Enhancing Digital Calibration Certificate (DCC) Workflow: Python-based DCC Tools Development and Implementation

G M Geronymo

Instituto Nacional de Metrologia, Qualidade e Tecnologia, Duque de Caxias, Brazil

gmgeronymo@inmetro.gov.br

Abstract. This paper presents the development of a set of open-source software tools using Python programming language for the efficient handling of Digital Calibration Certificates (DCCs), called *DCC Tools*. The proposed tools not only generate DCCs compliant with the XML schema proposed by Physikalisch-Technische Bundesanstalt (PTB), but also offer the capability to seamlessly embed and extract the human-readable (PDF) version of the calibration certificate into/from the DCC. By incorporating this functionality, the tools provide a comprehensive solution for managing both machine-readable and human-readable versions of calibration certificates within a unified framework. This contribution enhances the accessibility, convenience, and interoperability of DCCs, streamlining the calibration process for a wide range of applications.

1. Introduction

Calibration Certificates play a critical role in ensuring the accuracy of measurement instruments, which is essential for maintaining the quality of products and services. In today's digital era and Industry 4.0 environment, where automation and data-driven processes are becoming increasingly prevalent, the need for true digitalization (and not just digitization)[1] of calibration processes and associated documents is more crucial than ever. However, despite advancements in digitalization, calibration results are often linked to printed documents or manually transcribed into digital formats, slowing down the seamless integration with information systems and industrial automation. This raises important questions about the purpose of internet-connected equipment if calibration results remain in a paper format and how laboratories can demonstrate traceability through digital media [1].

To address these challenges, Digital Calibration Certificates (DCCs) have emerged as electronic documents utilizing the Extensible Markup Language (XML) and validated against specific schema

[1]Digitization refers to the process of converting physical information into digital formats, for example, converting the text from a printed paper to PDF document. On the other hand, digitalization refers to the use of digital technologies to improve the processes and create value for customers, for a example, by storing and distributing data in a machine-readable format that can be integrated into management and production systems in a automated way.

definitions (XSD). DCCs play a vital role in the digitalization of processes in various industries and research fields, particularly for the calibration and subsequent use of calibrated measuring instruments. However, the current practice of reporting calibration results to customers involves providing a formal document, typically in the form of a paper certificate or a digitally signed PDF/A. These formats, though useful for human interpretation, fall short of meeting the requirements of machine-readable data formats necessary for seamless integration with digitalized processes [1][2].

The inherent limitations of closed, non-machine-readable digital formats, such as PDF files, do not contribute to the true digitalization of calibration certificates. Analogue calibration data, whether on paper or as closed PDF documents, contribute minimally to improving the production process of a company. The manual conversion of analogue calibration data into digital formats specific to manufacturing plants is time-consuming and error-prone, hampering the efficiency and accuracy of subsequent processes.

In light of these challenges, this paper aims to address the need for efficient handling and integration of digital calibration certificates (DCCs) within modern calibration workflows. We present a set of open-source tools, called *DCC Tools*, implemented using the Python programming language, which not only generate DCCs compliant with the XSD schema proposed by Physikalisch-Technische Bundesanstalt (PTB) [3] but also enables the seamless embedding and extraction of human-readable (PDF) versions of the calibration certificates into/from the DCCs. By developing these tools, we aim to enhance accessibility, convenience, and interoperability, thereby streamlining calibration processes in line with the demands of digitalization, Industry 4.0, and improved data-driven decision-making.

The remainder of this paper is organized as follows. Section 2 introduces the concept of Digital Calibration Certificates (DCCs) and discusses their role in the digitalization of processes. Section 3 presents related work on tools for DCCs and highlights the limitations of existing practices. Section 4 describes the proposed Python-based *DCC Tools* for generating and handling DCCs. Finally, Section 5 concludes the paper and discusses potential future directions for research and development in this field.

2. Digital Calibration Certificates (DCCs)

The transition from analogue calibration certificates, which involved printing measurement results on paper, to digital calibration certificates (DCC) goes beyond a simple transfer of content (digitization). Converting paper-based certificates into formats like PDF merely facilitates human readability, whereas DCCs are designed to be machine-readable and interpretable (true digitalization). This is achieved through the adoption of a clearly defined data storage format. Unlike traditional methods, DCCs offer a more precise and efficient way to store a substantial amount of data.

