to a local directory (Linux or macOS™), but then the codes written using such pseudo references become machine dependent unless the same mapping/mounting standards are followed by everyone.

The project can contain various types of documents and should be organized appropriately. Some commonly used folders include:

- `.git`, or equivalent, exists if a version control repository is used.
- `admin` (optional) has project management information.
- `code` for all codes. Subfolders can be assigned based on the type of code. For example, `m` for MATLAB files, `lv`

[4]Base 62 is the highest base number using the digits '0'–'9', the letters 'A'–'Z' and 'a'–'z'.

for LabVIEW™, `cpp` for C or C++, etc..

- `docs` to store documentation, including presentations.
- `figs` to store all generated plots and graphics.
- `results` to store any intermediate data / results that may be useful to generate plots and tables.
- `pubs` for published material under appropriate subfolders.
- `refs` for all reference materials including papers and articles of interest.

Figure 4(a) shows an example of the discussed folder structure.

```
2018Q1_LKT_leakTest
├─.git
├─admin
├─code
│ ├─lv
│ ├─m
│ └─cpp
├─docs
├─figs
├─results
├─pubs
│ ├─2019_PHM_IEEE
│ └─2019_internalReport
└─refs
```

(a) Project Folder.

```
2018Q1_LKT_rawdata_leakTest
├─20180111
│ ├─testA
│ └─testB
├─20180112
│ ├─testA
│ └─testB
├─20180113
│ ├─testA
│ ├─testB
│ └─testC
└─20180114
  ├─testC
  └─testD
```

(b) Data Folder.

Fig. 4. Example folder structure.

4) Data Folder: The reader may have noticed that there is no provision for data in the project folder structure. Data folders tend to be big, especially in the era of big data. Those are best managed separately from a size, security and information retention policy [19] standpoint. So, every project needs to have at least one data folder. It helps in documentation and maintaining commonality if the structure of the data folder follows a certain convention. In our experience, some of the traditional approaches of having unlimited levels of subfolders with folder names in hexadecimal characters are not only meaningless, but also can hang up the server if the directory listing is requested. It is best to have a limit on the levels of subfolders and the number of files/folders in a folder. Keeping this in mind, a strategy as illustrated in Figure 4(b) could become useful.

The main folder could be named with the project code in conjunction with the `data` or `rawdata`. An example naming template is `<code>_rawdata_<descriptive name>`. Then, there could be folders arranged by date using a `yyyymmdd` format. This enables a chronological sorting of folders even when sorting by name, and also limits the number of folder one may have in a finite duration of the project. The experiment names could form subfolders inside the date folder, which would then contain the data from the experiments.

Using commercial data acquisition software often locks the user into their proprietary binary data files. While these are efficient in storage, they require licensed software or application interface to access them. A convenient way to store data then obviously becomes files that do not have a proprietary format. Using the widely used "CSV" or "Excel" is not an option for large volumes of data. A binary file is much preferred. Fortunately, there are options like ".mat" and hierarchical data format (HDF) files [20] that could be used for time series and multidimensional data. For images, videos, and audio there are already widely used standards.

Another commonly used option for saving data, specifically time series data, is variations of Historian databases [21], [22]. While various articles online discuss the pros and cons of these databases [23], the big shortcoming from our perspective is that the historians are proprietary and involve licensing costs. In addition, historians are not relational databases. While historian databases have their advantages, we do not recommend using any proprietary standards for storing data. One might consider using CLOB and BLOB in relational databases to store data but, as discussed in IV-C5, that is not a good option.

5) Schema in Database: A relational database brings all the project resources together in conjunction with copious details on the experiment. We start with the schema name that contains the code. Schema names may need to begin with a letter, in which case name of the group, a cost center, etc., are good choices, e.g., `MSR2018Q1_LKT` is a good schema name. A project schema should contain a list of commonly used essential tables with somewhat meaningful names as listed below.

