

Chemoinformatics : a textbook / Johann Gasteiger, Thomas Engel (Eds.) Weinheim: Wiley-VCH, 2003. ISBN 3-527-30681-1. 649 S. 69,- Euro

Das Jahrhundert der chemischen Literatur und deren Verschwinden

Das ganze 20. Jahrhundert hindurch erschienen Literaturführer zur chemischen Information, als letztes die 3. Auflage von Maizells „How to find chemical information“.¹ Heute ist durch die Verwendung moderner Computer- und Informationstechnik sowie aufgrund der weltweiten Vernetzung gerade die Chemie ein Beispiel für das verstärkte Verschwinden nicht-elektronischer Informationsmedien. Das hier zu besprechende Lehrbuch zur Chemoinformatik, also zur Anwendung von Methoden der Informatik bei der Lösung chemischer Probleme, verdeutlicht, dass neben den Biowissenschaften wohl keine andere Naturwissenschaft so weit auf dem Weg zu einer Informationswissenschaft vorangeschritten ist wie die Chemie.²

Das Fach Chemie und seine Informationsprobleme erforderten es frühzeitig, neue Ideen zu deren Lösung zu entwickeln. Oft waren sie dann Beispiel auch für andere Fachgebiete. Viele Pioniere der Information und Dokumentation kamen daher aus dem Bereich der Chemie.

Schon Anfang des 20. Jahrhunderts beschäftigten sich erste Zeitschriftenartikel und Bücher mit der chemischen Literatur und Information. Wilhelm Ostwalds Buch "Die chemische Literatur und die Organisation der Wissenschaft" (Leipzig, 1919) wird in der "International Encyclopedia of Information and Library Science" als ein frühes Beispiel spezifischer Literatur zur Informationswissenschaft erwähnt.³ In die gleiche Zeit fallen erste Aktivitäten zur Vermittlung des Umgangs mit chemischen Informationsmitteln im Rahmen der Lehre an Universitäten.⁴

Zur Geschichte der Chemoinformatik

Das Lehrbuch „Chemoinformatics“ erscheint zusammen mit dem ebenfalls von Johann Gasteiger mit herausgegebenen „Handbook of Chemoinformatics : from data to knowledge in 4 volumes“ (Wiley-VCH, 2003, 699 Euro). Das Lehrbuch stellt eine Art Auszug aus dem Handbuch dar und ist hauptsächlich von Mitgliedern der Arbeitsgruppe von Gasteiger am Computer-Chemie-Centrum der Friedrich-Alexander-Universität in Erlangen-Nürnberg verfasst worden. Der speziell an der chemischen Fachinformation interessierte Leser findet im Handbuch weitaus mehr wichtige Kapitel als im Lehrbuch.

Unter anderem enthält das „Handbook“ auch ein von Peter Willett verfasstes Kapitel zur Geschichte der „Chemoinformatics“ (Handbook, S. 6-20), die im Lehrbuch nur in 6 Zeilen gestreift wird (S. 9-10): Willett beschreibt die Entwicklung von Computertechniken, um chemische Strukturinformationen abzubilden. Aber auch die Verarbeitung chemischer Information und chemischer Stoffdaten, die anfangs bei der Herstellung gedruckter chemischer Informationsmittel erfolgte, kann zur Vorgeschichte der Chemoinformatik gezählt werden.⁵ Der Computereinsatz bei der Herstellung der Register gedruckter Bibliographien führte letztendlich auch zur Entwicklung der heutigen Online-Informationendienste.⁶

¹ Robert Edward Maizell: How to find chemical information : a guide for practicing chemists, educators, and students. 3rd ed. New York: Wiley, 1998. Für eine Liste dieser Werke siehe Melvin Guy Mellon: Chemical publications, their nature and use . 5th ed. New York : McGraw-Hill, 1982, S. 245f.

² Zur Situation in den Biowissenschaften siehe Timothy Lenoir, „Shaping Biomedicine as an Information Science“, in den Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems / Bowden, M.E., T. B. Hahn, and R. V. Williams (Eds.), Medford, NJ: Information Today, 1999, Volltext im Netz unter www.chemheritage.org/explore/ASIS_documents/ASIS98_main.htm

³ Im Artikel von R. T. Bottle, Information Science, in: J. Feather et al. (Eds.), International encyclopedia of information and library science, 2. ed., Routledge / London, 2003 (S.295-297), wird Ostwald auf S. 296 erwähnt.

⁴ Sparks, Marion E.: Chemical literature and its use. Science N.S. 47 (1918) No. 1216, 377-381.

