Multimodal Search in Chemical Documents and Reactions

Ayush Kumar Shah*
Abhisek Dey*
Rochester Institute of Technology
{as1211,ad4529}@rit.edu

Leo Luo*
University of Illinois
Urbana-Champaign
tluo9@illinois.edu

Bryan Amador*
Patrick Philippy
Rochester Institute of Technology
{ma5339,pmp2516}@rit.edu

Ming Zhong
Siru Ouyang
David Mark Friday
David Bianchi
Nick Jackson
University of Illinois
Urbana-Champaign
{mingz5,siruo2,friday1,davidmb2,jacksonn}@illinois.edu

Richard Zanibbi
Rochester Institute of Technology
rxzvcs@rit.edu

Jiawei Han
University of Illinois
Urbana-Champaign
hanj@illinois.edu

Abstract

We present a multimodal search tool that facilitates retrieval of chemical reactions, molecular structures, and associated text from scientific literature. Queries may combine molecular diagrams, textual descriptions, and reaction data, allowing users to connect different representations of chemical information. To support this, the indexing process includes chemical diagram extraction and parsing, extraction of reaction data from text in tabular form, and cross-modal linking of diagrams and their mentions in text. We describe the system's architecture, key functionalities, and retrieval process, along with expert assessments of the system. This demo[1] highlights the workflow and technical components of the search system.

CCS Concepts

• **Information systems → Chemical and biochemical retrieval**.

Keywords

multi-modal search, reactions, chemical diagrams

ACM Reference Format:
Ayush Kumar Shah, Abhisek Dey, Leo Luo, Bryan Amador, Patrick Philippy, Ming Zhong, Siru Ouyang, David Mark Friday, David Bianchi, Nick Jackson, Richard Zanibbi, and Jiawei Han. 2025. Multimodal Search in Chemical Documents and Reactions. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '25), July 13–18, 2025, Padua, Italy*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

*Authors who led and contributed equally to this research.
[1]https://www.cs.rit.edu/~dprl/reactionminer-demo-landing/

SIGIR '25, Padua, Italy

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

The scientific literature contains vast amounts of chemical knowledge represented in textual descriptions and diagrams. Roughly speaking, molecular and reaction diagrams represent the structure and relation of compounds participating in reactions, while the main text and text labels on reaction diagram entities describe how and why reactions occur (e.g., at what temperature, the yield of a product molecule, pertinent molecular properties, etc.). As a result, the full story of a reaction is often told using a combination of text and diagrams.

Traditional Chemical Information Retrieval (CIR) systems and commercial platforms such as SciFinder®[2] and Reaxys®[3] provide extensive text-based and structural search capabilities. However, these systems do not explicitly associate molecular figures with their textual descriptions, making it difficult to retrieve reactions found in diagrams with their details and context provided in text. Existing systems are also designed to return individual compounds or full documents as results, rather than returning text passages where reaction descriptions are found. These limitations present challenges for chemists, patent examiners, and researchers seeking to retrieve relevant compounds, reactions, or synthesis protocols along with their contexts efficiently.

To address these challenges, our system supports direct retrieval of relevant passages, which are returned along with their associated molecular structures. This includes structures extracted from molecular diagrams and referenced in the text by common name (e.g., 'chromene'), IUPAC name [19], or figure identifier (e.g., molecule '34', or molecule '4b'). Automatically extracted reaction records are also generated from passages using *ReactionMiner* [23], and provided alongside passages and their associated compounds. The reaction records enable researchers to explore linked reaction steps. Passages may be searched using text queries, molecular structure queries in SMILES (Simplified Molecular Input Line Entry System) strings [22], or a combination of the two. SMILES is frequently used in chemoinformatics, and can be readily generated by a number of commonly used drawing tools (e.g., ChemDraw or Marvin).

[2]SciFinder®: https://scifinder.cas.org/
[3]Reaxys®: https://www.reaxys.com/

This multimodal search model facilitates direct navigation between reaction text, associated molecule diagrams, and extracted chemical entities. The system supports text-based search using BM25, and SMILES-based molecular search using structural similarity and substructure matching provided by RDKit[4].