- The primary table is the `METADATA` table that holds all essential information regarding the experiment and the generated data. It typically contains a column corresponding to each metadata. The essential column is a unique identifier named `U_ID`, which could be a serial number of a part, or simply an auto-incremented column. It is the unique key for this table only. Avoid using `INDEX` or `UID` since they could be keywords for your database.
- The next essential table is the `ATTACHMENT` table that links signal files stored in the data folder to a record in this schema. This table should contain the same `U_ID` column that acts as the link to the `METADATA` table. It is best to keep the name of the indexing column identical in all tables, unless there are some special needs. This table should contain the following columns.
  - `U_ID`, not a unique key.
  - `RELPATH`, contains relative path to a file location.
  - `FILENAME`, contains the name of the file which contains multidimensional data.
  - `DESCRIPTION`, a useful column that describes what kind of data the file contains, since there could be multiple data files associated to one `U_ID`.

  This table replaces the need to store data directly in the database as `CLOB` or `BLOB`. Doing so would make the database very large, sluggish and hard to manage without a very complex and expensive IT infrastructure.
- The `BASEPATH` table can save base-paths to the data files archives. It could have three columns:
  - `BASEPATH` or simply `PTH` could be used to store

the base path,
  - DATE_START would store the date of experiment/data generation when this base-path was put to use, and
  - DATE_END would mark the last date any data was stored in that base path. By knowing the date of the experiment from METADATA, one could easily determine which base-path should be chosen. An empty value in this column would indicate that the corresponding path is currently in use.
- At times, certain variables/strings need defining in various tables, and one could create a LUT_DICT, a dictionary look up table. Columns TABLENAME, COL, VAL, and DESCRIPTION could be used to identify the table in which a certain column exists, and what a certain value in that column means. For example, a TABLENAME = "METADATA" and COL = "QUALITY" could contain VAL = 0 corresponding to DESCRIPTION = "Good". This kind of a dictionary table explains all the numeric/mnemonic codes that could be used in the all the tables in the schema.
- All the errors used in the coding and the process should be stored in LUT_ERROR with self-explanatory columns ERRCODE and DESCRIPTION.

We cannot stress enough the need for documentation, even in the case of a database. We highly encourage succinct and complete comments for every table and column in each schema.

Oracle has been our database of choice in this matter. The user could use any database, though we strongly suggest not to use a “pseudo” database like MS Access or SQLlite, for their lack of capabilities, security, and access management. MySQL would be recommended if open source database is required. The choice of database often goes beyond the financial aspect. Some of the databases today provide in-database analysis capabilities. For instance, Oracle offers Oracle R Enterprise [24], which could be a great advantage for R users. Databases like Oracle and MS SQL provide other in-memory analytics applications which could be advantageous.

D. Database Access Levels

Just having a database does not solve the problem. One would either need to have the skills of a database administrator (DBA) or employ one. The DBA undertakes responsibilities like backing up the database, tuning it for optimal performance, etc. Among them, the topic of relevance here is permission settings.

There are four levels of access that would be required to run this system.

1) *Database administrator* (DBA) is responsible for maintaining the entire database, performing backups, and keeping the system running. The DBA would be expected to create schemas on request, or restore tables on request but is not expected to do any data handling directly. A DBA is typically employed by the IT organization.
2) *Project administrator* has all the rights inside a schema, and could directly establish a connection to the project schema. The project administrator should be the one to create and maintain tables, set constraints, permissions to the tables, and so on. The project administrator should be able to access only the schema(s) he/she is designated to maintain. The project administrator is typically a researcher/engineer who owns the project or one who is the designated data manager for the project.
3) *Super user* is the one who has read/write access to tables from one or more schemas, depending on how many projects he/she is working on. The super user establishes connection using his/her private schema, to which one or more table collections have been shared by one or more project administrators. The super user should be able to insert values in tables and maintain the content of it, but should not be able to create/modify the tables. The schema name for a super user could be a unique user ID of him/her.
4) *User* is the one with the least access. He/she has read-only access to one or more table collections from one or more schemas, depending on how many projects he/she is assigned to. In an open data sharing model, the user could have access to tables from all schemas. The user would establish a connection using his/her own schema to which tables from various schemas have been shared. The schema name for a user could be the unique user ID assigned to him/her. If a user is also the super user for some schemas, write access to those tables should be granted selectively. In other words, an individual human being should only have one schema created based on the unique user ID and must have access to the various schemas based on the responsibilities. Typically, every consumer of the data in the database should have their own schema with their unique user ID.

