

Databáze v chemické a forenzní analýze

Tereza Uhlíková, tereza.uhlikova@vscht.cz

March 28, 2025

Tereza Uhlíková

Ústav analytické chemie

skupina teoretické spektroskopie

místnost A277

<https://web.vscht.cz/~uhlikovt/>

tereza.uhlikova@vscht.cz

O čem to bude

		Téma přednášky	Téma cvičení
1	11.3.	Opakování a seznamování	Hledání odpovědí pomocí AI
2	18.3.	Princip AI	AI databázové nástroje
3	25.3.	Chemické databáze	Analýza spekter pomocí databází
4	1.4.	Forenzní databáze	zkoušková témata
5	8.4.	společná databáze volně prodejných psych. látek ERA	
6	15.4.	společná databáze volně prodejných psych. látek VY a AI	
7	22.4.		
8	29.4.		
9	6.5.	ČTVRTEČNÍ ROZVRH = NENÍ PŘEDNÁŠKA	
10	...	Zkouška	

Základní chemická databáze

Periodic Table of the Elements

1 IA 1A	2 IIA 2A												13 IIIA 3A	14 IVA 4A	15 VA 5A	16 VIA 6A	17 VIIA 7A	18 VIIIA 8A
1 H Hydrogen 1.008													5 B Boron 10.811	6 C Carbon 12.011	7 N Nitrogen 14.007	8 O Oxygen 15.999	9 F Fluorine 18.998	10 Ne Neon 20.180
3 Li Lithium 6.941	4 Be Beryllium 9.012												13 Al Aluminum 26.982	14 Si Silicon 28.086	15 P Phosphorus 30.974	16 S Sulfur 32.066	17 Cl Chlorine 35.453	18 Ar Argon 39.948
11 Na Sodium 22.990	12 Mg Magnesium 24.305	3 IIIB 3B	4 IVB 4B	5 VB 5B	6 VIB 6B	7 VIIB 7B	8 VIII 8	9 VIII 8	10 VIII 8	11 IB 1B	12 IIB 2B		31 Ga Gallium 69.732	32 Ge Germanium 72.61	33 As Arsenic 74.922	34 Se Selenium 78.09	35 Br Bromine 79.904	36 Kr Krypton 84.80
19 K Potassium 39.098	20 Ca Calcium 40.078	21 Sc Scandium 44.956	22 Ti Titanium 47.88	23 V Vanadium 50.942	24 Cr Chromium 51.996	25 Mn Manganese 54.938	26 Fe Iron 55.933	27 Co Cobalt 58.933	28 Ni Nickel 58.693	29 Cu Copper 63.546	30 Zn Zinc 65.39		49 In Indium 114.818	50 Sn Tin 118.71	51 Sb Antimony 121.760	52 Te Tellurium 127.6	53 I Iodine 126.904	54 Xe Xenon 131.29
37 Rb Rubidium 84.458	38 Sr Strontium 87.62	39 Y Yttrium 88.906	40 Zr Zirconium 91.224	41 Nb Niobium 92.906	42 Mo Molybdenum 95.94	43 Tc Technetium 98.907	44 Ru Ruthenium 101.07	45 Rh Rhodium 102.906	46 Pd Palladium 106.42	47 Ag Silver 107.868	48 Cd Cadmium 112.411		81 Tl Thallium 204.383	82 Pb Lead 207.2	83 Bi Bismuth 208.980	84 Po Polonium [209, 92]	85 At Astatine 209, 97	86 Rn Radon 222, 018
55 Cs Cesium 132.905	56 Ba Barium 137.327	57-71 Lanthanide Series	72 Hf Hafnium 178.49	73 Ta Tantalum 180.948	74 W Tungsten 183.85	75 Re Rhenium 186.207	76 Os Osmium 190.23	77 Ir Iridium 192.22	78 Pt Platinum 195.08	79 Au Gold 196.967	80 Hg Mercury 200.59		113 Uut Ununtrium unknown	114 Fl Flerovium [289]	115 Uup Ununpentium unknown	116 Lv Livermorium [293]	117 Uus Ununseptium unknown	118 Uuo Ununoctium unknown
87 Fr Francium 223, 029	88 Ra Radium 226, 025	89-103 Actinide Series	104 Rf Rutherfordium [261]	105 Db Dubnium [262]	106 Sg Seaborgium [266]	107 Bh Bohrium [264]	108 Hs Hassium [269]	109 Mt Meitnerium [268]	110 Ds Darmstadtium [269]	111 Rg Roentgenium [272]	112 Cn Copernicium [277]							
			57 La Lanthanum 138.905	58 Ce Cerium 140.115	59 Pr Praseodymium 140.908	60 Nd Neodymium 144.24	61 Pm Promethium 144.913	62 Sm Samarium 150.36	63 Eu Europium 151.966	64 Gd Gadolinium 157.25	65 Tb Terbium 158.925	66 Dy Dysprosium 162.50	67 Ho Holmium 164.930	68 Er Erbium 167.26	69 Tm Thulium 168.934	70 Yb Ytterbium 173.04	71 Lu Lutetium 174.967	
			89 Ac Actinium 227.028	90 Th Thorium 232.038	91 Pa Protactinium 231.036	92 U Uranium 238.029	93 Np Neptunium 237.048	94 Pu Plutonium 244.064	95 Am Americium 243.061	96 Cm Curium 247.070	97 Bk Berkelium 247.070	98 Cf Californium 251.080	99 Es Einsteinium [254]	100 Fm Fermium 257.095	101 Md Mendelevium 258.1	102 No Nobelium 259.101	103 Lr Lawrencium [262]	
			Alkali Metal	Alkaline Earth	Transition Metal	Semimetal	Nonmetal	Basic Metal	Halogen	Noble Gas	Lanthanide	Actinide						

