A Combined Infrared Spectroscopy Database and Analysis Tool

Paul Jakob Jägerfeld, Hendrik Gossler, Johannes Riedel, Sofia Angeli*, Yuemin Wang, Sarah Bernart#, Jelena Jelic, Felix Studt and Olaf Deutschmann*

DOI: 10.1002/cite.202400150

The advances in density functional theory (DFT) made it possible to calculate the vibrational frequencies and peak intensities of adsorbates on extended heterogeneous catalyst surfaces. However, the peak broadening that naturally appears in experimental infrared spectra due to physical effects and instrumental limitations is usually not part of the quantum mechanical modeling method. Here, a new user-friendly application, CaRIn (Catalysis Research with Infrared Spectroscopy), within the CaRMeN platform is proposed, in which the peak width can be approximated by means of Gaussian functions. The peak broadening can be adjusted in real time and compared to both experimental and other simulated spectra. The application integrates functionality for better visual comparability of different spectra and a database of spectra to enhance workflow efficiency.

Keywords: Adsorption sites, Density functional theory, Heterogeneous catalysis, Infrared spectroscopy, Peak broadening

Received: November 14, 2024; *revised:* February 18, 2025; *accepted:* February 27, 2025

1 Introduction and Motivation

Infrared (IR) spectroscopy is one of the most important and powerful analytical methods in chemistry. Molecular vibrations of the sample under investigation are excited by interaction with electromagnetic radiation in the range of 2.5–25 µm (IR light). Quantum mechanical selection rules apply to the absorption processes. The selection rule states that the molecule only interacts with the electromagnetic wave if the dipole moment changes due to the excited molecular vibration. The specific vibrational bands of a compound are quantized and differ energetically due to the binding energy, making them characteristic of that compound. It is therefore possible to detect a compound qualitatively and, in some cases, quantitatively [1, 2].

An essential application of IR spectroscopy in heterogeneous catalysis is the analysis of the composition of gas-phase mixtures before, within, or after the catalyst to assess its catalytic activity [3–5]. For this purpose, transmission IR (TIR) is used, where a beam of IR light is passed through the gas-phase mixture to detect and quantify the components based on their characteristic absorption spectra. To gain deeper insights into the catalyst itself, diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) is employed. In this method, adsorbed molecules on the catalyst surface are examined in situ. The catalyst is characterized in a reaction atmosphere, which provides information about the intermediates formed on the surface. For example, for carbon monoxide (CO), one of the most prominent probe molecules, the C—O stretching vibrational mode can be used to determine the bond strength between adsorbate and metal (oxide), because the C—O bond strength is influenced through the dative bond to the surface (σ donation) and back donation into the antibonding $\pi*$-orbital [6]. In the case of CO adsorption on oxide surfaces, electrostatic interactions also play a role [7].

[1]Paul Jakob Jägerfeld https://orcid.org/0009-0000-7452-6230,
[1]Dr. Hendrik Gossler https://orcid.org/0000-0002-6863-0338,
[1]Johannes Riedel, [2]Dr. Sofia Angeli
https://orcid.org/0000-0001-8623-047X (sofia.angeli@kit.edu),
[3]Dr. Yuemin Wang https://orcid.org/0000-0002-9963-5473,
[2]Dr. Sarah Bernart https://orcid.org/0000-0001-5722-7634,
[2]Dr. Jelena Jelic https://orcid.org/0000-0002-2701-0765,
[2]Prof. Dr. Felix Studt https://orcid.org/0000-0001-6841-4232,
[1,2]Prof. Dr. Olaf Deutschmann
https://orcid.org/0000-0001-9211-7529 (deutschmann@kit.edu)
[1]Karlsruhe Institute of Technology (KIT), Institute for Chemical Technology and Polymer Chemistry (ITCP), Kaiserstraße 12, Karlsruhe 76131, Germany.
[2]Karlsruhe Institute of Technology (KIT), Institute of Catalysis Research and Technology (IKFT), Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen 76344, Germany.
[3]Karlsruhe Institute of Technology (KIT), Institute of Functional Interfaces (IFG), Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen 76344, Germany.
#Present address: Eindhoven University of Technology, Inorganic Materials & Catalysis, 5612 AZ Eindhoven, The Netherlands

By detecting reaction intermediates, it is possible to obtain information about reaction mechanisms. With the aid of isotopic substitution, further insights into mechanisms can be gained [8]. This is possible due to the high temporal resolution of the method. These investigations are mainly carried out on well-defined single-crystal surfaces [9–11]. However, substrate-adsorbate interactions vary depending on the crystal plane. Therefore, the properties of a real catalyst are often assumed to be a superposition of each of the different planes, even if this assumption is invalid in certain cases [12]. In contrast to the extensive IR reflection absorption spectroscopy (IRRAS) studies on metals [13], the application of IR spectroscopy to oxide single-crystal surfaces is significantly hindered by intrinsic experimental challenges posed by dielectric substrates. Recently, some of us have reported IRRAS investigations on various metal oxide surfaces [14–16], utilizing an advanced ultrahigh-vacuum system specially designed for IR experiments on both single crystals and nanoparticles [7, 17]. Reliable reference data, obtained from well-characterized oxide surfaces through the surface science approach that combines experiment and theory [18–21], are essential for gaining profound insights into the surface structure and chemistry of both pristine and metal-deposited oxide catalysts.