A person who does not have the necessary access to issue a structured query language query directly to a schema, and can instead access data only through a precompiled application interface, would be called a *customer*. We are customers when we use a web interface to read bank statements online.

E. Documentation

Documenting at every stage is of paramount importance, even when only one person is involved in a project. Adequate documentation enables easy technology hand over, facilitates collaboration, and prevents data loss from an organization standpoint in case the person working on the project moves on. The following enlists the some crucial documents, without which the project may cease to be valuable.

- All aspects of the experiment(s) must be maintained in one or more METADATA tables.
- A short narration explaining the goals of the experiment is important and can double as an executive summary.
- The codes should be documented *in situ* alongwith a separate document describing the architecture, especially if there are dependencies. Code documentation should allude to the mathematical concept being programmed.

- A narration explaining the various data files should also be provided, explaining what each file contains and how the data inside them are organized.
- A narration explaining any relevant information on any practical deviations from a prescribed experiment will aid the data analyst in properly interpreting the data, since no experiment is ever conducted under truly ideal conditions.

A good version control tool is highly recommended, for codes and documents. Recently GIT (https://git-scm.com/) has gained popularity and for good reasons.

F. Business Imperatives

In order to address III-C, leadership must be supportive. Suppliers should provide connectivity options and can price accordingly. Engineering and purchasing should work in tandem to agree on purchasing agreements. There would be long term benefits at the expense of potentially higher initial investments.

V. Ethics

In a highly collaborative and connected NDE environment, sharing credit where credit is due and acknowledging the support from others are essential. Sharing appropriate credit fosters collaboration through trust and gratitude. In the modern NDE landscape, there are often multiple contributors who can be grouped as:

- experimentalist(s) — who designs and/or performs the experiments, including preparation and procurement of samples on which the experiments must be conducted;
- modeler(s) — who prepares models (mathematical, physical, chemical, etc.) necessary for the study;
- analyst(s) — who analyses and interpreting the outcome of the experimentally obtained data by utilizing the model(s);
- sponsor(s) — who provides funding, and/or materials and equipment (provided the person/entity has not been compensated for such resources);
- project manager — who sets funds and deadlines judiciously, and ensures that these constraints are met;
- programmer(s)/software developer(s) — who provides coding support;
- editor/typesetter — who edits and typesets manuscript(s) and manages the communication;
- data owner — who legally owns the data;

though one individual could assume one or more of these roles.

This raises some inevitable questions regarding due-credit. Do all these people deserve authorship of the resulting NDE documentation? Who earns the inventorship if there is any intellectual property (patent(s), defensive publication(s), or trade secret(s))?

John Hopkins Medicine clarified some of this in a web note [25]. The following excerpt adequately defines the distinction between authorship and inventorship.

> "The basic difference between inventorship and authorship is that inventorship is legally defined and authorship is subjectively decided. Authors are added to research publications sometimes out of professional courtesy or deference with no fear of invalidation of the science presented. If the inventors listed on a patent are not correct, or left off, the patent can be deemed invalid; a very serious consequence resulting in loss of intellectual property rights."

The United States Patent and Trademark Office (USPTO, https://www.uspto.gov) has published inventorship guidelines in [26] and improper naming of inventors in [27].

Authorship on the other hand is not bound by such legalities and thus needs some clarification. Various communities (e.g. International Committee of Medical Journal Editors (ICMJE), and Institute of Electrical and Electronics Engineers (IEEE)) require the author(s) to have done all of the following [28]–[30]:

1) made a significant intellectual contribution to the theoretical development, system or experimental design, prototype development, and/or the analysis and interpretation of data associated with the work;
2) contributed to drafting the article or reviewing and/or revising it for intellectual content;
3) approved the final version of the article as accepted for publication, including references; and
4) agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Based on these guidelines, to be an NDE document author:

- the modeler must provide the model(s) along with a writeup explaining it;
- the experimentalist must design and/or conduct the experiments to gather data, and contribute to writing the resulting documentation;
- the analyst must analyze the data and note outcome(s).

Of course, one author could play one or more of these roles. Ghost authorship and honorary authorship [28] are unethical, and must be strongly discouraged. It is also worth noting that the primary author (when authors are not listed alphabetically) is the one who prepares most of the document and/or is the primary contributor to the intellectual content. It is recommended that the authors discuss this matter before starting to draft a publishable document.

