Development of the NIST X-ray Photoelectron Spectroscopy (XPS) Database, Version 5

PRACTICE PAPER

ANGELA Y. LEE ⓘ

CEDRIC J. POWELL ⓘ

JUSTIN M. GORHAM ⓘ

ADAM MOREY ⓘ

JOHN HENRY J. SCOTT ⓘ

ROBERT J. HANISCH ⓘ

*Author affiliations can be found in the back matter of this article

]u[ubiquity press

ABSTRACT

It has been over 20 years since the National Institute of Standards and Technology (NIST) launched the first web version of its X-ray Photoelectron Spectroscopy (XPS) database (Lee et al., 2002; Wagner, 1991) which has approximately 1000 active users every day. This database was recently redesigned to meet NIST security requirements and to include new features and enhancements. The new application is built upon a relational database using a cross-platform and open-source framework (Henning et al., 2002). In addition to a modern design interface, the new features include custom-built components for displaying formatted molecular formulas and spectral lines, for sorting spectral lines, and for graphical display of chemical shifts of binding energies, Auger-electron kinetic energies, and Auger parameters for elements in different compounds.

CORRESPONDING AUTHOR:

Angela Y. Lee

Office of Data and Informatics, National Institute of Standards and Technology, Gaithersburg, MD 20899-8520, US

leea@nist.gov

1. INTRODUCTION

X-ray photoelectron spectroscopy (XPS) is one of the most powerful techniques for surface chemical analysis. Technologically, surface, and interface properties are crucial for the fabrication and performance of a wide range of advanced materials such as ceramics, composites, alloys, polymers, superconductors, diamond-like thin films, biomaterials, and nanomaterials as well as for semiconductor devices, optoelectronic materials, high-density magnetic-storage media, sensors, thin films, and coatings. In a typical XPS study, a specimen is irradiated with near-monochromatic X-rays and emitted photoelectrons and Auger electrons are detected with an electron energy analyzer (Jablonski et al., 2020). Measurements are made of core-electron binding energies and, for some applications, of Auger-electron kinetic energies. Elements (except for H and He) and their chemical states can be identified from the measured electron energies and of small shifts in these energies for elements in different chemical states. The information depth for XPS measurements is generally between 0.5 nm and 5 nm, depending on the sample material and the measurement conditions (Powell, 2020). Over 200,000 papers with XPS results have been published during the past 30 years.

The XPS Database, specifically SRD 20, is one of NIST's scientific and technical Standard Reference Databases (SRDs) which contain data that have been critically evaluated by experts in the relevant fields (Kaiser et al., 2018). The SRD 20 consists of over 33,000 records containing data for different parameters that are useful for interpreting XPS spectra. The original online version was created in 2000. We have recently developed a new version of SRD 20 with the latest web development technologies, and to address issues related to NIST security requirements and to incorporate new features.

In this paper, we report on the newly implemented NIST XPS Application (app). We first describe the design and show example screens of its features. Finally, we discuss the challenges we encountered and how they were overcome.

2. DISCUSSION

The NIST XPS app is required to run on a variety of platforms (cross-platform) including Windows, MacOS, and Linux. We decided to use a cross-platform, open-source, high-performance framework (Ofoegbu, 2023) with a fast development cycle. The unified .NET[1] developer platform has a single set of base class libraries, a software development kit, and tools for building modern and cloud-based applications. The database-driven NIST XPS app described here utilizes the latest framework .NET 7 Core (Materials and Methods: 1), Entity Framework Core, and the Blazor user interface component suite. A component-based development approach was adopted to implement the app for reusability, uniformity, and maintainability (Basha and Moiz, 2014). Modern continuous integration, continuous delivery, and continuous deployment (CI/CD) pipelines in GitLab[2] were used to build and deploy a .NET Docker image by automating the .NET build processes and to include transitive dependency[3] vulnerability scanning.[4]

2.1 ENTITY FRAMEWORK CORE

The Entity Framework (EF) Core is a framework for storing data that allows developers to work with data at a higher level of abstraction to eliminate the programming need for most of the data-access codes. The main advantage of the EF core is its ability to utilize built-in classes in the framework and store instances in different types of data stores. To be specific, the data model, a collection of classes in object-oriented programming (OOP), is the center of the application. The data model describes how the XPS data are stored along with the related logic for the application and data validation constraints. Object-to-data store mapping technology coined EF Code is a cross-platform data-access framework for .NET, while an object-relational mapper (ORM) is designed to handle data access. ORM applies a mapping definition to associate

---

1   Microsoft is a registered trademark of Microsoft Corporation, and Windows, SQL Server, .Net Framework are trademarks of Microsoft Corporation.