Experimental IR techniques can be complemented by computational methods such as the density functional theory (DFT) to elucidate molecular mechanisms, investigate adsorbate-catalyst interactions and simulate IR spectra. DFT calculations are performed in order to find stable minima of adsorbates such as CO on, e.g., transition metal surfaces, which are characterized by forces on each atom that are close to zero [22–24]. Vibrational frequencies of the adsorbate on the catalyst surface can then be calculated easily by (partial) Hessians, usually keeping the atoms of the surface fixed at their optimized positions within the harmonic-oscillator approach. After obtaining the vibrational frequencies and normal modes of the system, the simulated IR spectra can be generated [25, 26]. However, experimental spectra exhibit broader peaks whereas theory gives one specific value. This is due to a combination of instrumental limitations (instrumental broadening) and physical phenomena. First and foremost, these include natural broadening, due to the uncertainty principle, which defines the minimum possible line width in the shape of a Lorentzian. Further effects include Doppler broadening, caused by the Maxwell distribution of relative velocities between atoms/molecules and the observer, resulting in a Gaussian line shape, as well as surface heterogeneity and adsorbate-adsorbate interactions (proximity broadening). Overlapping due to broadening can result in obscured peaks, shoulder features, and shifts in peak maxima. As a result, the visual appearance of a DFT-simulated spectrum can considerably differ from that of an experimentally obtained spectrum, making comparisons difficult [1, 27, 28].

A common approach to improving the comparability of simulated and experimental spectra is to artificially broaden the calculated peaks by convoluting the calculated spectra with Gaussian functions. The width of these Gaussian functions is typically expressed as the half-width at half-maximum (HWHM). In the case of the Gaussian\16 software, a state-of-the-art program for electronic structure modeling, the HWHM is set to a constant value of 135.00 cm^{-1} and cannot be adjusted [29].

However, one significant challenge in this field is the difficulty of finding appropriate experimental spectra for the specific systems being studied. The complexity of heterogeneous catalysis often results in experimental data that are either scarce or not directly comparable to the systems modeled in simulations. Therefore, having access to large amounts of simulated and experimental spectra would be of great benefit to the scientific community. Not only the quantity but also the quality of the data is of great importance. To utilize the benefits of research data management (RDM), the FAIR (findability, accessibility, interoperability, and reusability) principles suggested by Wilkinson et al. [30] should be followed. This can be achieved by providing comprehensive and standardized metadata. RDM offers multiple advantages, including improved data integrity, quality, preservation, and security [31]. Tools such as Adacta place an additional emphasis on traceability, to ensure reusability of data already produced [32]. The CaRMeN platform provides a database for experimental data along with their metadata and enables a direct comparison with simulation data in an automated fashion [33–36]. With the Novel Materials Discovery (NOMAD) project, an extensive database and analysis tool for computational data in material science is freely available [37]. IR spectra for organic compounds are publicly available through the Spectral Database for Organic Compounds (SDBS), the NIST Chemistry WebBook, or various other substance databases [38, 39]. However, to our knowledge, there exists no such database in heterogeneous catalysis for experimental spectra of adsorbates.

In this paper, a novel software application is presented that enables the direct comparison of experimentally obtained IR spectra with DFT calculations with real-time adjustment of the peak broadening. The application follows the strategy development within the CaRMeN platform and is freely accessible through the browser [33, 35, 36, 40]. Furthermore, the new tool consists of a database of experimental and simulated IR spectra, with a strict separation between public and private data and with clearly defined metadata. Implementing the application accelerates the workflow of validation of DFT against experimental data.

2 Overview of Features

2.1 Database

This new application can be accessed by a web browser [35, 36]. Every registered user can upload their own spectra