Related Works. Early CIR systems such as ChemXSeer [12] focused on extracting and indexing chemical names from tables in PDFs, in order to allow users to search by molecule name or formula. Later efforts, such as TREC-CHEM [9], introduced the concept of document-level retrieval for chemistry-specific tasks, curating a dataset of patents and employing manual relevance assessments to evaluate retrieval effectiveness. Recent advances in deep learning have focused on cross-modal learning to align structured molecular representations with textual descriptions. Both Text2Mol [3] and MoleculeSTM [7] adopt joint learning approaches that embed chemical structures and text into a shared embedding space, facilitating retrieval across modalities. Text2Mol employs graph neural networks (GNNs) [15] and perceptron models to predict the most appropriate SMILES representation given a textual query, addressing the task of molecule retrieval from natural language descriptions. Similarly, MoleculeSTM is designed for structure-text retrieval, retrieving chemical structure from textual descriptions and vice versa by leveraging a multimodal transformer within a contrastive learning framework. However, these systems focused on document-level indexing rather than passage-level search, whereas our system enables structured retrieval by directly linking extracted molecular diagrams, textual mentions, and reaction descriptions. OpenChemIE [4] extracts reaction data from text, tables, and figures using modality-specific models, similar to our approach of combining text-based reaction extraction and molecular diagram parsing. However, unlike OpenChemIE, our system integrates extraction with passage-level retrieval.

Our demonstration showcases how structured indexing and linking of chemical information across text and figures can improve chemical information retrieval and provide a more comprehensive and flexible search framework for chemists and researchers.

2 Indexing Extracted Compounds and Reactions

In our system, textual and graphical content are processed through two parallel pipelines, whose outputs are later used to build a unified index of chemical entities, paragraph texts, and reactions.

Text mining for extracting reaction information. We employ *ReactionMiner [23]*, a pipeline that processes text extracted from PDF documents to isolate and categorize reaction-related content. First, the text is segmented into reaction-related sentences through product-indicative keywords and topic modeling [2] for defining the contextual boundary. A large language model, LLaMA3.1-8b [5], fine-tuned with LoRA [6], is then used to identify relevant chemical entities such as reactants, products, and catalysts, along with key conditions (e.g., temperature, reagents, or solvents). Each reaction mention is associated with segmented text bounding boxes in the PDF (see Figure 1), enabling direct navigation to the underlying paragraph. By grouping identified mentions into coherent reaction units, a structured record for each reaction step present in the text is established.

Extracting SMILES from document text. In addition to the reaction records extracted by *ReactionMiner*, we also extract individual compounds from document text. These additional SMILES annotations ensure that compounds mentioned both inside and outside of reaction descriptions can be retrieved through the search interface. For indexing compounds in text, we first use PyTesseract to convert document page images to text, which is then passed to *ChemDataExtractor2.0's* [11] Chemical Named Entity Recognition (CNER) system to identify molecule names. Each recognized name is converted into a SMILES string via *OPSIN* [8], and any passage containing at least one valid SMILES is retained for indexing.

Molecular diagram extraction and parsing. We use YOLOv8, an improved version of scaled YOLOv4 [21], for detecting molecular regions in documents. We then employ *ChemScraper* [16] to parse molecular diagrams from detected PDF regions (see Figure 1) through two complementary pipelines: a *born-digital* approach for vector images representing characters and geometric objects, and a *visual parsing* [18] approach for pixel-based raster images. In the born-digital parser, *SymbolScraper* [17] accesses low-level PDF drawing commands to extract lines, polygons, and characters directly from the PDF. The visual parser works on raster images, applying the *Line Segment Detector (LSD)* [20] and watershed algorithm to detect line primitives and text regions. Together, these methods yield a set of graphical elements (e.g., atoms, bond lines, named functional groups) and their local connections.