© 2013 Todd Helmenstein
chemistry.about.com
scienceonline.org

Jak vyhledávat - obecně

- Pro všechny platí - nejdůležitější je si přečíst **návod**

(jak se používá, jaké jsou jednotky, jak se značí přechody, jaké je rozlišení, v čem se měřilo)

- rozmyslet si **otázku**, co vlastně hledám

(informací je mnoho, ale která je ta správná) – HITRAN x Cologne – přesné linie x sp. konstanty

- použít **různé** přístupy/otázky

(jméno amoniak x čpavek, frekvence přechodu – rozmezí, intenzita, disociace)

- dohledání **autora**, zdroje (citace u spekter)
- používat **různé** databáze

(od různých lidí, z různých zemí)

- **obnovené** (přeměřením spektra se přechod přiřadí jinému kv. číslu NO_3) - ale i **staré** verze (čím více dat tím více chyb)
- spousty, ale dost se překrývají (HITRAN x GEISA)
- jedna centrální neexistuje
- někdy lze použít i nepublikované

High-resolution TRANsmission molecular absorption database

<https://hitran.org/>

- Nejen soupis molekul, ale i jejich vlastností – specializované
- Části, aneb co je důležité pro atmosféru:
 - 1 soupis linií z vysoce rozlišené IR a mikrovlnné spektroskopie + intenzity
 - 2 infračervené absorpční průřezy (pro velmi hustá spektra)
 - 3 soupis linií a průřezů pro UV oblast
 - 4 index lomu aerosolů
 - 5 srážková absorpce
 - 6 obecná data a software pro filtrování a kompilaci

- Kdo ji tvoří?

Dr Laurence S. Rothman - zakladatel v 60-tých letech

- Proč ji tvoří?

chemické složení a možné reakce v atmosféře a ve vesmíru

- Komu slouží?

všem...

- Kolik stojí? & Kolik pracujících lidí?

různé laboratoře (desítky) dodávají naměřená spektra

- Jací lidé jsou potřeba?

odborní spektroskopici - experimentátoři i teoretici + programátoři
databází

- Kolik času?

- Jak je veliká?

celá databáze má velikost cca 10 GB

- Kolik molekul?

55 molekul a jejich izotopologů

- kolik experimentálních linií?
- kolik predikcí?

Jednotky (nejen SI, opět specializované)

Měřeno ve vakuu, vzduchu (popř. vodě, rozpouštědle, tuhé fázi, jako pára)

Jaká chyba měření

...