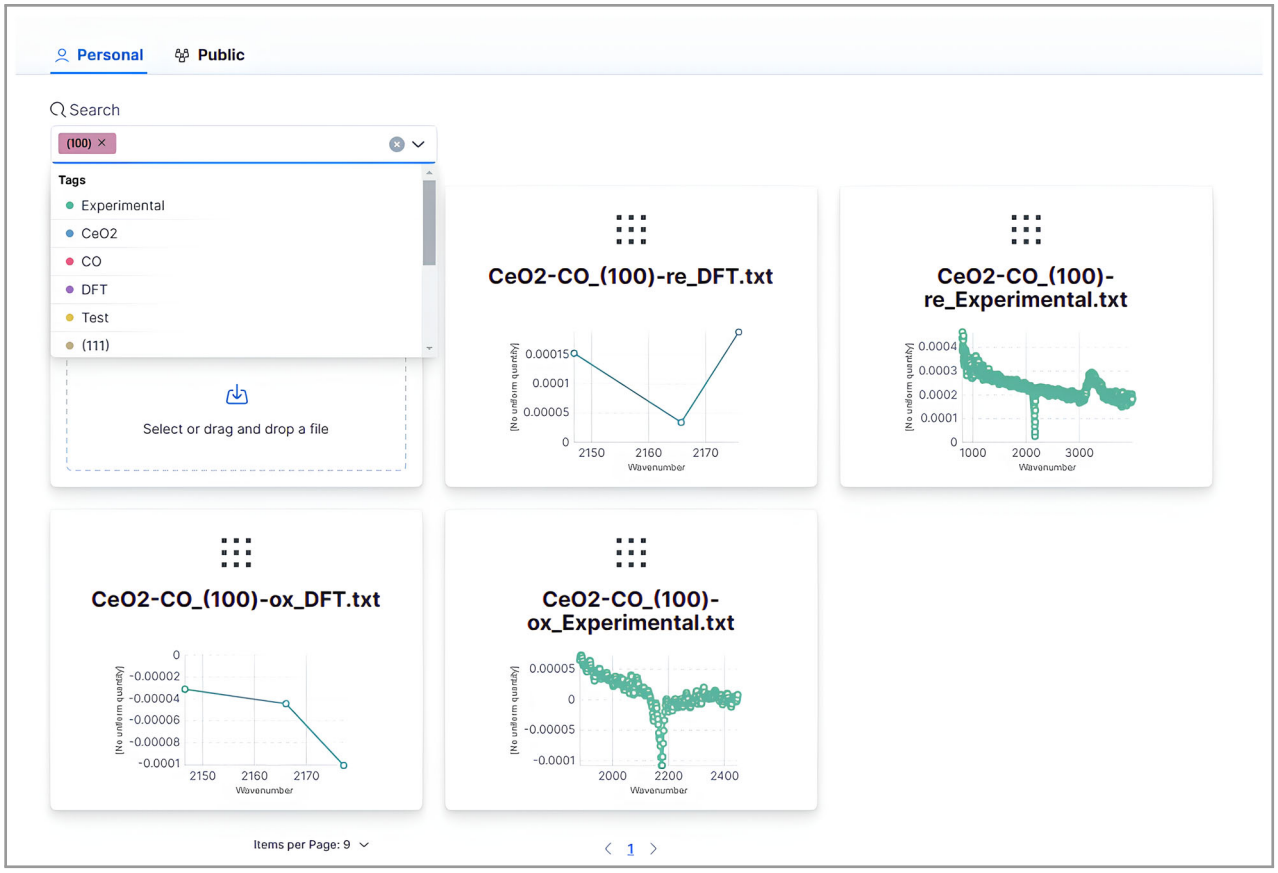

Figure 1. Screenshot of the overview page, where all personally uploaded spectra can be accessed. To improve accessibility, data can be filtered via tags or atoms and searched for with a full-text search engine.

as comma-separated values (CSV) files into the application through an import wizard. These files are parsed and converted into a standardized format, while the original data is preserved for traceability. By default, read and write access is restricted to the original creator of the data. However, users have the option to make individual spectra publicly accessible within the application. Once data is made public, all users, including unregistered ones, gain read access, while the original creator retains exclusive rights to modify the content. This same framework applies to entire projects and workspaces, referred to as "views", which will be elaborated further in the following section.

The database is designed following the FAIR guiding principles for scientific data management and stewardship [30]. Obligatory metadata for every spectrum, including digital object identifiers (DOI) for published data, user details, and atomic composition, is complemented by optional metadata, which are molecular details, user-defined tags, images, and free-text description. Spectra can be filtered by tags and atoms and searched for with a full-text search engine [41], and previews of the spectra are rendered inside the user interface (Fig. 1).

2.2 Plotting

Imported spectra can be quickly plotted into a so-called "view", i.e., a workspace where multiple spectra can be modified and compared (Fig. 2). When a spectrum consists of less than 50 data points, it is assumed to be simulated and is therefore plotted as a bar chart. Because this is an arbitrary boundary, the user can freely swap between bar and line plot. Wavenumbers are set as the standard x-axis and cannot be changed in the current version. However, since IR spectra typically utilize either absorbance or transmittance on the y-axis, adjustments to the y-values of individual plots are often necessary. This is especially the case for simulated spectra, where absolute intensities may not align with experimental data. Adjustments such as inversion, addition of a constant (with an auto-adjustment function available), or multiplication by a constant can be made.

Moreover, basic features for organizing the views are implemented. For example, each plot can be renamed, zoomed, or hidden. These settings persist in the database, but users can utilize a reset function to revert to the original settings upon opening the view. Finally, users have the

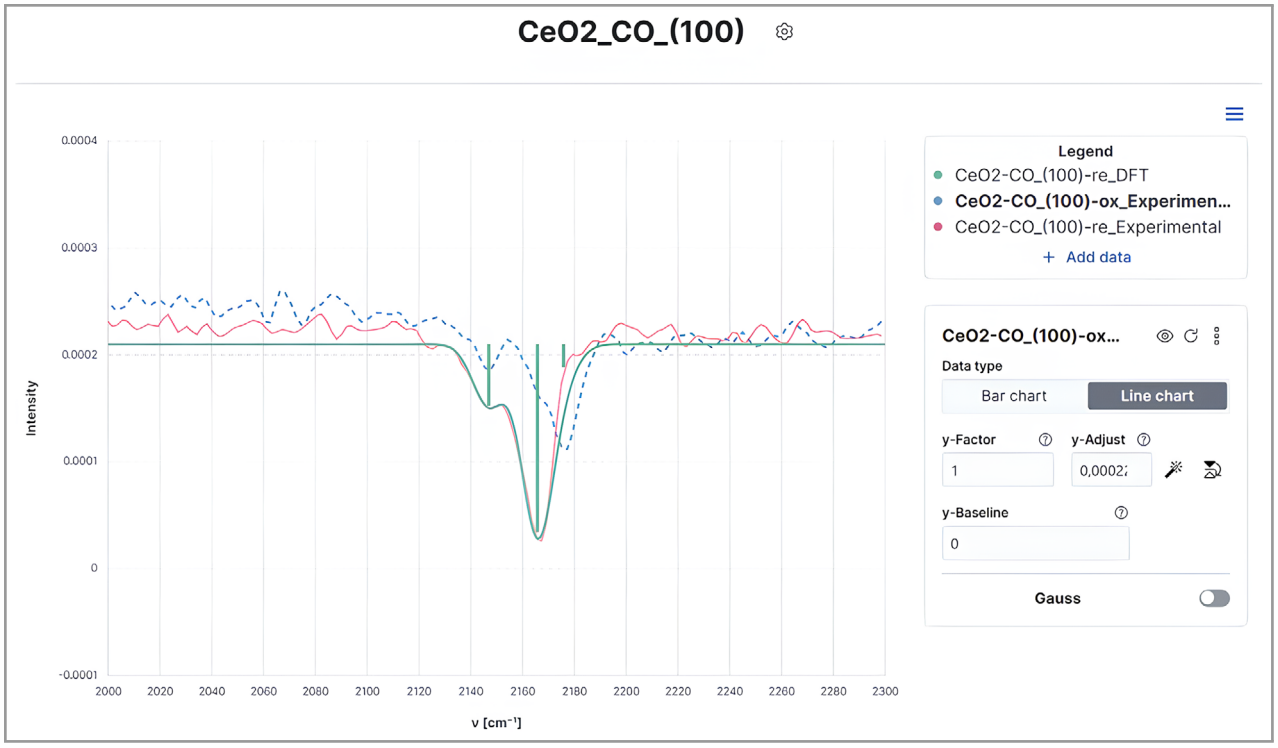

Figure 2. Screenshot of the "view" feature. One DFT spectrum (green) and two experimental IR spectra of CO adsorbed on oxidized (blue) and reduced (red) CeO_2-(100) surfaces by Yang et al. [15] loaded into a view. The peak broadening for the DFT spectrum is adjusted to fit one of the experiments. On the right-hand side, available chart settings are displayed.

option to export the view in CSV or PNG (portable network graphics) formats. In particular, the CSV data is convenient to generate publication-ready figures.