In the unique scenario of NDE 4.0 where the analyst obtains data from the experimentalist (with data owners' permission), the experimentalist should be requested to provide a writeup about the experiment in order to be a co-author in the document. In the event the experimentalist requests an analyst to interpret data, the analyst must provide a detailed analysis that gets included in the document and the analyst must be named a co-author. Similarly, if the modeler is providing model(s) he/she must provide an explanation to be included in the document in order to be named a co-author. Some papers are experiment-focused, some are modeling-focused, and others are analysis-focused. The primary author should be the lead person (decided ahead of time) in the primary focus area of the paper. Typically a lead experimentalist designs, supervises, and verifies the accuracy of the experiment; the lead analyst provides the mathematical foundation and verifies the outcome; and the lead modeler lays out the parameters and design of the model. If there are any individual(s) whose only role is to (a) prepare a software interface, (b) write a code for analysis/modeling, (c) perform typesetting, (d) perform editorial corrections, (e) maintain legal ownership of the data, (f) acquire data, (g) provide technical consultation, (h) provide experimentation/technical support, (i) manage, or (j) sponsor, they should be acknowledged.

VI. Conclusion

A new envisioned generation of NDE, termed NDE 4.0 or Smart NDE, was presented and discussed. Conventional NDE practices are antiquated, cumbersome, and inefficient, though not from want of technological advances. To keep pace with Industry 4.0 practices, a paradigm shift must occur. The various considerations which accompany any paradigm shift were considered and discussed, including implementation, resistance, and ethics. Thought was given to the impact this paradigm shift would have on infrastructure, people, equipment, documentation, and data. This philosophy would have significant work-flow and behavioral impact in the NDE lab environment and would be an excellent return on investments for businesses. While NDE 4.0 mitigates many challenges of its four components, it also inherits the unaddressed issues. For example, cybersecurity and network infrastructure burdens are inherited from "Connected Equipment" and "Network Storage," respectively. Having said all, we acknowledge that there could be scenarios that we have not envisioned despite the due diligence. The reader is therefore encouraged to provide feedback to the authors stating how this has or has not worked.

Acknowledgments

The authors would like to thank Dr. Teresa Rinker, Dr. Ryan Sekol, Dr. Amberlee Haselhuhn, Dr. Michael Wincek, and Mark Noto for their insightful discussions on data collection and interpretation. The authors would also like to thank Dr. Jeffrey Abell and Dr. Gary Smyth for supporting this initiative.

References

[1] J. Jang-Jaccard and S. Nepal, "A survey of emerging threats in cybersecurity," *Journal of Computer and System Sciences*, vol. 80, no. 5, pp. 973 – 993, 2014, special Issue on Dependable and Secure Computing. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022000014000178

[2] C. A. Escobar, M. A. Wincek, D. Chakraborty, and R. Morales-Menendez, "Process-monitoring-for-quality — applications," *Manufacturing Letters*, vol. 16, pp. 14 – 17, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213846318300166

[3] J. A. Abell, D. Chakraborty, C. A. Escobar, K. H. Im, D. M. Wegner, and M. A. Wincek, "Big data driven manufacturing — process-monitoring-for-quality philosophy," *Journal of Manufacturing Science and Engineering*, vol. 139, no. 10, pp. 101 009–1 – 101 009–12, October 2017.

[4] D. E. Bray and R. K. Stanley, *Nondestructive evaluation: a tool in design, manufacturing and service*. CRC Press, 1998.

[5] N. G. Meyendorf, L. J. Bond, J. Curtis-Beard, and S. Heilmann, "Nde 4.0 — nde for the 21st century — the internet of things and cyber physical systems will revolutionize nde," in *15th Asia Pacific Conference for Non-Destructive Testing (APCNDT2017), Singapore*, 2017.