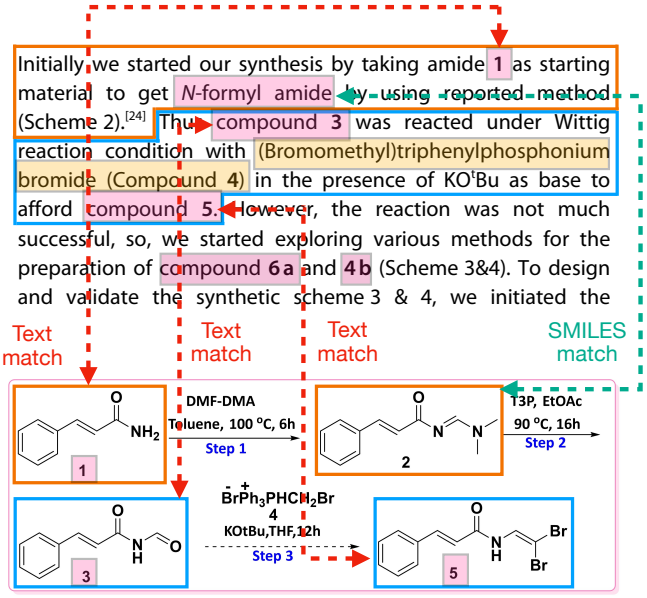
ChemScraper then constructs a *visual structure graph* using a Minimum Spanning Tree (MST) and rewrite rules for born-digital diagrams and a segmentation-aware, multi-task neural network for raster images. The visual graph is then converted to a molecular graph, where bonds become edges and implicit carbon atoms are inferred from line intersections. The final molecular graph is stored in *CDXML format* to retain both chemical and visual structure. This format is converted into SMILES for indexing.

Compound-Passage Linking and Multimodal Indexing. Once text-based reaction information and text and diagram-based molecular structures are extracted, they must be unified into a cohesive representation that supports flexible querying. There are two types of passages in our system: (1) those extracted by *ReactionMiner* (boxed text in Figure 1), and (2) those extracted from general text regions using PyTesseract (unboxed).

For passages extracted by *ReactionMiner*, relevant text fields include reactants, products, catalysts, and yields. We focus on linking the reactants and products in reaction records with their corresponding molecular diagrams using two approaches: *1. Token-based text matching:* Text mentions of diagram labels for reactants and products are identified using regular expressions, and then matched with the nearest diagram text label by a normalized Levenshtein similarity ratio[5], ensuring minor variations in naming do not prevent linkage. An example is shown in red in Figure 1, where 'compound 5' is matched with the diagram labeled '5.' *2. SMILES-based fingerprint matching:* The text tokens are first processed through *ChemDataExtractor2.0* [11] and *OPSIN* [8], following the same approach described earlier for non-reaction text, while SMILES from diagrams are extracted by *ChemScraper* [16]. Each SMILES representation undergoes molecular fingerprinting, producing a binary

[4]https://www.rdkit.org

[5]https://rapidfuzz.github.io/Levenshtein/levenshtein.html#ratio



Scheme 2. Initial Attempted Synthesis of Lansiumamide A- Amino acid hybrid.

Figure 1: Text/diagram compound extraction and compound–passage linking. Two passage types are shown: (1) reaction passages from *ReactionMiner* (2 boxes) and (2) a single text passage containing both extracted compounds. Highlighted text denotes extracted chemical entities: pink for mentions in molecular diagrams, yellow for unmatched mentions. Matches come from (1) *Text matching* via Levenshtein distance and (2) *SMILES matching* via *Tanimoto Similarity*. Highlight colors (e.g., orange and blue) indicate molecules & reaction text linked to the same reaction passage.