Views, like individual spectra, can also be made public within the application. Other users can access these public views and utilize the available functionalities to adjust them. However, any modifications made by these users will not be saved and will only be visible during their session.

2.3 Peak Broadening

To simulate the effect of peak broadening for DFT-generated spectra, a Gaussian curve of the same standard deviation σ is calculated for each peak (characterized by its center μ_i and intensity I_i). The sum of all Gaussian curves is plotted into the view together with the bars. A schematic representation of this process is illustrated in Fig. 3. The user is provided with the flexibility to conveniently adjust the standard deviation (σ) with a slider for precise control over the peak shape to match experimental data.

For each peak (μ_i, I_i) of a calculated spectrum, a Gaussian function is calculated as

$$f_i(\tilde{\upsilon}) = I_i \cdot e^{-\frac{1}{2}\left(\frac{\tilde{\upsilon}-\mu_i}{\sigma}\right)^2} \tag{1}$$

Because a slider is used to change the standard deviation, it is important to ensure that the underlying calculations resulting from swift changes in σ are computationally efficient. Otherwise, the performance loss would cause a non-responsive user interface. The sum of all Gauss functions is then calculated as

$$g(\tilde{\upsilon}) = \sum_i \begin{cases} f_i(\tilde{\upsilon}) & \text{if } f_i(\tilde{\upsilon}) > I_{i,\min} \\ 0 & \text{if } f_i(\tilde{\upsilon}) \le I_{i,\min} \end{cases} \tag{2}$$

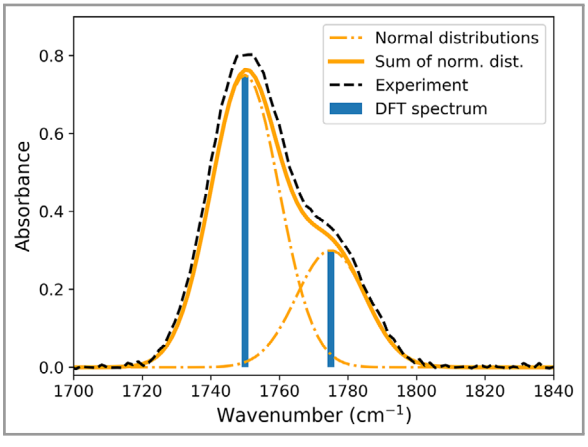

Figure 3. Schematic representation of the peak broadening effect applied to a DFT spectrum.

In this equation, a constant $I_{i,\min}$ (presently set to 0.001) representing the minimum intensity is introduced in the application for performance reasons, as it allows to recalculate $g(\tilde{\upsilon})$ faster in the event that the spectra contain many low-intensity peaks that are not important for the visualization. To further increase the speed of recalculating $g(\tilde{\upsilon})$, the threshold is defined in terms of $\tilde{\upsilon}$ not $f(\tilde{\upsilon})$. This optimization avoids computationally expensive calculations such as division, exponentiation, and square roots for each adjustment of σ. Therefore, the condition $[f_i(\tilde{\upsilon}) > I_{\min}]$ is reformulated as

$$[\mu_i + \sigma \cdot L > \tilde{\upsilon}\,] \wedge [\,\mu_i - \sigma \cdot L < \tilde{\upsilon}] \tag{3}$$

The formula checks if $\tilde{\upsilon}$ lies within a specific range centered around μ_i. The range extends from $[\mu_i + \sigma \cdot L > \tilde{\upsilon}]$ to $[\mu_i - \sigma \cdot L < \tilde{\upsilon}]$. If $\tilde{\upsilon}$ is within this range, the formula evaluates to true; otherwise, it evaluates to false. L is calculated once for each peak and does not need to be recalculated for each adjustment of σ:

$$L = \sqrt{-2 \ln\left(\frac{I_{\min}}{I_i}\right)} \tag{4}$$

In the user interface, the user can set the step size, denoted as h, for $\tilde{\upsilon}$. The function $g(\tilde{\upsilon})$ is then calculated and plotted for each μ_i as well as every $\tilde{\upsilon}$ within the specified range and with intervals of h. The idea is for the user to roughly adjust σ with a slider, maintaining a responsive user interface due to low calculation times of the low-resolution curve, and then reduce the step size, yielding a smoother higher-resolution curve for fine-tuning the standard deviation.

3 Illustrative Example

In general, a common approach for evaluating DFT data is the comparison of the numerically predicted data with experimental IR spectra. However, there are several challenges involved in this process, which the software presented in this paper overcomes. The first challenge is finding suitable experimental data, as this requires extensive literature research or conducting experiments themselves. Although literature data is readily available, it is often not in a standardized format. This necessitates manual data processing, which can be time-consuming and difficult.

The second challenge is performing the comparison between the simulated and experimental data. It is often the case that the intensities of the simulated spectra must be scaled, and peak broadening has to be considered, in order to visually compare the two spectra.

CaRIn (Catalysis Research with Infrared Spectroscopy) [35, 36] addresses these challenges by providing a user-friendly interface that allows for easy access to a large database of experimental IR spectra. The spectra are categorized using metadata that includes, e.g., atoms, tags, DOI, images, and the importing user, making it easy to search for specific data. The second challenge is addressed by providing tools that allow users to conveniently compare simulated and experimental spectra, including options to scale intensities and adjust peak broadening, all while providing immediate visual feedback.