2   GitLab is an open-core DevSecOps product and a registered trademark of GitLab Inc./GitLab B.V.

3   Transitive dependency – Wikipedia.

4   .NET Handbook Best Practices/Vulnerable Dependencies (infinum.com).

columns in database tables with properties of the various Classes. There are two kinds of EF Code:

- Database-First: if there is an existing database, reverse engineering can be used to build the data model and context.
- Code-First: if a database does not exist, one must define a model by creating classes in C# and then use the EF Code to generate a new database that matches the structure and features of model classes.

As this was an update to the previous version (Lee et al., 2002), the Database-First approach was selected to build a data model. It is important to note that our data model comprises a series of classes with each one corresponding to a table in the relational database (Codd, 1970). All data model classes must interact with a built-in *database context class* (*DbContext)* in the EF Code. Database context class represents a session with the database, understands how to communicate with the database, and dynamically generates SQL (structured query language) statements by executing a language-integrated query (LINQ) to query and manipulate data.

2.2 USER INTERFACE

The .NET Core Blazor is the web framework of the .NET Core ecosystem for developing user-interfaces (UIs). For speedy development, the open-source Radzen[5] Blazor component library (Materials and Methods: 3) was chosen to build interactive web UIs. Figure 1 shows an overview of the Blazor Server architecture. SignalR is an open-source, real-time communication library that moves data back and forth between the web server and the client-side browser bidirectionally as well as manages real time UI and data updates. The Blazor components are rendered on the Blazor Server which can directly access the database with an EF Core.
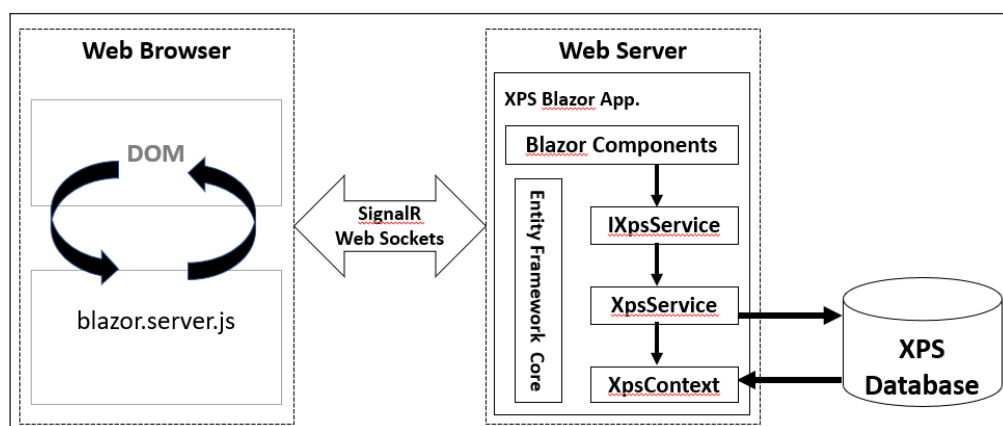

Figure 1 Overview of the Blazor server architecture.

To use Blazor in the application, three items were needed:

1. A reference to blazor.server.js,
2. Balzor services in the service container (Dependency injection, DI), where DI is 'a technique for accessing services configured in a central location'.
3. Blazor's SignalR hub

The database context, *XPSContext*, connects the XPS relational database directly between the Blazor components and database access. There is a small custom-built Data Service layer called *XpsService* which implements the Interface, *IXpsService*, so that we can reuse our developed data access codes for the entire SRD 20 development.