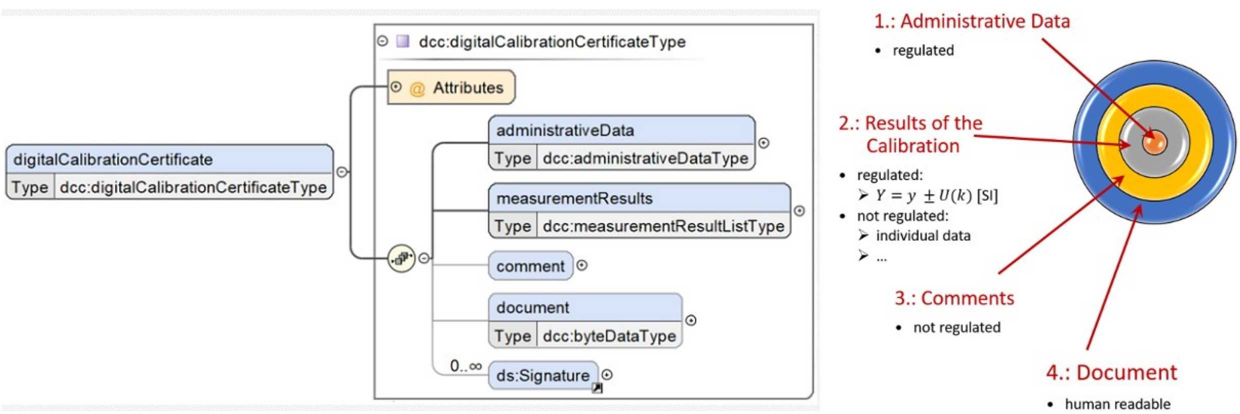

Figure 1. DCC structure (adapted from [5]).

Figure 1 shows the general structure of the DCC, divided into four areas: administrative data, measurement results, comments (optional) and human readable output (optional). The fifth element presented on the figure is a optional field for a digital cryptographic signature. The digital format chosen by PTB was Extensive Markup Language (XML), following the Calibration Data Exchange format proposed in the german standard VDI/VDE 2623 [6]. The structure of a XML file can be validated against a XML Schema Definition (XSD), and the schema proposed for the DCC is available on PTB DCC website [3].

2.1. Digital SI (D-SI)

To enable the seamless integration of measurement data into digital calibration certificates (DCC) using a machine-readable format, the development of a novel XML format called Digital System of Units (D-SI) [4] has been undertaken. The D-SI format is designed to incorporate crucial information such as numerical values for measurement results, units associated with those values, and comprehensive details regarding measurement uncertainties, including both the uncertainty value and its distribution. To further enhance the standardization and advancement of the D-SI format, ongoing efforts are being carried out by the "CIPM Task Group on the Digital SI" [7][8]. Figure 2 shows a visual representation of the D-SI data model and Figure 3 shows an application example for a resistance value in ohms.

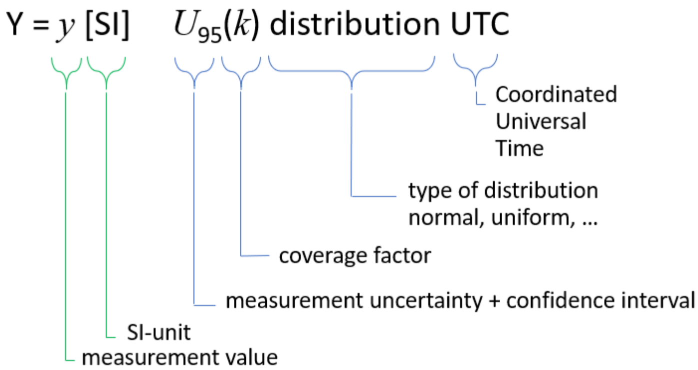

Figure 2. D-SI data model (adapted from [9]).

```
<si:quantity>
  <si:real>
    <si:value>0.99999397</si:value>
    <si:unit>\metre\tothe{2}\kilogram\second\tothe{-3}\ampere\tothe{2}</si:unit>
    <si:expandedUnc>
      <si:uncertainty>2.10E-07</si:uncertainty>
      <si:coverageFactor>2</si:coverageFactor>
      <si:coverageProbability> 0.9545 </si:coverageProbability>
    </si:expandedUnc>
    <si:dateTime>2023-02-09T13:37:37-03:00</si:dateTime>
  </si:real>
</si:quantity>
```

Figure 3. Example of a fully compliant D-SI quantity statement. The electrical resistance unit (ohm) is stated in term of the SI base units (m^2 kg s^{-3} A^2) for a "platinum" quality class in terms of machine-readability. Alternatively, the unit can be represented as "\ohm" for a "gold" quality class [4].