An illustrative example serves to showcase the new software by describing the workflow with a typical paper from the literature. In the study on the CO/CeO_2 system by Yang et al. [14, 19] and Lustemberg et al. [18, 42], the individual single-crystal surfaces of ceria were systematically investigated via IR/SLIR (surface-ligand IR) with CO as the probe molecule [43]. Powder ceria materials are the most important form of ceria in industrial applications. The different facets of the crystal system exhibit very different structures and therefore chemical activities. Additionally, the oxidative or reductive potential of the reaction atmosphere further influences adsorbate-catalyst interaction through vacancy formation [19]. Depending on the oxidation state of the surface and the facet, a blue shift of existing CO bands or an appearance of additional CO peaks takes place in the typical wavenumber range above 2143 cm^{-1}. However, due to the weak bond between the oxide and CO, the vibrational modes are very close to each other, with a maximum of 2176 cm^{-1}. In the study, six experimental spectra were obtained for CO adsorbed on CeO_2 surfaces, specifically on the (100), (110), and (111) facets, in both reduced and oxidized forms [18]. Furthermore, a novel hybrid DFT approach, using mixtures of DFT(GGA) (generalized gradient approximation) and Hartree-Fock exchange energies (HSE06 functional), was developed to better simulate the CO-CeO_2 interactions [18]. This paper is representative of the described situation often found in the literature: It contains high-quality data but does not offer them in a machine-readable format, e.g., as supplemental material. In the following, we lay out the steps required to work with such data.

A researcher conducting further DFT calculations may be interested in the experimental data contained in the paper to validate their simulations. However, because the paper only contains the data in the form of a figure, the spectral data must be pulled out using data extraction software tools such as Datathief [44]. Even if the raw spectrum had been made directly available, e.g., as supplementary material, it would likely still have to be converted into a standard format such as CSV, with properly calibrated axis units.

Because the six experimental IR spectra of CO on the (100), (110), and (111) facets of ceria in both reduced and oxidized forms, along with the corresponding simulated spectra, were uploaded to the application, they were automatically converted into a standardized format and made searchable. By using the search engine and filtering by tags such as "CeO_2," "Experimental," the specific catalyst surface (e.g., "(100)"), and the adsorbate "CO," the available spectra are already narrowed down to two (Fig. 4).

Figure 4. Screenshot illustrating a filter example: All available spectra are filtered by four tags (CeO_2, (100), Experimental, CO) inside the search bar. As a result, only two spectra are shown.

From the viewpoint of an experimentalist attempting to establish a new post-treatment for CeO_2, they might wish to investigate whether their probe aligns more closely with the reduced or oxidized state of the catalyst. They might either find the simulated data by Lustemberg et al. [18, 42] within the application's database or access another service such as NOMAD [37]. Subsequently, several of the DFT spectra are loaded into a single view, alongside with their experiment. Given that there will be most certainly differences between their experimental setup and those of others, the peak broadening effect observed in their spectra will change as well.

For instance, in a reduced state, the CeO_2(110) surface exhibits an intense peak at 2175 cm^{-1} and a smaller peak at 2170 cm^{-1}, whereas on the oxidized catalyst surface, the peak at 2170 cm^{-1} has a higher intensity. If the broadening effect is sufficiently pronounced, the two peaks will merge, making it challenging to distinguish between the reduced and the oxidized state (Fig. 5). By adjusting the peak width for the DFT spectra while comparing them to the experimental spectrum of the researcher's own sample, it is possible to determine the state of reduction or oxidation of the catalyst. This will be even more pronounced if the wavenumbers themselves are slightly off and it becomes challenging to differentiate between the various cases due to the smaller peaks being obscured by the larger ones in the form of shoulders.

4 Methodology/Technical Implementation

The presented software is a TypeScript-based web application. Data is managed in the backend in a Structured Query Language (SQL) database, ensuring efficient and secure information storage. Raw files are kept in an S3 (simple storage service)-compatible global object storage. The backend exposes a GraphQL application programming interface (API) to facilitate efficient queries from the client side, which utilizes the React and Relay frameworks to provide a responsive and user-friendly interface. Authentication and authorization mechanisms are enforced using JavaScript Object Notation (JSON) Web Tokens, with access levels managed both in the frontend and within the GraphQL resolvers. The application includes unit tests for key functions, focusing on critical logic and high-risk areas to prevent regression and bugs. The application is hosted on the edge [45], leveraging a network of globally distributed servers to minimize latency and improve scalability.

Building on this technological foundation, API, SQL schemas, and the user interface were utilized, modified and extended to support metadata (Fig. 6), intricate search functionalities, and data sharing. Additionally, plotting, importing, and exporting tools were tailored for the comparison of experimental and DFT spectra.

5 Summary and Outlook

In this paper, we have shown the effectiveness of being able to adjust peak broadening for simulated IR spectra with real-time feedback through a new user interface, the app CaRIn, which is freely accessible through the browser [35, 36]. Furthermore, the proposed application includes a FAIR database feature that enables researchers to search and retrieve spectra from a shared repository, with a clear distinction between published and private data. In contrast to the peak broadening, the usefulness of the database feature is dependent on the network effect, because the more populated the database is, the easier it will be for a researcher to find significant spectra through the proposed search functionalities. However, this will come with the risk of low-quality data being introduced into the dataset, which will require continuous stewardship. Especially for experimental data, sample quality is often unknown and depends