2.3 SEARCH STRATEGIES AND IMPLEMENTATIONS

The NIST XPS App has a modern UI design (see Supporting information (SI) Figure S1) with several new features that will now be briefly described. The first new feature is a custom-built data validator. If a user wished to identify an unknown line in a measured XPS spectrum, they

---

5   Radzen is a registered trademark of Radzen Company. Radzen Blazor Components are free and open source native Blazor UI components.

would click on the first option of the menu, select an energy type, and enter an observed energy and its estimated uncertainty, as shown in Figure 2. The observed energy is then checked with the custom-built data validator (Section 1 in the SI) to ensure that it is within the allowed data range for the selected energy type. If one or more matches are found, the relevant data records are listed that show the element, spectral line, chemical formula, and energy value as well as two hyperlinks (see Figure S2). One hyperlink (Details) provides metadata from the data record with information on the specimen material, the measurement conditions, and the literature citation (shown in Figure 3). The second link (Other Data) can be used to immediately find other data in the database from the same paper, as shown in Figure 4. A user can click on the column headers in the displays of Figures S2 and 4 to sort the information alphabetically (element, formula, and spectral line) or numerically (atomic number and energy).

Figure 2 User interface for identification of an unknown spectral line.

Figure 3 shows an example of the Details Summary screen that displays selected metadata from a chosen data record. This Detail screen has five tabs that provide information on the specimen material, citation details for the paper from which the data originated, data processing procedures, measurement conditions, and the specimen material. One new feature of the Citation tab is a link entitled 'All Records in this Publication' which, when clicked, pulls a list of compounds and other data in the database from the same publication, thus providing a convenient means for accessing related data, as shown in Figure 4. Figure S2, Figures 3 and 4, and Figure S2 have the same layout for displaying a list of compounds and associated data, thus providing an excellent example for our use of the same reusable and custom-built component.

Figure 3 Details summary screen for a specific entry from a publication. Additionally, 'All Records in this Publication' had been selected on a previous screen to provide all data from the selected publication. The DOI may be clicked which will direct the user to the journal homepage for the article.

The component-based software-development approach (Basha and Moiz, 2014) was applied throughout the application development. For example, digital object identifiers (DOIs) (Gahan, 2023) were added to most of the references where XPS data were obtained. A component was built to advise a user that they would then be leaving the NIST website. Clicking on the DOI link in Figure 3 (right) first opens an alert window to advise users that they would then be leaving the NIST website, as illustrated in the upper right of Figure 3. Users would then be directed to the external site recorded in the DOI for the reference of interest.

2.4 CHEMICAL SHIFT PLOTS

A useful new feature of the XPS App is the addition of chemical-shift plots. This type of plot displays the variation in measured binding energies, Auger-electron energies, or Auger parameters for a specified line representative of a selected element in different chemical states. We utilized the third-party Blazor Extensions library (Materials and Methods: 4) that contains the HTML5 Canvas graphic element for this purpose. We previously used a similar technology to display Wagner plots (Rumble et al., 1992; Wagner, 1977) in earlier versions of SRD 20. A new algorithm was designed and implemented (SI Section 6) to draw the chemical-shift plot using the HTML5 Canvas interface. After selecting chemical shift from the plot options on the home page, a user interface with a periodic table is displayed for selecting an element and energy type, as shown in Figure S3 (see SI). The user must next choose a specific line for their plot. The database then generates a list of materials that meet the selected requirements. Figure 5 provides an example screen showing the materials in the database that contain niobium and that have binding energy data for the $3d_{5/2}$ line. The user can then select up to 12 of these materials and click 'Graph' to produce the chemical shift plot shown in Figure 6 (left). If more than 20 compounds were identified in the search, a menu for further filtering compounds by selection of desired elements in the compounds is displayed in the right panel of Figure 5. The horizontal lines in Figure 6 (left) represent the range of reported binding energies for the Nb $3d_{5/2}$ photoelectron lines for each of the selected Nb compounds. The vertical lines indicate values of the median binding energy for each compound with the values presented below the vertical lines. A Summary Table is also produced (Figure 6 (right)) that lists the median binding energies for each compound and provides further information on the data source. A user could then click on a 'Details' link in the right most column of Figure 6 (right) to obtain information from that record similar to that shown in Figure 4.