[6] D. Jacob, "Fresh thinking for quality in the digital era," in *Quality Digest*, Jul. 2017. [Online]. Available: https://www.qualitydigest.com/inside/management-article/quality-40-072417.html

[7] N. G. Meyendorf, "Nde for the 21st century: industry 4.0 requires nde 4.0 (conference presentation)," pp. 10 171 – 10 171 – 1, 2017. [Online]. Available: https://doi.org/10.1117/12.2263326

[8] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18 – 23, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S221384631400025X

[9] M. Brettel, N. Friederichsen, M. Keller, and M. Rosenberg, "How virtualization, decentralization and network building change the manufacturing landscape: An industry 4.0 perspective," *International Journal of Information and Communication Engineering*, vol. 8, no. 1, pp. 37 – 44, 2014.

[10] N. Jazdi, "Cyber physical systems in the context of industry 4.0," in *2014 IEEE International Conference on Automation, Quality and Testing, Robotics*, May 2014, pp. 1–4.

[11] D. Ivanov, A. Dolgui, B. Sokolov, F. Werner, and M. Ivanova, "A dynamic model and an algorithm for short-term supply chain scheduling in the smart factory industry 4.0," *International Journal of Production Research*, vol. 54, no. 2, pp. 386–402, 2016. [Online]. Available: https://doi.org/10.1080/00207543.2014.999958

[12] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, "Industry 4.0," *Business & Information Systems Engineering*, vol. 6, no. 4, pp. 239–242, Aug 2014. [Online]. Available: https://doi.org/10.1007/s12599-014-0334-4

[13] "Methods of Inspecting and Testing Wood Poles," Occupational Safety and Health Standard, OSHA Standard 1910.269(D), 1910.

[14] M. Lederman, "The early history of radiotherapy: 18951939," *International Journal of Radiation Oncology, Biology, Physics*, vol. 7, no. 5, pp. 639 – 648, 1981.

[15] C. N. MacLeod, S. G. Pierce, M. Morozov, R. Summan, G. Dobie, P. McCubbin, C. McCubbin, S. Dearie, and G. Munro, "Automated metrology and nde measurements for increased throughput in aerospace component manufacture," *AIP Conference Proceedings*, vol. 1650, no. 1, pp. 1958–1966, 2015. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.4914823

[16] F. Shafait, S. M. Imran, and S. Klette-Matzat, "Fault detection and localization in empty water bottles through machine vision," in *E-Tech 2004*, July 2004, pp. 30–34.

[17] A. Ng, "Ai transformation playbook how to lead your company into the ai era dec 13, 2018," Online, Dec. 2018. [Online]. Available: https://landing.ai/ai-transformation-playbook/

[18] A. Rowland, "10 tips for choosing the perfect domain name," Nov. 2018. [Online]. Available: https://www.godaddy.com/garage/10-tips-for-choosing-the-perfect-domain-name/

[19] D. Dunkelberger, "5 key steps to developing a solid data retention policy," Apr. 2018. [Online]. Available: https://www.ispartnersllc.com/blog/5-steps-developing-data-retention-policy/

[20] "What is HDF5®." [Online]. Available: https://www.hdfgroup.org/

[21] "Operational historian," Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Operational_historian

[22] "Data historian overview." [Online]. Available: http://www.automatedresults.com/PI/data-historian-overview.aspx

[23] G. Nunan, "Data historians vs time-series: which is better for data analysis?" Aug. 2017. [Online]. Available: https://www.nukon.com.au/blog/time-series-vs-data-historians-which-is-better-for-big-data-analysis

[24] "Oracle R Enterprise." [Online]. Available: https://www.oracle.com/technetwork/database/database-technologies/r/r-enterprise/overview/index.html

[25] Apr. 2005. [Online]. Available: https://www.hopkinsmedicine.org/webnotes/newsletter_draft/licensing/0504.cfm

[26] "2137 Pre-AIA 35 U.S.C. 102(f) [R-08.2017]." [Online]. Available: https://www.uspto.gov/web/offices/pac/mpep/s2137.html

[27] "2157 improper naming of inventors [R-07.2015]." [Online]. Available: https://www.uspto.gov/web/offices/pac/mpep/s2157.html

[28] M. Panter, "Credit where credit is due: Best practices for authorship attribution." [Online]. Available: https://www.aje.com/dist/docs/Authorship_Attribution_EN.pdf

[29] "IEEE publication services and products board operations manual 2019," Nov. 2018. [Online]. Available: https://pspb.ieee.org/images/files/files/opsmanual.pdf

[30] "Defining the role of authors and contributors." [Online]. Available: http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html