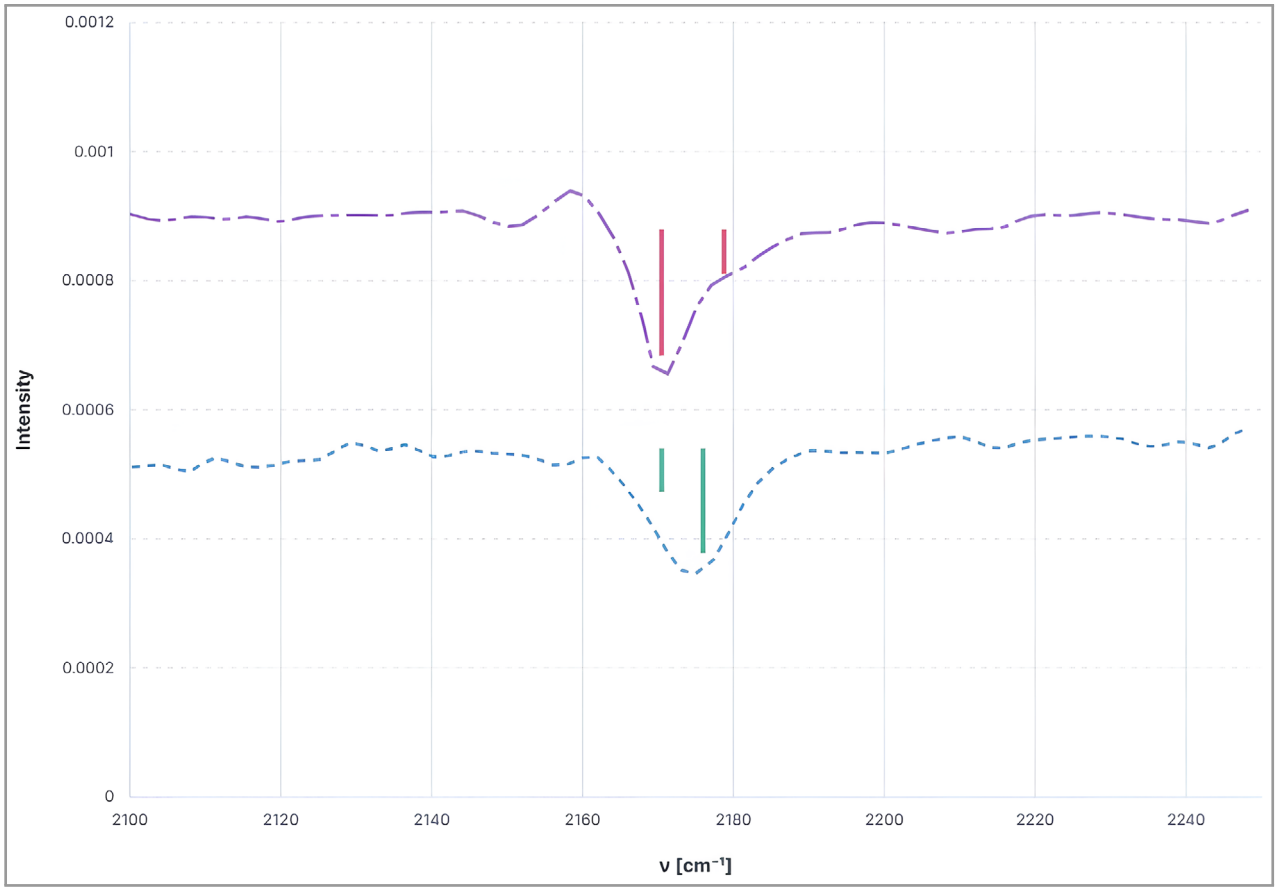

Figure 5. Experimental spectra of CO adsorbed on oxidized and reduced CeO_2(100) single-crystal surfaces [14, 15]. Underlying vibrational modes are shown as bars. Peak broadening effects lead to obfuscation of distinct vibrational features associated with the CO molecules.

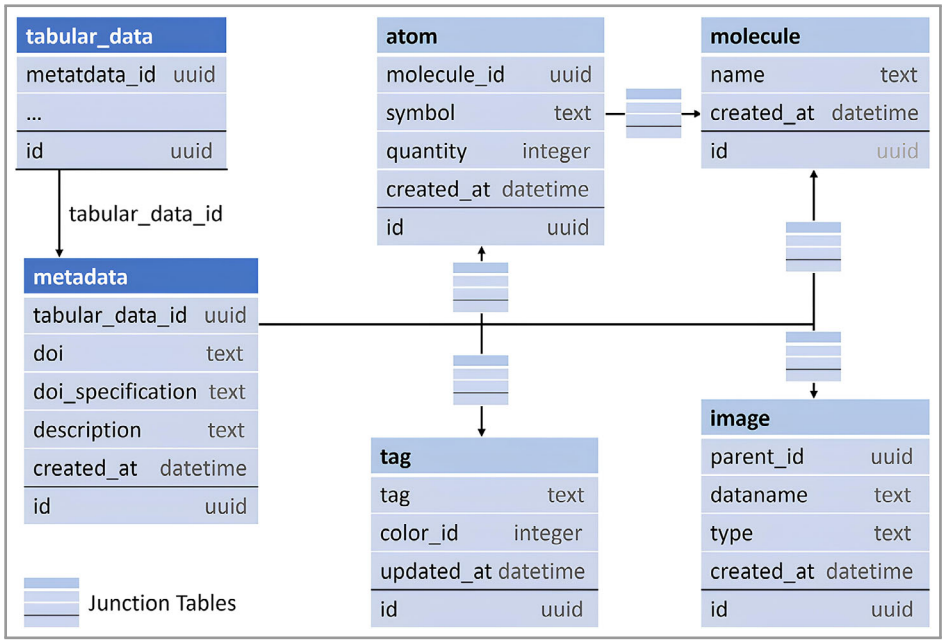

Figure 6. SQL diagram of the metadata entity that is linked to each spectrum (tabular data).

on its history, which is seldom recorded [37]. Novel tools like Adacta, by creating a digital twin of an experimental setup, produce a traceable history of all samples [32]. Coupling such tools to the platform CaRMeN, our application could offer researchers a comprehensive approach to managing and analyzing experimental spectra, enhancing the reliability and utility of the generated data.