🏠 / Scientific Citation

Kamada M., Ichikawa K., Aita O., *Phys. Rev. B 47, 3511* (1993) **DOI:** 10.1103/PhysRevB.47.3511

Instruction:
- *Click the column to sort in ascending or descending order.*
- *Enter a string in the input box in the column header and press the enter to filter the search result in the selected column.*
- *Click on the hyperlink in the column for more information.*

Total Records: 18

ELEME…	ATOMI…	FORMULA≡	SPECTRAL LINE	ENERGY (EV)	DETAILS
	⌃⌄				
Ba	56	$BaCl_2$	$N_{45}O_{23}V$	71.20	📄
Ba	56	$BaCl_2$	$N_{45}O_{23}O_{23}$	58.40	📄
Ba	56	$BaCl_2$	$N_{45}O_{23}O_{23}$	56.00	📄
Ba	56	$BaBr_2$	$N_{45}O_{23}O_{23}$	55.40	📄
Ba	56	$BaBr_2$	$N_{45}O_{23}V$	71.40	📄
Ba	56	$BaBr_2$	$N_{45}O_{23}O_{23}$	57.80	📄
Ba	56	$BaCl_2$	$5p_{3/2}$	11.40	📄
Ba	56	$BaCl_2$	$5p_{1/2}$	13.60	📄

Figure 4 Matches from a search for additional data from a selected citation.

In physical chemistry, molecular formulas and spectral data are traditionally presented with subscripts. Since this information is stored as simple text in the database, two new components for formatting displays of molecular formulas (SI Section 3) and spectral lines (SI Section 4)

were developed. This was accomplished by first introducing HTML5 Canvas markup tags and then processing them. The molecular formatter converts a molecular formula from plain text into a standard chemical formula such as those shown in Figure 6 (right) while Figure S4 (see SI) is an example of the orbital formatter converting orbital text into a physically meaningful format for display.

Figure 5 List of Nb-containing materials and compounds from the database with data for the Nb 3d5/2 orbital.

Ideally, there should be little dispersion of the measured binding energies for each compound. For the example chemical-state plot for Nb compounds in Figure 6 (left), however, the 12 reported binding energies for the Nb $3d_{5/2}$ photoelectron line of Nb_2O_5 range from 206.6 eV to 208.2 eV with a median value of 207.35 eV. It is difficult in retrospect to explain a spread of 1.6 eV in reported binding energies for the same compound. Possible explanations include drifts in the calibration status of each XPS instrument after its reported calibration. While there are now formal procedures for calibration of the binding-energy scales of XPS instruments (Powell, 2004), these were not available at the time that most data were acquired for the XPS App. Another possible explanation for the spread in reported binding energies for Nb_2O_5 is that the charge-neutralization procedures were inadequate for this nonconductive oxide. Again, improved procedures are now available for this purpose (Powell, 2004).

Figure 6 (left) Display of a binding-energy chemical-shift plot for Nb and selected Nb compounds for the Nb $3d_{5/2}$ photoelectron line. (right) Summary table for Nb and 11 Nb compounds used for the chemical-shift plot in Figure 5 (left).