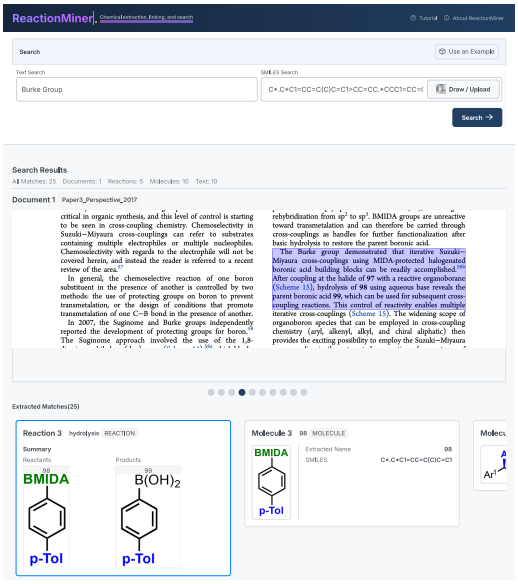


Figure 2: Multi-modal search results for a text and Reaction SMARTS query. Results are organized by document, with matched passages linked to extracted reactions, molecular structures, and highlighted text mentions. Key reaction details, including reactant ('98') and product ('99') in both text and diagrams, along with their predicted SMILES representations, are displayed. Users can navigate directly to relevant sections within each document, with highlighted passages indicating the corresponding matches.

vector encoding the molecular graph's connectivity patterns. To determine similarity, we compute the *Tanimoto Similarity* [1] measuring the overlap between fingerprint vectors, giving a similarity score between 0 and 1. The diagram with the highest *Tanimoto Similarity* score with a passage SMILES are linked. An example is shown in blue in Figure 1, where 'N-formyl amide' is matched with the diagram labeled '2.' If a compound can be linked by both strategies, the strategy with the higher score between the normalized *Levenshtein similarity ratio* and the *Tanimoto Similarity* is chosen.

In the final index, passages are annotated with associated SMILES and reaction entities. This enables retrieval of specific compounds as well as reaction-related information, linking molecular structures to their reaction context. During indexing, IUPAC names in text are also tokenized into constituent groups [13] to improve retrieval. For example, an IUPAC name, 'N-((E)-2-bromo-2-phenylvinyl)-cinnamamide' would be tokenized as 'N E 2 bromo 2 phenyl vinyl cinnamamide'.

3 Multimodal Search: Suzuki Coupling Papers

Our system supports three primary search models based on different query types: text search, molecular structure/substructure search (from SMILES), and multimodal search combining text and molecular queries. Text search is handled using BM25 [14] as implemented in PyTerrier [10], allowing users to search for compounds, reaction conditions, or chemical properties in text. Molecular structure search is provided by RDKit, enabling both exact molecule matching and substructure searching. Multimodal search integrates results from both models, by re-ranking results to prioritize passages that contain both textual and molecular matches.

For this demonstration, we indexed seven research papers and six supplementary information documents related to Suzuki coupling reactions, provided by chemists at the University of Illinois. The dataset includes a total of 1282 extracted passages, out of which 538 are indexed (passages without links to a reaction or compound name are removed), 383 unique SMILES string, and 219 extracted reactions. These passages are linked to diagrams and SMILES as described in the previous section. Molecule SMILES are also searchable within detected chemical entities in passages and reactions. This structured linking allows users to retrieve molecular and reaction information, whether entities appear in text descriptions or are represented as molecular diagrams.

Molecular SMILES and SMARTS Search. Query SMILES and candidate SMILES in the index are converted into fingerprint vector representations using Morgan Fingerprinting (2048 bits), and are ranked using *Tanimoto Similarity*, described earlier. The system also accommodates structured queries specifying entire chemical

reactions using *Reaction SMARTS* (SMILES Arbitrary Targets Specification) notation. This format extends SMILES strings to include '>' separating reactants, reagents, and products, while '.' separates individual SMILES compounds within each category, as seen in Figure 2. To process these queries, SMILES compounds are extracted from the SMARTS string, and each is matched against indexed molecular structures and reaction passages.

This allows retrieval of passages describing reactions described in SMARTS, while also providing access to details provide in *ReactionMiner* records linked to passages. Information such as reactants, products, catalysts, reaction conditions, and temperature are presented for retrieved passages. This enables users to search for entire reactions rather than just isolated compounds, improving retrieval of contextually relevant reaction information.