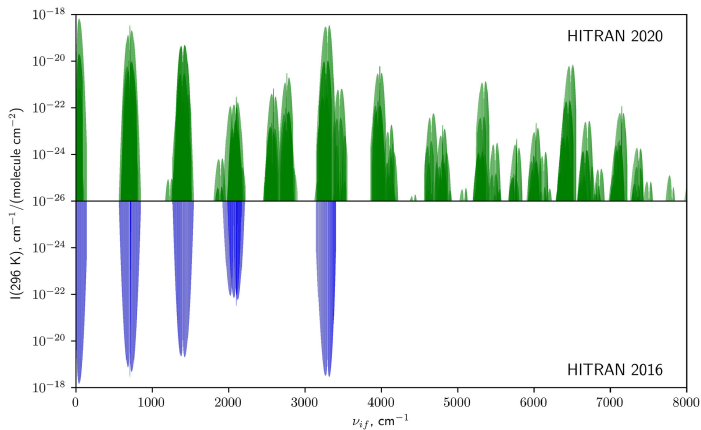
Updated bands for $^{16}\text{O}_3$.

Band	Number of lines	Spectral range (cm^{-1})	S_ν	References for ν , S
022–000	1336	3297.46–3478.53	9.936	[71,71]
121–000	1611	3383.04–3483.38	62.720	[71,71]
113–100	658	3490.53–3565.76	4.038	[72,71]
014–001	1136	3520.70–3605.15	12.251	[72,71]
014–100	13	3533.85–3562.08	0.029	[72,71]
113–001	12	3543.34–3604.91	0.036	[72,71]
213–000 ^a	954	5447.73–5526.30	9.627	[71,71]

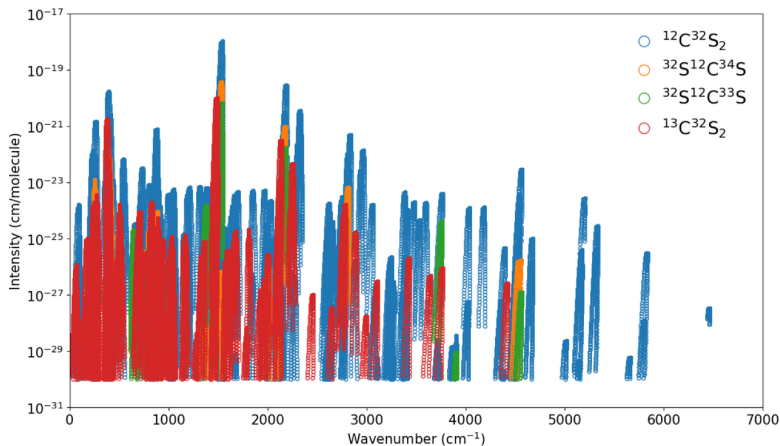
Note: S_ν is the sum of line intensities in units of $10^{-23} \text{ cm}^{-1}/(\text{molecule cm}^{-2})$ at 296 K for the corresponding bands included in the line list, ν is the line position, and S is the line intensity.

^a This band was labeled as 015–000 in the previous version of HITRAN.