In theory, signal processing algorithms in combination with a sufficiently populated database could be used to automatically identify unknown materials or specific features of IR spectra [46, 47]. However, this objective is beyond the scope of this work; the same applies to automated fitting of peak broadening, which could be implemented via a mathematical optimization with the objective function, trying to minimize the difference between the newly created curve of overlapping Gaussian curves and the experimental function within a certain wavenumber range. If needed by the community, these features could be implemented in the future. Finally, we would like to make you aware of the fact that scientifically based understanding of data is required, because manipulating the spectra through chart settings features can significantly decrease the scientific validity of the original simulation or experiment.

Acknowledgements

The authors thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for financial support via CRC-1441 (Project ID 426888090) and NFDI4Cat (Project ID 441926934). We very much acknowledge omegadot software & consulting GmbH, Limburgerhof, Germany, for use of the strategy of the CaRMeN platform and very fruitful support and discussions on our software development. S.B., J.J., and F.S. acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant no. INST 40/575-1 FUGG (JUSTUS 2 cluster, RVs bw17D011).

Funding by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) state of Baden-Württemberg.

Open access funding enabled and organized by Projekt DEAL.

Symbols Used

f_i	[–]	Gaussian function
g	[–]	sum of all Gaussian functions
I_i	[–]	peak intensity
$I_{i,\min}$	[–]	minimum threshold for intensity
L	[cm^{-1}]	peak distance threshold

Greek symbols

σ	[cm^{-1}]	standard deviation
μ_i	[cm^{-1}]	peak center
$\tilde{\upsilon}$	[cm^{-1}]	wavenumber

Sub-/superscripts

i	[–]	peak

Abbreviations

API	application programming interface
CSV	comma-separated values
DFT	density functional theory
DOI	digital object identifier
DRIFTS	diffuse-reflectance infrared Fourier transform spectroscopy
FAIR	findability, accessibility, interoperability, and reusability
HWHM	half-width at half-maximum
IR	infrared
JSON	JavaScript Object Notation
NOMAD	Novel Materials Discovery
S3	simple storage service
SDBS	Spectral Database for Organic Compounds
SQL	Structured Query Language

References

[1] D. C. Harris, M. D. Bertolucci, *Symmetry and Spectroscopy*, Dover Publications, New York **1989**.
[2] T. J. Toops, D. B. Smith, W. P. Partridge, *Appl. Catal., B* **2005**, *58* (*3/4*), 245–254. DOI: https://doi.org/10.1016/j.apcatb.2004.10.021
[3] R. Horn, K. A. Williams, N. J. Degenstein, L. D. Schmidt, *J. Catal.* **2006**, *242* (*1*), 92–102. DOI: https://doi.org/10.1016/j.jcat.2006.05.008
[4] D. Livio, C. Diehm, A. Donazzi, A. Beretta, O. Deutschmann, *Appl. Catal., A* **2013**, *467*, 530–541. DOI: https://doi.org/10.1016/j.apcata.2013.07.054
[5] M. Cavers, J. M. Davidson, I. R. Harkness, L. V. C. Rees, G. S. Mcdougall, *J. Catal.* **1999**, *430*, 426–430.
[6] G. Blyholder, *J. Phys. Chem.* **1964**, *68* (*10*), 2772–2777. DOI: https://doi.org/10.1021/j100792a006
[7] Y. Wang, C. Wöll, *Chem. Soc. Rev.* **2017**, *46* (*7*), 1875–1932. DOI: https://doi.org/10.1039/c6cs00914j
[8] T. V. W. Janssens, F. Zaera, *Surf. Sci.* **1995**, *344* (*1/2*), 77–84. DOI: https://doi.org/10.1016/0039-6028(95)00836-5
[9] G. A. Somorjai, *Surf. Sci.* **1994**, *299/300* (*C*), 849–866. DOI: https://doi.org/10.1016/0039-6028(94)90702-1
[10] G. Ertl, *Angew. Chem. Int. Ed.* **2008**, *47* (*19*), 3524–3535. DOI: https://doi.org/10.1002/anie.200800480
[11] M. Andersen, X. Yu, M. Kick, Y. Wang, K. Reuter, *J. Phys. Chem. C* **2018**, *122* (*9*), 4963–4971. DOI: https://doi.org/10.1021/acs.jpcc.8b00158
[12] V. Gorodetskii, J. Lauterbach, H. H. Rotermund, J. H. Block, G. Ertl, *Nature* **1994**, *370* (*6487*), 276–279. DOI: https://doi.org/10.1038/370276a0
[13] M. Trenary, *Annu. Rev. Phys. Chem.* **2000**, *51* (*1*), 381–403. DOI: https://doi.org/10.1146/annurev.physchem.51.1.381

[14] C. Yang, X. Yu, S. Heißler, A. Nefedov, S. Colussi, J. Llorca, A. Trovarelli, Y. Wang, C. Wöll, *Angew. Chem. Int. Ed.* **2017**, *56* (*1*), 375–379. DOI: https://doi.org/10.1002/anie.201609179

[15] X. Yu, P. Schwarz, A. Nefedov, B. Meyer, Y. Wang, C. Wöll, *Angew. Chem. Int. Ed.* **2019**, *58* (*49*), 17751–17757. DOI: https://doi.org/10.1002/anie.201910191

[16] S. Chen, P. N. Pleßow, Z. Yu, E. Sauter, L. Caulfield, A. Nefedov, F. Studt, Y. Wang, C. Wöll, *Angew. Chem. Int. Ed.* **2024**, *63* (*27*), e202404775. DOI: https://doi.org/10.1002/anie.202404775

[17] Y. Wang, A. Glenz, M. Muhler, C. Wöll, *Rev. Sci. Instrum.* **2009**, *80* (*11*), 113108. DOI: https://doi.org/10.1063/1.3257677