FORMULA	SPECTRAL LINE	MEDIAN	DETAILS?
Nb	$3d_{5/2}$	202.23 eV	
Nb_2O_5	$3d_{5/2}$	207.35 eV	
K_2NbF_7	$3d_{5/2}$	209.40 eV	
$KNbO_3$	$3d_{5/2}$	206.50 eV	
$BaNbO_3$	$3d_{5/2}$	206.50 eV	
$Ca_2Nb_2O_7$	$3d_{5/2}$	206.70 eV	
$CaNb_2O_6$	$3d_{5/2}$	206.80 eV	
NbH	$3d_{5/2}$	203.20 eV	
Nb_3N_4	$3d_{5/2}$	204.00 eV	
$NbCl_5$	$3d_{5/2}$	208.00 eV	
$SnNbS_3$	$3d_{5/2}$	202.80 eV	
$NbTe_4$	$3d_{5/2}$	203.80 eV	
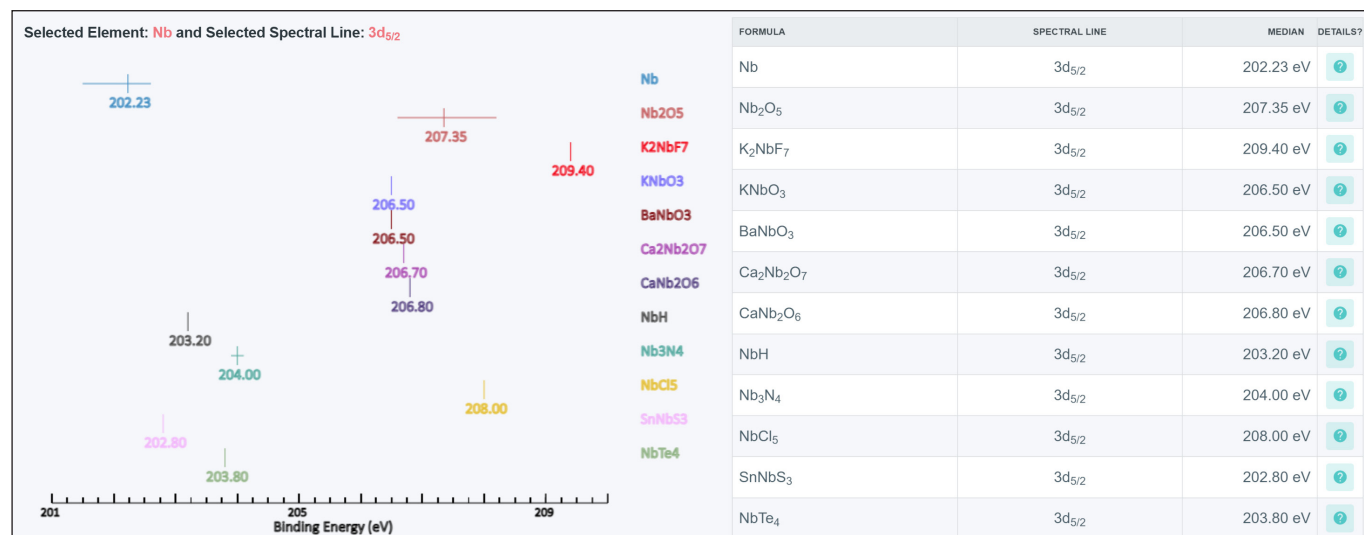
3. CONCLUSIONS

In summary, several challenging issues were encountered during the development of the XPS App. To address these issues, we developed the algorithms and implemented the custom components listed in Table 1. Further details are given in the SI.

NEW TO SRD 20 V 5.0		
FEATURE	**DESCRIPTION**	**LOCATION DESCRIBED**
Web Architecture	Cross-platform, open-source .NET framework, EF Core framework, Blazor Server frame	Section 2.2
Data Validator	Custom validators for checking users' input	Section 2.3, SI Section 1
Chemical Shift Plots	New algorithm was designed to draw the chemical shift plot in the HTML5 Canvas	Section 2.4, SI Section 6
Molecular Formatter	Converts a molecular formula from plain text into a standard chemical formula	Section 2.4, SI Section 3
Orbital Formatter	Converts orbital text into a physically meaningful format	Section 2.4, SI Section 4

- data validators for user inputs for selected spectral type and for each allowed energy range (SI Section 1).
- sorting functions to arrange spectral lines according to their electron orbitals instead of ordering alphabetically (SI Section 2),
- formatters to display molecular formulas (SI Section 4) and spectral lines (SI Section 5),
- graphical presentation of chemical-shift plots. Since there was no graphical package available for our purpose, we designed an algorithm to draw lines on an empty canvas using real-time data from the database (SI Section 6).

With open-source frameworks we successfully built a platform-independent, robust, and scalable Version 5 of the NIST XPS application (SRD 20) which can be deployed on different operating systems.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Supporting Information.** Figures S1 to S4 and Sections 1 to 6. DOI: https://doi.org/10.5334/dsj-2024-045.s1

Sections in the Supplementary Information

Section 1: Data validator

Section 2: Spectrum Sort Service

Section 3: Formatting Service

Section 4: Molecule Formatter

Section 5: Orbital Formatter

Section 6: Graphical Representation of Chemical-shift plots

ACKNOWLEDGEMENTS

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Angela Y. Lee ⓘ orcid.org/0009-0001-9174-2880
Office of Data and Informatics, National Institute of Standards and Technology, Gaithersburg, US

Cedric J. Powell ⓘ orcid.org/0000-0001-8990-2286
Nano Materials Research Group, National Institute of Standards and Technology, Gaithersburg, US

Justin M. Gorham ⓘ orcid.org/0000-0002-0569-297X
Nano Materials Research Group, National Institute of Standards and Technology, Gaithersburg, US

Adam Morey ⓘ orcid.org/0009-0008-9616-9017
Office of Data and Informatics, National Institute of Standards and Technology, Gaithersburg, US