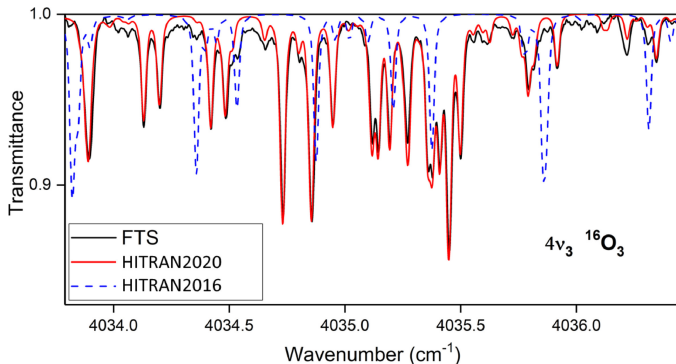
HITRAN data



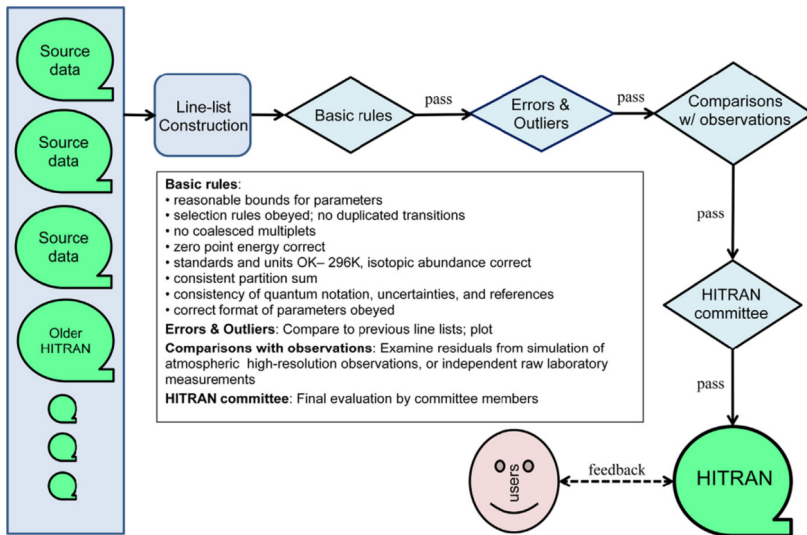
HITRAN data



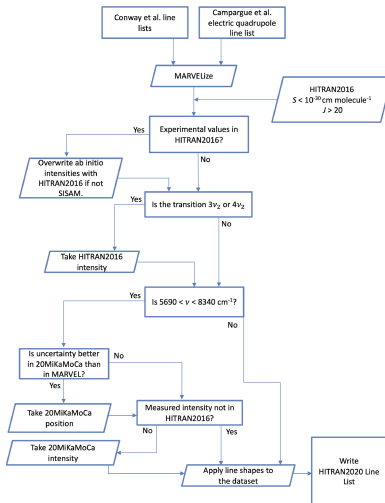
Výpočet transmittance v okolí 4000 cm^{-1} použitím S&MPO_20d (HITRAN2020). Srovnání výpočtu predikcí HITRAN2016 A HITRAN2020 a experimentu.

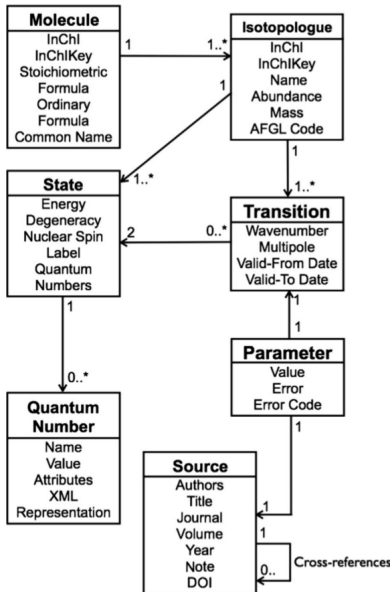


HITRAN vnější mašinérie



Získání intenzit konzistentních s Hitranem 2020





CDMS - Koeln

Molecules in the Interstellar Medium or Circumstellar Shells

<https://cdms.astro.uni-koeln.de/classic/predictions/daten/Cyanamide/>

https://cdms.astro.uni-koeln.de/classic/predictions/pdfs/CDMS_AA.pdf

NIST - National Institute of Standards and Technology

<http://www.nist.gov/>

Stránky celého institutu pro standard – nejen databáze, dělají i výzkum a vše standardizují - základní veličiny, podle čeho se má měřit a kalibrovat přístroje

Otevřená (prozatím), placená z daní USA, databáze v rámci jednotlivých ústavů
2017 rozpočet byl \$1 000 million $\times 25 = 25$ miliard Kč CR má 1 250.9 miliard Kč
<https://www.nist.gov/about-nist/our-organization/budget-planning>

Seznam volně přístupných databází v NISTu

Free Standard Reference Databases by SRD Number

<http://www.nist.gov/srd/onlinelist.cfm>

Jejich popis

<http://srdata.nist.gov/gateway/gateway?dblist=1>

Nás jako chemiky zajímá zejména chemická spektroskopická data

<http://webbook.nist.gov/>

Databáze obsahuje zejména organické sloučeniny a několik malých anorganických, pouze takové pro které mají v NIST data.