2.2. Digitalization of the calibration process

A proposal of a measurement workflow using DCCs is presented in [10], based on the European Metrology Cloud (MC) [11]. In this proposal the authors present Excel and Python tools to retrieve and save DCCs to the MC. Although a metrology cloud is the most desirable path for the digitalization of calibration processes, it is not strictly necessary. Even if DCCs are shared to customers by simpler methods, like by e-mail for example, it will be a huge gain in digitalization compared to the current practice of sending paper issued or PDF certificates. One of the main advantages in the use of DCCs is the possibility of real-time action on the instrumentation. In a smart factory there will be a large number of low-cost sensors that can benefit from the use of DCCs for real-time check of validity of measurement data [12]. The *DCC Tools* proposed on this paper can be used both in the initial stage of adoption of DCCs, before the establishment of a MC, and in future scenarios, integrated to the MC.

3. Existing solutions

PTB provides two sample tools in the DCC's project downloads page [5]: GEMIMEG Tool [13] and Excel Tool [14][15]. GEMIMEG Tool is open-source under EUPL license, and is developed using modern Javascript to run entirely on the user's web browser, with no data being sent to a third-party server. GEMIMEG tool enables a comfortable, manual creation of the administrative data in the DCC using several input masks [3]. The tool is provided as a sample application and is meant to be used mainly for educational and testing purposes. On the other hand, Excel Tool (also open-source under EUPL license) is a macro-enabled Excel spreadsheet that can generate the DCC XML from building blocks. It is an interesting approach since many laboratories use Excel spreadsheets in their calibration workflow, but it requires some manual work to transfer the calibration data to the Excel Tool worksheet.

Another proposed approach for DCC with machine readable data involves the inclusion of metadata (CSV, JSON or XML) inside a PDF/A-3 document [16]. This approach contrasts with PTB's DCC, where human readable version of the calibration certificate can be included inside the XML file as a base64 encoded string.

The proposed *DCC Tools* under development aim to achieve compliance with PTB's DCC schema, similar to the existing GEMIMEG Tool and Excel Tool. However, to better accommodate production systems, we believe that a cloud-based architecture is a good choice. This adoption of a cloud-based approach presents several advantages, including enhanced scalability, reliability, and accessibility. By harnessing the capabilities of cloud computing, the tools will be better equipped to handle the demands encountered within production environments.

The development of a Python library, PyDCC, is currently underway as part of the GEMIMEG-II project and is intended to be publicly available as an open-source software [1]. The primary objective of PyDCC is to enable effortless access to the contents of the DCC and facilitate their seamless processing. By doing so, PyDCC empowers users to leverage the inherent advantages of the DCC in a straightforward and user-friendly manner, ensuring its readiness for digital manufacturing. The PyDCC library simplifies the extraction and manipulation of DCC data, enabling users to effortlessly harness the benefits offered by digital manufacturing processes. This ability to easily extract measurement data information from the DCC represents a pivotal step towards the comprehensive digitalization of calibration processes. It is not hard to implement a simple software tool to access the contents of the DCC using Python, but since PyDCC is already under development and will provide a more complete solution, our *DCC tools* will leverage PyDCC once it is available for this essential functionality, ensuring a streamlined and efficient integration of measurement data into the calibration workflow.

4. DCC Tools

Our proposed *DCC Tools* package is freely available on GitHub [17] under GPL version 3 license. It was implemented using Python programming language and designed to run as a modern solution using cloud-based technologies such as AWS Lambda functions (but it can be easily adapted for local execution if desired). Up to now there are three main tools: *dccGenerator*, a tool to generate a PTB DCC compliant XML file, taking as input a JSON array (in the future it will be extended to accept XLSX / ODS spreadsheets for easier integration with current calibration systems), *dccAddPdf*, a tool to add a human readable PDF version of the calibration certificate to the XML file and *dccGetPdf*, a tool to extract the human readable PDF from the XML file. When PyDCC becomes available, we plan to implement a tool (*dccReader*) to extract measurement data from the DCC XML files and recover it as JSON arrays or XLSX / ODS spreadsheets. The GitHub repository contains also some CURL scripts to test the tools and some examples of DCCs. Up to now two tools are implemented based on examples of real calibration certificates of Inmetro's Electrical Standardization Metrology Laboratory (Lampe). Some work is needed to extend it to more general use cases. Figure 4 shows a diagram representation of *DCC Tools*.