Multimodal Search. For multimodal queries, containing both text and SMILES (or Reaction SMARTS), as shown in Figure 2, the initial SMILES candidates are obtained using *Sub-structure search*. We find that including all possible sub-structures as valid candidates for re-ranking text candidates leads to better performance as *Tanimoto Similarity* can be unpredictable for very specific information needs when combined with text. If multiple SMILES are provided, the retrieved passages for each individual SMILES query are aggregated. Results from text-based and SMILES-based retrieval are combined. A fusion step adjusts rankings to prioritize passages containing a higher number of matched SMILES. Retrieved passages are re-ranked based on their BM25 text relevance score and the presence of SMILES matches.

Reaction Navigation. The system also provides a dedicated reaction navigation feature for each retrieved document. When a user selects a passage, they can inspect all extracted reactions from the associated document in a structured list, with each reaction entry pointing to the relevant PDF page and bounding-box highlights, as shown in Figure 2. This approach lets users explore multiple reaction mentions in context, making it easier to follow complex procedures, compare alternative synthetic routes, or identify recurring reagents and intermediates within a single publication.

4 Expert Evaluation by Chemists

To evaluate the system's effectiveness for chemists, we conducted an expert assessment with researchers at the University of Illinois. The system effectively retrieved relevant chemical information, linking molecular diagrams and text-based reaction details to chemical names or SMILES queries. For example, as shown in Figure 2, a multi-modal search with the text query 'Burke group' and a Reaction SMARTS string successfully retrieved passages with relevant reactions matching the SMILES query, which were associated with the 'Burke group', as shown in the highlighted text and corresponding reaction diagrams displayed below the document image. Chemists found the ability to click on molecule 'cards' and navigate directly to the corresponding section in the document particularly useful. The structured reaction output captured key experimental details such as yield, catalysts, solvents, and temperature, enhancing the accessibility of reaction data. The reaction and molecule cards serve as a structured extractive summary of the paper, while also providing navigation links to their original context. Note that the 'Reaction 3' in Figure 2 does not show these additional details as they were not available in the text. Notably, the system retrieved derivatives of a queried molecule, such as 'benzo[b]thiophen-2-ylboronic acid,' which was relevant to the SMILES query 'C1=CC=C2C(=C1)C3=CC=CC=C3S2' (dibenzothiophene) but not explicitly searched for. Overall, the chemists were able to find the information they were hoping for, and the retrieved results were useful for their research.

While reaction details were generally well extracted, experts recommended incorporating additional metadata such as 'equivalents' and mol% of catalysts, which are essential for exporting data to electronic lab notebooks. When using a combination of text and SMILES queries, users found it challenging to determine whether retrieved results were more influenced by text-based or structure-based matching, suggesting a need for greater transparency in the effect of text vs. SMILES on ranking. Additionally, filtering options to view reactions, molecules, and text separately would improve usability, allowing chemists to focus on the most relevant data. Another key area for improvement is enhanced diagram-text linking, as some extracted text mentions were not associated with their corresponding molecular diagrams. Addressing these issues would further enhance the system's utility for chemical research.

5 Conclusion and Future Work

This work presents a multimodal search system that integrates text and molecular structure retrieval, enabling passage-level search with structured linking between chemical entities, molecular diagrams, and reaction descriptions. By combining BM25 for text, RDKit-based molecular similarity search, and a fusion mechanism for multimodal queries, our system improves access to chemical knowledge in scientific literature. The expert evaluation with chemists demonstrated the system's usability, with researchers successfully retrieving relevant chemical information, including useful molecular derivatives and structured reaction details.