Forenzní portál

<http://www.nist.gov/forensics-portal.cfm>

<http://webbook.nist.gov/>

- Termochemická data pro více jak 7000 organických sloučenin
- IČ spektra pro 16000 Hmotnostní spektra pro 33000
- UV/Vis spektra pro 1600
- Data pro plynovou chromatografii pro 27000
- Electronická a vibrační spektra pro 5000
- Spektroskopické konstanty pro 600 diatomik (Herzberg - 1960)
- Ionizační energie pro 16000
- Termofyzikální vlastnosti pro 74 kapalin

<http://webbook.nist.gov/chemistry>

Možnosti hledání: přímé

Vzorec

Jméno

katalogové číslo

Ionizační energie

Elektronová afinita

Protonová afinita

Kyselost

Energie vzniku produktů

Vibrační energie a Elektronická energie

Nakreslená struktura Struktura v souboru

Třída struktur

Molekulová hmotnost

Reakce

Autor

Možnosti hledání: nepřímé

Doklikat se

- Vzorec – nezávisí na pořadí, mezerách, velikostech $nO \times No$,? - jakýkoli atom, možnost přidat libovolný atom (opatrně), izotopolog, větší počet atomů
- Jméno - nezávisí na pořadí, mezerách, velikostech,* - neúplné
- katalogové číslo
- Ionová energie
 - Ionizační energie – eV, rozmezí (není pro všechny sloučeniny)
 - Elektronová afinita – stejné jako IE
 - Protonová afinita – stejné jako IE
 - Kyselost – definice – gibbspva volná energie
 - Energie vzniku produktů – struktura často není známa—pouze vzorec
- Vibrační a elektronová spektra
 - Vibrační energie - vlnočet
 - Elektronická energie

- Struktura
 - Applet Based Structure Search - přímo se nakreslí
 - File Based Structure Search - nahrát soubor
 - Structure Class Search - charakterizuje se struktúra podle vazeb, atomů, kruhů...
- Molekulová hmotnost - nejvíce zastoupený izotop
- Reakce - přímá reakce
- Autor - reference

Jednotky – SI nebo založené na kalorii (kalorie a atmosféra místo Joule a bar)

Typy dat, která se obdrží: - když se nic nezvolí, zobrazí se pouze základní informace

- temochemická data

- plyná fáze
- kondenzovaná fáze
- fázová přeměna
- reakce
- ionizační energie
- energie pro tvoření iontových klastrů

- ostatní data

- IR spektra
- hmotnostní spektra
- UV/Vis spektra
- vibrační a elektronová spektra
- konstanty dvoutatomových molekul - vibrační, rotační a rotačně-vibrační
- Henryho zákon - udává souvislost parciálního tlaku páry dané látky nad roztokem a jejího podílu v tomto roztoku

Jak data vypadají

<http://webbook.nist.gov/chemistry/form-ser.html>

Tabulky: společné rysy – více kodů (text, html, pdf, ascii), reference, komentáře

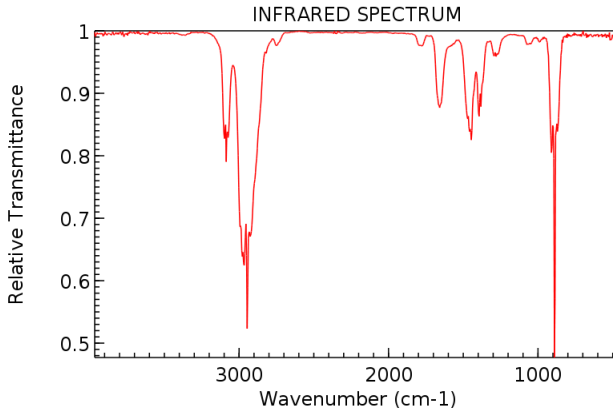
Grafy: Java

Spektra: online x applet

Vyzkoušíme C₆H₆ – není pouze benzen

[https://webbook.nist.gov/cgi/cbook.cgi?Formula=C₆H₆&NoIon=on&Units=SI](https://webbook.nist.gov/cgi/cbook.cgi?Formula=C6H6&NoIon=on&Units=SI)

Čemu náleží toto spektrum - první pík je 3086 cm^{-1}



NIST Chemistry WebBook (<http://webbook.nist.gov/chemistry>)

Atomová databáze <http://www.nist.gov/pml/data/asd.cfm>

jsou dány ve viditelné oblasti tyto linie v nm ve vzduchu,

648.14 |

648.44 |

651.73 |

653.00 |

655.12 |

Jaké mu prvku náleží a vygenerujte celé spektrum

SpecDB = Specification Database Design

<https://britastro.org/specdb/>

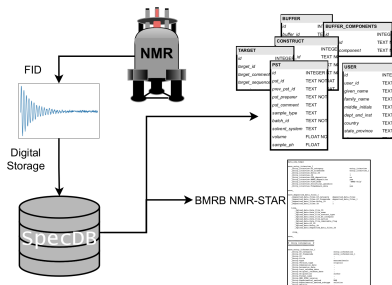
<https://github.com/markasoftware/SpecDB>