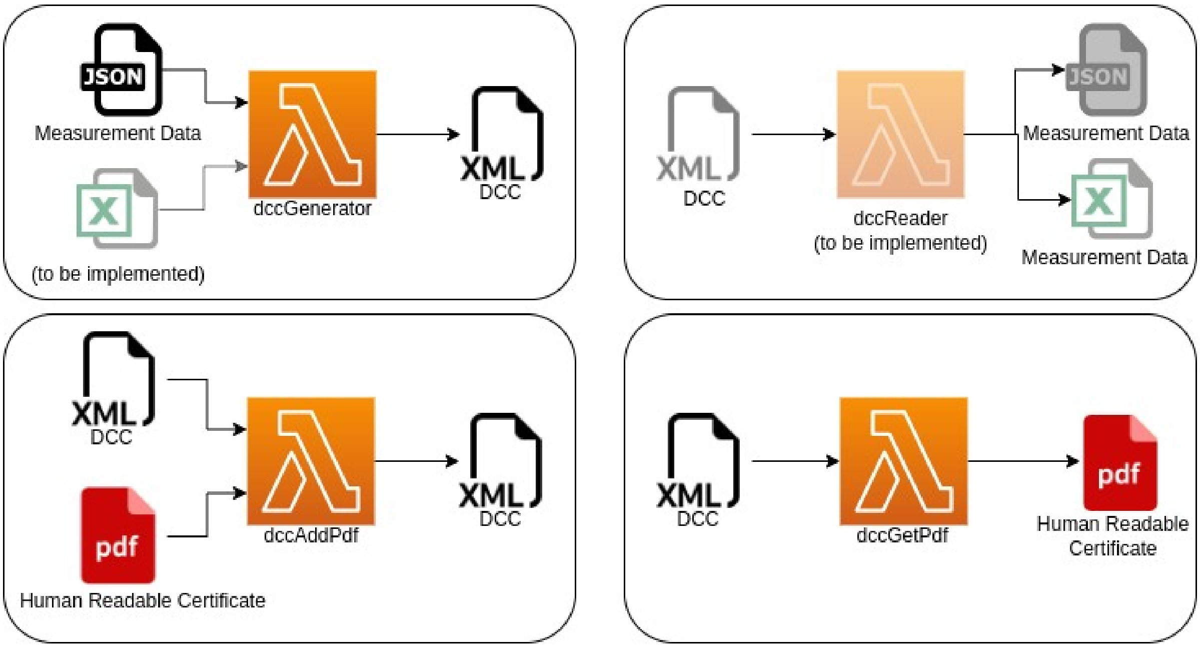

Figura 4. *DCC Tools* diagram.

4.1. System architecture.

The DCC Tools system architecture follows a serverless design paradigm, using AWS Lambda as the core compute service. This architecture provides several advantages, such as automatic scaling, pay-as-you-go pricing, and reduced operational overhead. In contrast to utilizing the AWS API Gateway as the entry point for external systems or users, for *DCC Tools* we have opted to keep with the simplicity and cost-effectiveness of AWS Lambda function URLs. By directly invoking the Lambda functions using their unique URLs, our application simplify the communication process and reduce unnecessary overhead associated with API Gateway configuration and management. This approach allows for a more lightweight and straightforward architecture, particularly suitable for scenarios where the system primarily interacts with a limited number of trusted clients or internal systems. The Lambda function URLs provide a direct and efficient means for data submission and retrieval, facilitating seamless integration with existing workflows and reducing the overall operational costs. Additionally, this approach aligns with the serverless nature of the *DCC Tools*, using the pay-as-you-go pricing model of Lambda functions without incurring additional expenses associated with API Gateway usage. Furthermore, for the typical usage of a calibration laboratory, the cost of operating the *DCC Tools* within the AWS Lambda free tier limits is expected. By staying within these free tier limits, calibration laboratories can use *DCC Tools* without incurring additional costs, making it a cost-efficient solution for digital calibration certificate generation. This further enhances the appeal and accessibility of the *DCC Tools* for calibration laboratories of various sizes and budgets.

4.2. Criptographic Signatures.

The DCC XML document can be signed digitally to add trust between the calibration institute and the customer. Adding digital signatures to XML files is a common practice in the software industry, and there are several tools available for this purpose, some of them free of charge. In Brazil we can use for free *Assinador Serpro* [18], developed by a federal state-owned company, but the signer must own a valid digital certificate issued by a certification authority.