⁵ Gerade zur Vorgeschichte der Chemoinformatik in der chemischen Dokumentation finden sich interessante Beiträge auf zwei Konferenzen zur Geschichte wissenschaftlich technischer Informationssysteme, siehe den Proceedings-Band zur 1. Konferenz in Anmerkung 2 sowie den Band The History and Heritage of Scientific and Technological Information Systems: Proceedings of the 2002 Conference / W. Boyd Rayward and Mary Ellen Bowden, (Eds.). Medford, N. J.: Information Today, 2004, Volltext im Netz unter <http://www.chemheritage.org/events/asist2002/proceedings.html>

⁶ Näheres zu deren Entwicklung siehe bei Charles P. Bourne, Trudi Bellardo Hahn: A history of online information services, 1963-1976. Cambridge, Mass.: MIT Press, 2003.

In mancher Hinsicht kann das Handbuch auch als eine „versteckte“ Aktualisierung von Teilen der Encyclopedia of computational chemistry (ed.-in-chief: Paul von Ragué Schleyer, Chichester: Wiley, 1998) angesehen werden. Dies zeigt, dass die Abgrenzung beider Gebiete nicht ganz einfach ist. Chemoinformatik kann als Teil der Computerchemie angesehen werden, deren Ursprünge ja in der Berechnung quantenchemischer (Molekül-) Modelle mit Datenverarbeitungsanlagen liegen.⁷

Zum Lehrbuch

Die Suche nach Objekten in einer Datenbank erfordert eine maschinenlesbare Repräsentation dieser Literaturzitate oder chemischen Stoffe. So beginnt das Lehrbuch mit einem Kapitel zur Repräsentation chemischer Verbindungen (Nomenklatur, Notationen, Verknüpfungstabellen als Darstellung der Struktur von Verbindungen). Dies geht bis zu modernen elektronischen Zeichenprogrammen für Moleküle und den dreidimensionalen Darstellungen auch komplexer Moleküle, wie z.B. Proteinen, die besonders in der Bioinformatik eine Rolle spielen. Watson und Crick nutzten zur Darstellung der Strukturen der DNA noch reale Molekülmodelle. Die elektronische Darstellung von Molekülen als Partner bei chemischen Reaktionen ist Voraussetzung für die Suche in chemischen Reaktionsdatenbanken und wird in einem eigenen Kapitel beschrieben.

Den chemischen Informationsspezialisten interessieren vor allem die Kapitel 5 „Databases and data sources in chemistry“ und 6 „Searching chemical structures“. Die restlichen Kapitel des Lehrbuches sind dann der Berechnung chemischer Daten sowie von Strukturparametern gewidmet. Ausserdem werden differenziert Methoden der Datenanalyse und Anwendungen - z.B. zur Modellierung von Stoffeigenschaften, chemischen Reaktionen und Arzneimittelsynthesen - beschrieben. Kurze Abschnitte führen in Grundlagen maschinellen Lernens, genetischer Algorithmen und des Data Mining ein. Deutlich wird, dass chemische Forschung ohne den Einsatz von Methoden der Informatik kaum noch auskommt.

Das fünfte Kapitel enthält einen zusammenfassenden Überblick über alle wichtigen chemischen Datenbanken. Eingebettet sind kleine spezielle Tutorials, z. B. zur Nutzung des Systems der Chemical Abstracts, zur Nutzung der Beilstein-Datenbank oder zur Suche im Internet bzw. zur Suche nach Umweltinformationen.

Ergänzend sollten hier auf jeden Fall auch Kapitel des „Handbooks“ konsultiert werden, die von renommierten, auch internationalen Fachleuten geschrieben wurden: Gary Wiggins gibt einen „Overview of Databases“ (S. 496-506), Andreas Barth schreibt über bibliographische Datenbanken (S. 507-522), William Fisanick und Eric R. Shively über das „CAS Information System“ (S. 556-607). Auch umfangreiche Kapitel zu speziellen Datenbanken, zu Beilstein (Alexander J. Lawson), zu Reaktionsdatenbanken (Engelbert Zass), zur Umweltinformation (Kristine Voigt), zu Patenten und Biodatenbanken sowie zur Chemie im Internet finden sich im Handbook, dessen Preis allerdings dazu führt, dass es sicher nicht in zu vielen Bibliotheken vorhanden sein wird. Das Lehrbuch ist für den an hohe Monographie-Preise mittlerweile leider gewohnten Chemiker, der sich mit der Chemoinformatik beschäftigen will, eine gute Alternative.

Lehrbuch und Handbuch demonstrieren den Umfang heutiger Chemie-Information, die sich längst beträchtlich von der reinen Suche nach chemischen Informationen verabschiedet hat. Beide Werke sind nötig, um auf diesem komplexen und sich schnell weiterentwickelnden Gebiet den Überblick zu behalten. Eine chemisch orientierte Fachbibliothek sollte beide ihren Kunden anbieten.

Thomas Hapke

⁷ Siehe auch Peyerimhoff, Sigrid D.: The development of computational chemistry in Germany. Reviews in Computational Chemistry 18 (2002) 257-291.