<https://specdb.readthedocs.io/en/latest/>

<https://ui.adsabs.harvard.edu/abs/2018AAS...23136215K/abstract>

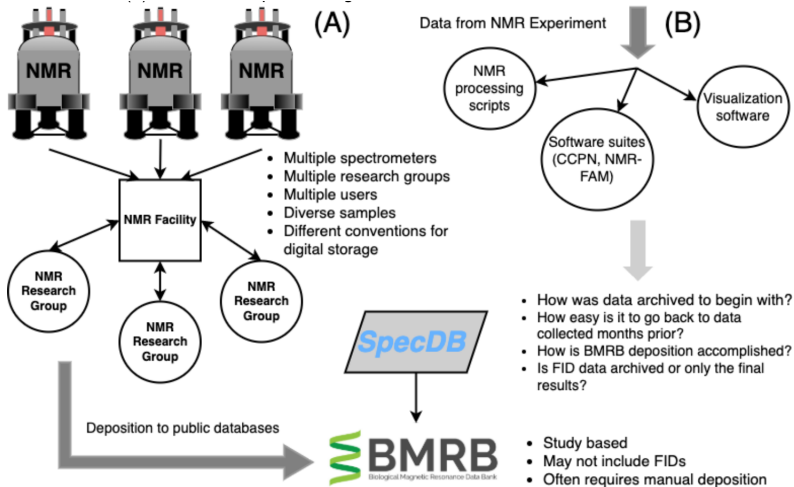
https://link.springer.com/chapter/10.1007/11676935_29

<https://doi.org/10.1016/j.jmr.2022.107268>

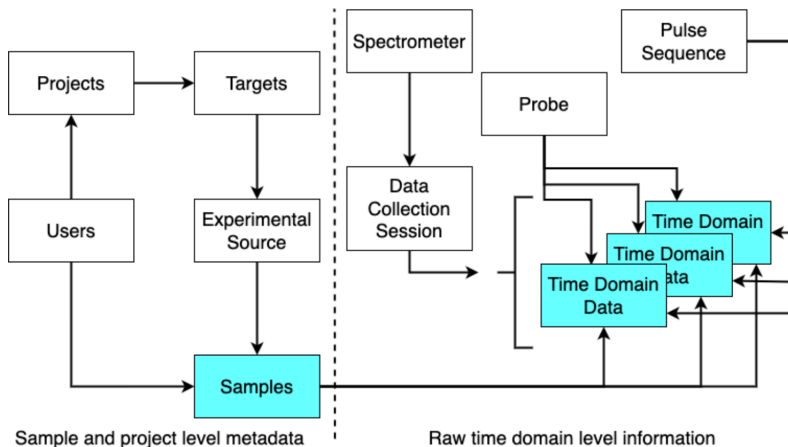


<https://doi.org/10.1016/j.jmr.2022.107268>

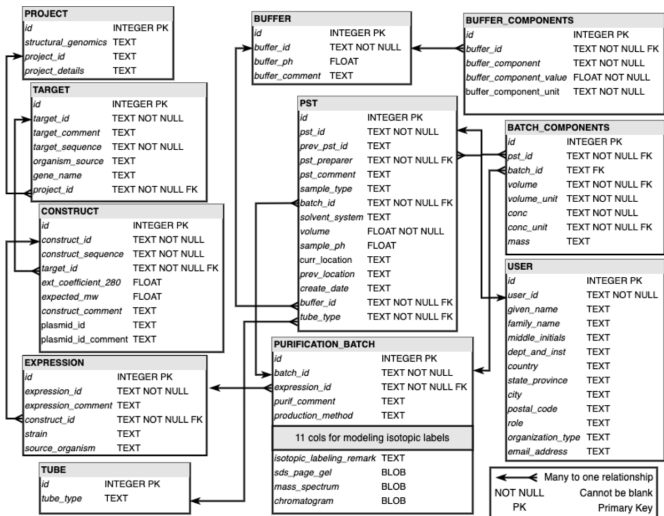
Data ecosystem



SpecDB dvoukřídlové schéma



Relační diagram I



Relační diagram II