4.3. Privacy and confidentiality concerns

The proposed architecture for *DCC Tools* does not include any kind of data storage on the cloud servers. The AWS Lambda functions are used just to process the data and return it to the user. Since the access to the lambda function is done by encrypted HTTPS requests, the confidentiality of the transmitted data is guaranteed. However, if the laboratory has concerns about confidentiality issues by using an external provider (like AWS) the functions can be easily adapted to run on a local server.

5. Conclusion

This paper introduces a pragmatic solution for the efficient handling of DCCs through the utilization of a cloud-based serverless application. The software tools presented in this study offer significant potential for enhancement and customization to cater to the specific requirements of calibration laboratories, seamlessly integrating into their calibration workflows. This integration contributes to a genuine digitalization of calibration processes. By adopting DCCs, calibration laboratories can experience substantial improvements in productivity by eliminating the laborious task of manual transcription of calibration data. This, in turn, mitigates the undesired repercussions arising from inevitable transcription errors, thereby fostering accuracy and efficiency in calibration operations.

References

[1] S. Hackel, S. Schönhals, L. Doering, T. Engel, and R. Baumfalk, 'The Digital Calibration Certificate (DCC) for an end-to-end digital quality infrastructure for industry 4.0', *Sci*, vol. 5, no. 1, p. 11, Mar. 2023.

[2] M. S. Gadelrab and R. A. Abouhogail, 'Towards a new generation of digital calibration certificate: Analysis and survey', *Measurement (Lond.)*, vol. 181, no. 109611, p. 109611, Aug. 2021.

[3] DCC_PTB. Available online: https://www.ptb.de/dcc (accessed on 19th June 2023).

[4] Hutzschenreuter, D., Lin, S., Loewe, J. H., Scheibner, A., Klobucar, R., Acko, B., Müller, B., & Heindorf, L. (2021). SmartCom Digital-SI (D-SI) XML exchange format for metrological data

version 2.0.0 (2.0.0). Zenodo. https://doi.org/10.5281/ZENODO.3366901

[5] DCC – Wiki – Downloads. Available online: https://dccwiki.ptb.de/en/downloads (accessed on 22th June 2023).

[6] VDI/VDE 2623, "Format for Data Exchange in Management of Measuring and Testing Equipment", ICS 03.120.10, 35.240.99 (February 2012).

[7] T. Mustapaa, J. Autiosalo, P. Nikander, J. E. Siegel, and R. Viitala, 'Digital metrology for the internet of things', in *2020 Global Internet of Things Summit (GIoTS)*, Dublin, Ireland, 2020.

[8] CIPM Task Group on the Digital SI. Available online: https://www.bipm.org/en/committees/ci/cipm/wg/cipm-tg-dsi (accessed on 26th June 2023).

[9] D-SI Online Guide. Available online: https://gitlab1.ptb.de/d-ptb/d-si/xsd-d-si/-/blob/master/wiki/d-si-online-guide.md (accessed on 26th June 2023).

[10] T. Bruns, J. Nordholz, D. Röske, and T. Schrader, 'A demonstrator for measurement workflows using digital calibration certificates (DCCs)', *Measur. Sens.*, vol. 18, no. 100208, p. 100208, Dec. 2021.

[11] Thiel, F., & Wetzlich, J. (2019). The European Metrology Cloud: Impact of European Regulations on Data Protection and the Free Flow of Non-Personal Data. In S. Gazal (Ed.), 19th International Congress of Metrology (CIM2019). EDP Sciences. https://doi.org/10.1051/metrology/201901001

[12] M. Marques, J. A. Sousa, and L. Ribeiro, 'Calibration 4.0 – Information system for usage of digital calibration certificates', in *19th International Congress of Metrology (CIM2019)*, Paris, France, 2019.

[13] GEMIMEG Tool. Available online: https://www.gemimeg.ptb.de/gemimeg-tool (accesed on 22th June 2023).

[14] Excel Tool. Available online: https://dccwiki.ptb.de/en/Tools/Excel-Tool (accessed on 22th June 2023).

[15] D. Röske, 'A visual tool for generating digital calibration certificates (DCCs) in Excel', *Measur. Sens.*, vol. 18, no. 100175, p. 100175, Dec. 2021.

[16] G. Boschung *et al.*, 'PDF/A-3 solution for digital calibration certificates', *Measur. Sens.*, vol. 18, no. 100282, p. 100282, Dec. 2021.

[17] DCC Tools. Available online: https://github.com/gmgeronymo/dcc_tools (accessed on 21th June 2023).

[18] Assinador Serpro. Available online: https://www.serpro.gov.br/links-fixos-superiores/assinador-digital/assinador-serpro (accessed on 28th June 2023).