[18] P. G. Lustemberg, C. Yang, Y. Wang, C. Wöll, M. V. Ganduglia-Pirovano, *J. Chem. Phys.* **2023**, *159* (*3*), 034704. DOI: https://doi.org/10.1063/5.0153745

[19] C. Yang, M. Capdevila-Cortada, C. Dong, Y. Zhou, J. Wang, X. Yu, A. Nefedov, S. Heißler, N. López, W. Shen, et al., *J. Phys. Chem. Lett.* **2020**, *11* (*18*), 7925–7931. DOI: https://doi.org/10.1021/acs.jpclett.0c02409

[20] M. V. Ganduglia-Pirovano, A. Martínez-Arias, S. Chen, Y. Wang, P. G. Lustemberg, *Mater. Today Sustain.* **2024**, *26*, 4–7. DOI: https://doi.org/10.1016/j.mtsust.2024.100783

[21] J. Vazquez Quesada, P. S. Bernart, F. Studt, Y. Wang, K. Fink, *J. Chem. Phys.* **2024**, *161* (*22*), 224707. DOI: https://doi.org/10.1063/5.0231189

[22] J. K. Nørskov, F. Abild-Pedersen, F. Studt, T. Bligaard, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (*3*), 937–943. DOI: https://doi.org/10.1073/pnas.1006652108

[23] P. J. Feibelman, B. Hammer, F. Wagner, M. Scheffler, R. Stumpf, R. Watwe, J. Dumesic, *J. Phys. Chem. B* **2001**, *105*, 4018–4025.

[24] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77* (*18*), 3865–3868. DOI: https://doi.org/10.1103/PhysRevLett.77.3865

[25] J. Neugebauer, M. Reiher, C. Kind, B. A. Hess, *J. Comput. Chem.* **2002**, *23* (*9*), 895–910. DOI: https://doi.org/10.1002/jcc.10089

[26] J. Greeley, M. Mavrikakis, *Surf. Sci.* **2003**, *540* (*2/3*), 215–229. DOI: https://doi.org/10.1016/S0039-6028(03)00790-8

[27] B. H. Stuart, in *Infrared Spectroscopy: Fundamamentals and Applications*, John Wiley & Sons, New York **2004**.

[28] J. M. Hollas, in *Modern Spectroscopy*, 4th ed., John Wiley & Sons, Chichester **2003**.

[29] M. J. Frisch, G. W. Trucks, H. B. Schlegel, Gaussian 16 user's reference, https://gaussian.com/man/ (accessed on November 14, 2024).

[30] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al., *Sci. Data* **2016**, *3*, 1–9. DOI: https://doi.org/10.1038/sdata.2016.18

[31] C. Wulf, M. Beller, T. Boenisch, O. Deutschmann, S. Hanf, N. Kockmann, R. Kraehnert, M. Oezaslan, S. Palkovits, S. Schimmler, et al., *ChemCatChem* **2021**, *13* (*14*), 3223–3236. DOI: https://doi.org/10.1002/cctc.202001974

[32] H. Gossler, J. Riedel, E. Daymo, R. Chacko, S. Angeli, O. Deutschmann, *Chem. Ing. Tech.* **2022**, *94* (*11*), 1798–1807. DOI: https://doi.org/10.1002/CITE.202200064

[33] H. Gossler, L. Maier, S. Angeli, S. Tischer, O. Deutschmann, *Phys. Chem. Chem. Phys.* **2018**, *20*, 10857. DOI: https://doi.org/10.1039/c7cp07777g

[34] H. Gossler, L. Maier, S. Angeli, S. Tischer, O. Deutschmann, *Catalysts* **2019**, *9* (*3*), 1–11. DOI: https://doi.org/10.3390/catal9030227

[35] www.detchem.de/carmen/IR (accessed on November 14, 2024).

[36] https://www.itcp.kit.edu/deutschmann/english/Research.php (accessed on November 14, 2024).

[37] C. Draxl, M. Scheffler, *MRS Bull.* **2018**, *43* (*9*), 676–682. DOI: https://doi.org/10.1557/mrs.2018.208

[38] National Institute of Advanced Industrial Science and Technology (AIST), Spectral Database for Organic Compounds (SDBS), https://sdbs.db.aist.go.jp/Disclaimer.aspx (accessed on November 14, 2024).

[39] P. J. Linstrom, W. G. Mallard (Eds), *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, National Institute of Standards and Technology, Gaithersburg, **n.d.**

[40] https://detchem.de/carmen (accessed on November 14, 2024).

[41] L. Ongaro, MiniSearch, https://lucaong.github.io/minisearch/ (accessed on November 14, 2024).

[42] P. G. Lustemberg, P. N. Plessow, Y. Wang, C. Yang, A. Nefedov, F. Studt, C. Wöll, M. V. Ganduglia-Pirovano, *Phys. Rev. Lett.* **2020**, *125* (*25*), 256101. DOI: https://doi.org/10.1103/PhysRevLett.125.256101

[43] C. Yang, H. Idriss, Y. Wang, C. Wöll, *Acc. Chem. Res.* **2024**, *57* (*22*), 3316–3326. DOI: https://doi.org/10.1021/acs.accounts.4c00529

[44] B. Tummers, Datathief III, https://datathief.org/ **2006**.

[45] K. Cao, Y. Liu, G. Meng, Q. Sun, *IEEE Access.* **2020**, *8*, 85714–85728. DOI: https://doi.org/10.1109/ACCESS.2020.2991734

[46] S. G. Lambrakos, L. Huang, L. Massa, A. Shabev, *Nav. Res. Lab.* **2019**, (NRL/MR/6394–19-9850).

[47] S. W. Smith, in *Digital Signal Processing*, California Technical Publishing, San Diego, CA **2003**, pp. 123–140. DOI: https://doi.org/10.1016/b978-0-7506-7444-7/50044-3