Future work will focus on enhancing retrieval effectiveness with dense embeddings and cross-modal search, leveraging transformer-based models to improve ranking and semantic matching across text and molecular representations. While the system currently matches chemical names to diagrams via SMILES translation, inspired by Text2Mol [3], we aim to explore query expansion within an aligned multimodal embedding space. This approach would expand text queries to incorporate corresponding molecular diagrams or SMILES representations and extend SMILES queries to include relevant text-based descriptors, improving retrieval flexibility. Additional directions include scaling the system to index larger collections and integrating external chemical databases, and refine filtering mechanisms to improve user experience. A more mature version of our system may find use in chemical research, industry, and patent analysis, reducing time spent on literature review and supporting efficient retrieval of structured chemical information.

Acknowledgments

This work was supported by the National Science Foundation USA (Grant #2019897, Molecule Maker Lab Institute). We also thank Matt Berry, Kate Arneson, Bingji Gao, Sara Lambert, and other members from the NCSA team who helped create the online system.

References

[1] Adrià Cereto-Massagué, María José Ojeda, Cristina Valls, Miquel Mulero, Santiago Garcia-Vallvé, and Gerard Pujadas. 2015. Molecular fingerprint similarity search in virtual screening. *Methods* 71 (2015), 58–63. doi:10.1016/j.ymeth.2014.08.005

[2] Freddy Y. Y. Choi. 2000. Advances in domain independent linear text segmentation. In *1st Meeting of the North American Chapter of the Association for Computational Linguistics*. https://aclanthology.org/A00-2004/

[3] Carl Edwards, ChengXiang Zhai, and Heng Ji. 2021. Text2Mol: Cross-Modal Molecule Retrieval with Natural Language Queries. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih (Eds.). Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 595–607. doi:10.18653/v1/2021.emnlp-main.47

[4] Vincent Fan, Yujie Qian, Alex Wang, Amber Wang, Connor W. Coley, and Regina Barzilay. 2024. OpenChemIE: An Information Extraction Toolkit for Chemistry Literature. *Journal of Chemical Information and Modeling* 64, 14 (July 2024), 5521–5534. doi:10.1021/acs.jcim.4c00572

[5] Aaron Grattafiori et al. 2024. The Llama 3 Herd of Models. doi:10.48550/arXiv.2407.21783 arXiv:2407.21783 [cs].

[6] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685* (2021).

[7] Shengchao Liu, Weili Nie, Chengpeng Wang, Jiarui Lu, Zhuoran Qiao, Ling Liu, Jian Tang, Chaowei Xiao, and Animashree Anandkumar. 2023. Multi-Modal Molecule Structure–Text Model for Text-Based Retrieval and Editing. *Nature Machine Intelligence* 5, 12 (Dec. 2023), 1447–1457. doi:10.1038/s42256-023-00759-6

[8] Daniel M. Lowe, Peter T. Corbett, Peter Murray-Rust, and Robert C. Glen. 2011. Chemical Name to Structure: OPSIN, an Open Source Solution. *Journal of Chemical Information and Modeling* 51, 3 (2011), 739–753. doi:10.1021/ci100384d arXiv:https://doi.org/10.1021/ci100384d PMID: 21384929.

[9] Mihai Lupu, Jimmy Huang, Jianhan Zhu, and John Tait. 2009. TREC-CHEM: large scale chemical information retrieval evaluation at TREC. *SIGIR Forum* 43, 2 (Dec. 2009), 63–70. doi:10.1145/1670564.1670576

[10] Craig Macdonald, Nicola Tonellotto, Sean MacAvaney, and Iadh Ounis. 2021. PyTerrier: Declarative Experimentation in Python from BM25 to Dense Retrieval. In *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management, Virtual Event, Queensland, Australia, November 1 - 5, 2021*, Gianluca Demartini, Guido Zuccon, J. Shane Culpepper, Zi Huang, and Hanghang Tong (Eds.). ACM, 4526–4533. doi:10.1145/3459637.3482013

[11] Juraj Mavračić, Callum J. Court, Taketomo Isazawa, Stephen R. Elliott, and Jacqueline M. Cole. 2021. ChemDataExtractor 2.0: Autopopulated Ontologies for Materials Science. *Journal of Chemical Information and Modeling* 61 (9 2021), 4280–4289. Issue 9. doi:10.1021/acs.jcim.1c00446