John Henry J. Scott ⓘ orcid.org/0000-0002-3796-2110
Materials Measurement Science Division, National Institute of Standards and Technology, Gaithersburg, US

Robert J. Hanisch ⓘ orcid.org/0000-0002-6853-4602
Office of Data and Informatics, National Institute of Standards and Technology, Gaithersburg, US

REFERENCES

Basha, J.N.M. and **Moiz, S.A.** (2014) 'Component based software development: a state of art', *IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012)*, Nagapattinam, Tamil Nadu, India, March 30–31, 2012, arXiv, Cornell University, 14 June 2014. Available at: https://doi.org/10.48550/arXiv.1406.3728

Codd, E.F. (1970) 'A relational model of data for large, shared data banks,' *Communications of the ACM*, 13, pp. 377. Available at: https://doi.org/10.1145/362384.362685

Gahan, C. (2023). *What is a DOI? Finding and Using Digital Object Identifiers* Scribbr. Available at: https://www.scribbr.com/citing-sources/what-is-a-doi/ (Accessed: 14 May 2024).

Henning, H., Sebastian, H. and **Majchrzak, T.A.** (2002) 'Comparing Cross-platform Development Approaches for Mobile Applications', *International Conference on Web Information Systems and Technologies*. Porto, Portugal, 18–21 April. SciTePress, pp. 299–311. Available at: https://doi.org/10.5220/0003904502990311

Jablonski, A. and **Powell, C.J.** (2020) 'Effective attenuation lengths for different quantitative applications of X-ray photoelectron spectroscopy', *J. Phys. Chem. Ref. Data,* 49, 033102. Available at: https://doi.org/10.1063/5.0008576

Kaiser, D.L. and **Hanisch, R.J.** (2018) *Technical review of NIST's free standard reference data products* (NIST Special Publication NIST SP-1223). Available at: https://doi.org/10.6028/NIST.SP.1223 (Accessed: 3 September 2024).

Lee, A.Y., Blakeslee, D.M., Powell, C.J. and **Rumble, J.R.** (2002) 'Development of the web-based NIST X-ray Photoelectron Spectroscopy (XPS) Database', *Data Sci. J.*, 1, pp. 1–12. Available at: https://doi.org/10.2481/dsj.1.1

Ofoegbu, J. (2023) *15 Important considerations for choosing a web dev framework*. Available at: https://code.tutsplus.com/15-important-considerations-for-choosing-a-web-dev-framework--net-8035t. (Accessed: 29 May 2023).

Powell, C.J. (2004) 'Improvements in the Reliability of X-ray Photoelectron Spectroscopy for Surface Analysis', *J. Chemical Education*, 81, pp. 1734. Available at: https://doi.org/10.1021/ed081p1734

Powell, C.J. (2020) 'Practical guide for inelastic mean free paths, effective attenuation lengths, mean escape depths, and information depths in x-ray photoelectron spectroscopy', *J. Vac. Sci. Technol. A.,* 38(2), 023209. Available at: https://doi.org/10.1116/1.5141079

Rumble, J.R., Bickham, D.M. and **Powell, C.J.** (1992) 'The NIST X-Ray Photoelectron Spectroscopy Database', *Surf. Interface Anal.,* 19, pp. 241. Available at: https://doi.org/10.1002/sia.740190147

Wagner, C.D. (1977) 'A new approach to identifying chemical states comprising combined use of Auger and photoelectron lines', *J. Electron Spectroscopy and Related Phenomena*, 10, pp. 305. Available at: https://doi.org/10.1016/0368-2048(77)85028-7

Wagner, C.D. (1991) *The NIST X-Ray Photoelectron Spectroscopy (XPS) Database* (NIST Technical Note, 1289). Available at: https://doi.org/10.6028/NIST.TN.1289

MATERIALS AND METHODS

Frameworks and tooling

1. .NET 7 (.NET 7 Core) 7.0.14
2. Entity Framework Core 7.0.12
3. Radzen Blazor (https://blazor.radzen.com/) 4.19.0
4. Blazor Extensions Canvas 1.1.1

5. SQL Server 2016
6. Visual Studio 2022 v17.6.4 or better
7. SQL Server Management Studio (SSMS) v18.11.1
8. GitLab