[12] Prasenjit Mitra, C. Lee Giles, Bingjun Sun, and Ying Liu. 2007. ChemXSeer: a digital library and data repository for chemical kinetics. In *Proceedings of the ACM First Workshop on CyberInfrastructure: Information Management in EScience* (Lisbon, Portugal) *(CIMS '07)*. Association for Computing Machinery, New York, NY, USA, 7–10. doi:10.1145/1317353.1317356

[13] Kohulan Rajan, Achim Zielesny, and Christoph Steinbeck. 2021. STOUT: SMILES to IUPAC names using neural machine translation. *Journal of Cheminformatics* 13 (12 2021). Issue 1. doi:10.1186/s13321-021-00512-4

[14] Stephen E. Robertson, Steve Walker, Susan Jones, Micheline Hancock-Beaulieu, and Mike Gatford. 1994. Okapi at TREC-3. In *Text Retrieval Conference*. https://api.semanticscholar.org/CorpusID:41563977

[15] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2009. The Graph Neural Network Model. *IEEE Transactions on Neural Networks* 20, 1 (Jan. 2009), 61–80. doi:10.1109/TNN.2008.2005605

[16] Ayush Kumar Shah, Bryan Amador, Abhisek Dey, Ming Creekmore, Blake Ocampo, Scott Denmark, and Richard Zanibbi. 2024. ChemScraper: leveraging PDF graphics instructions for molecular diagram parsing. *International Journal on Document Analysis and Recognition (IJDAR)* 27, 3 (2024), 395–414. doi:10.1007/s10032-024-00486-7

[17] Ayush Kumar Shah, Abhisek Dey, and Richard Zanibbi. 2021. A Math Formula Extraction and Evaluation Framework for PDF Documents. In *Document Analysis and Recognition – ICDAR 2021 (Lecture Notes in Computer Science)*, Josep Lladós, Daniel Lopresti, and Seiichi Uchida (Eds.). Springer International Publishing, Cham, 19–34. doi:10.1007/978-3-030-86331-9_2

[18] Ayush Kumar Shah and Richard Zanibbi. 2023. Line-of-Sight with Graph Attention Parser (LGAP) for Math Formulas. In *2023 International Conference on Document Analysis and Recognition (ICDAR)* (San José, CA, USA). Springer-Verlag, Berlin, Heidelberg, 401–419. doi:10.1007/978-3-031-41734-4_25

[19] Stanislaw Skonieczny. 2006. The IUPAC rules for naming organic molecules. *Journal of Chemical Education* 83 (2006), 1633–1637. Issue 11. doi:10.1021/ed083p1633

[20] Rafael Grompone von Gioi, Jérémie Jakubowicz, Jean-Michel Morel, and Gregory Randall. 2012. LSD: A Line Segment Detector. *Image Processing On Line* 2 (March 2012), 35–55. doi:10.5201/ipol.2012.gjmr-lsd

[21] Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. 2021. Scaled-YOLOv4: Scaling Cross Stage Partial Network. In *2021 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 13024–13033. doi:10.1109/CVPR46437.2021.01283

[22] David Weininger. 1988. SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules. *Journal of Chemical Information and Computer Sciences* 28, 1 (1988), 31–36. doi:10.1021/ci00057a005

[23] Ming Zhong, Siru Ouyang, Yizhu Jiao, Priyanka Kargupta, Leo Luo, Yanzhen Shen, Bobby Zhou, Xianrui Zhong, Xuan Liu, Hongxiang Li, Jinfeng Xiao, Minhao Jiang, Vivian Hu, Xuan Wang, Heng Ji, Martin Burke, Huimin Zhao, and Jiawei Han. 2023. Reaction Miner: An Integrated System for Chemical Reaction Extraction from Textual Data. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Yansong Feng and Els Lefever (Eds.). Association for Computational Linguistics, Singapore, 389–402. doi:10.18653/v1/2023.emnlp-demo.36

Received 18 February 2025