# Pharmaceutical Chemistry

## QSAR through the OCHEM portal


SAPIENZA
UNIVERSITÀ DI ROMA

# Introduction

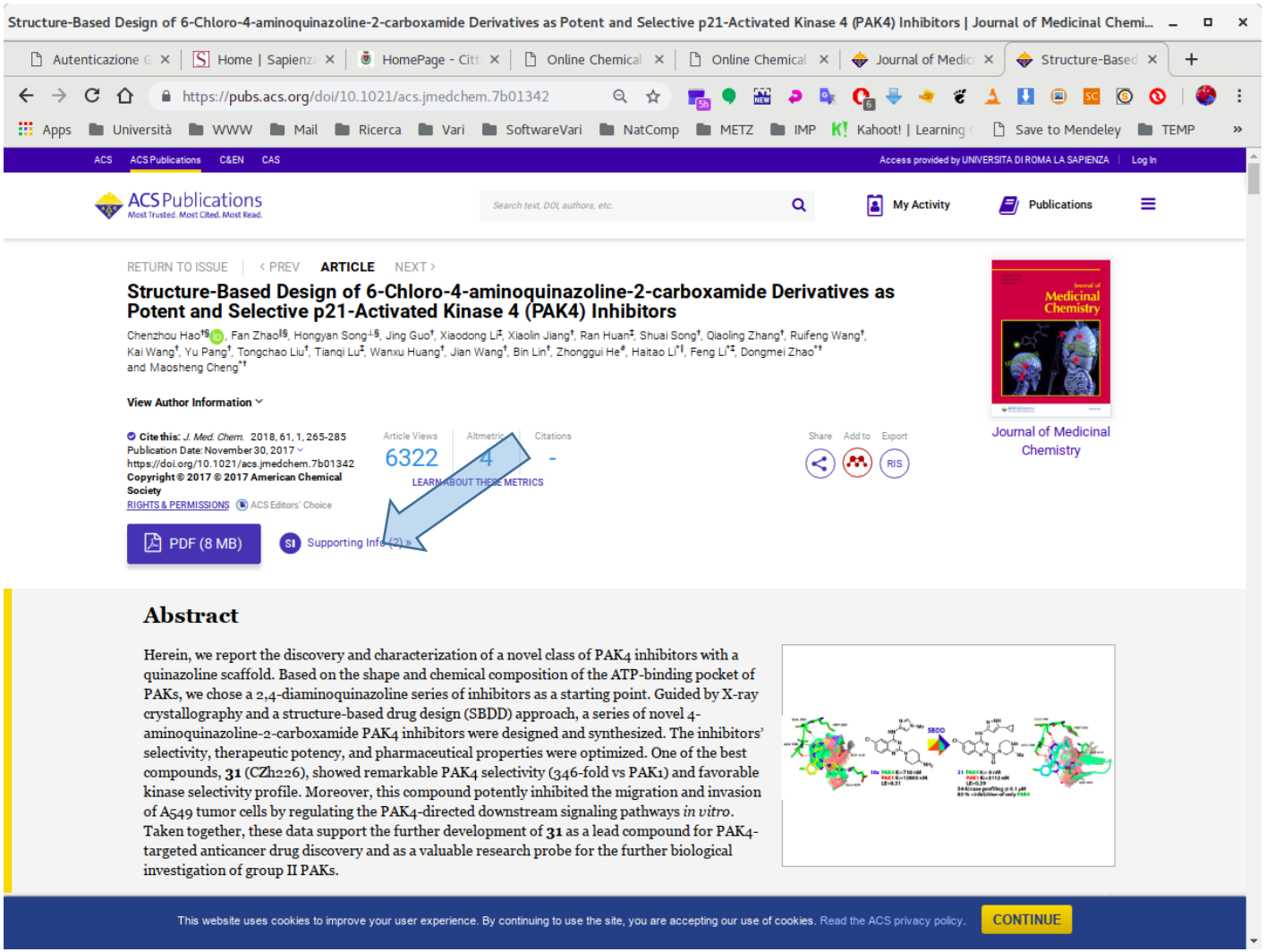
J Comput Aided Mol Des (2011) 25:533–554
DOI 10.1007/s10822-011-9440-2

**Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information**

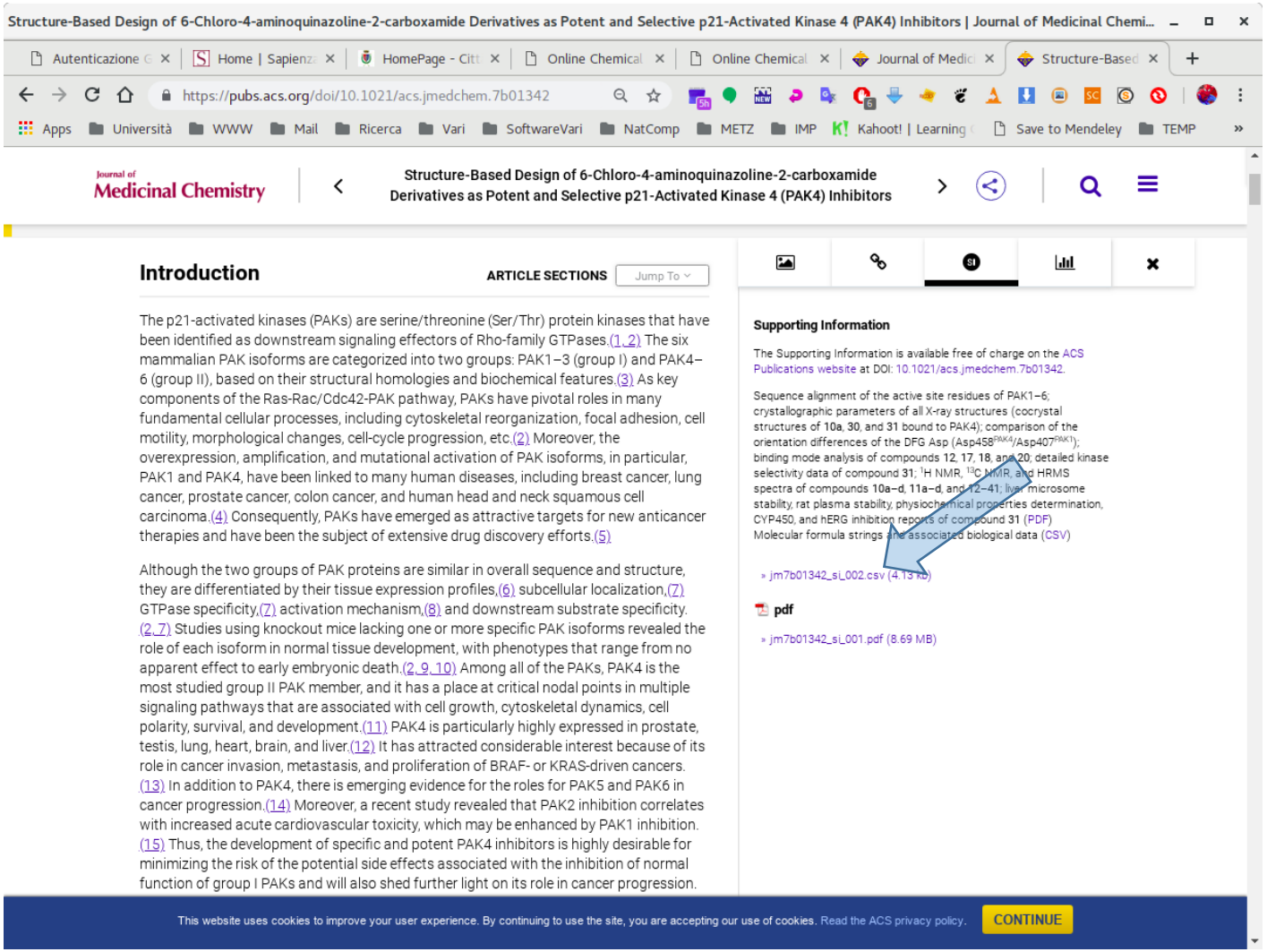
The Online Chemical Modeling Environment is a web-based platform that aims to automate and simplify the typical steps required for QSAR modeling. The platform consists of two major subsystems: the database of experimental measurements and the modeling framework. A user-contributed database contains a set of tools for easy input, search and modification of thousands of records. The OCHEM database is based on the wiki principle and focuses primarily on the quality and verifiability of the data. The database is tightly integrated with the modeling framework, which supports all the steps required to create a predictive model: data search, calculation and selection of a vast variety of molecular descriptors, application of machine learning methods, validation, analysis of the model and assessment of the applicability domain. As compared to other similar systems, OCHEM is not intended to re-implement the existing tools or models but rather to invite the original authors to contribute their results, make them publicly available, share them with other users and to become members of the growing research community. Our intention is to make OCHEM a widely used platform to perform the QSPR/QSAR studies online and share it with other users on the Web. The ultimate goal of OCHEM is collecting all possible chemoinformatics tools within one simple, reliable and user-friendly resource. The OCHEM is free for web users and it is available online at http://www.ochem.eu

.

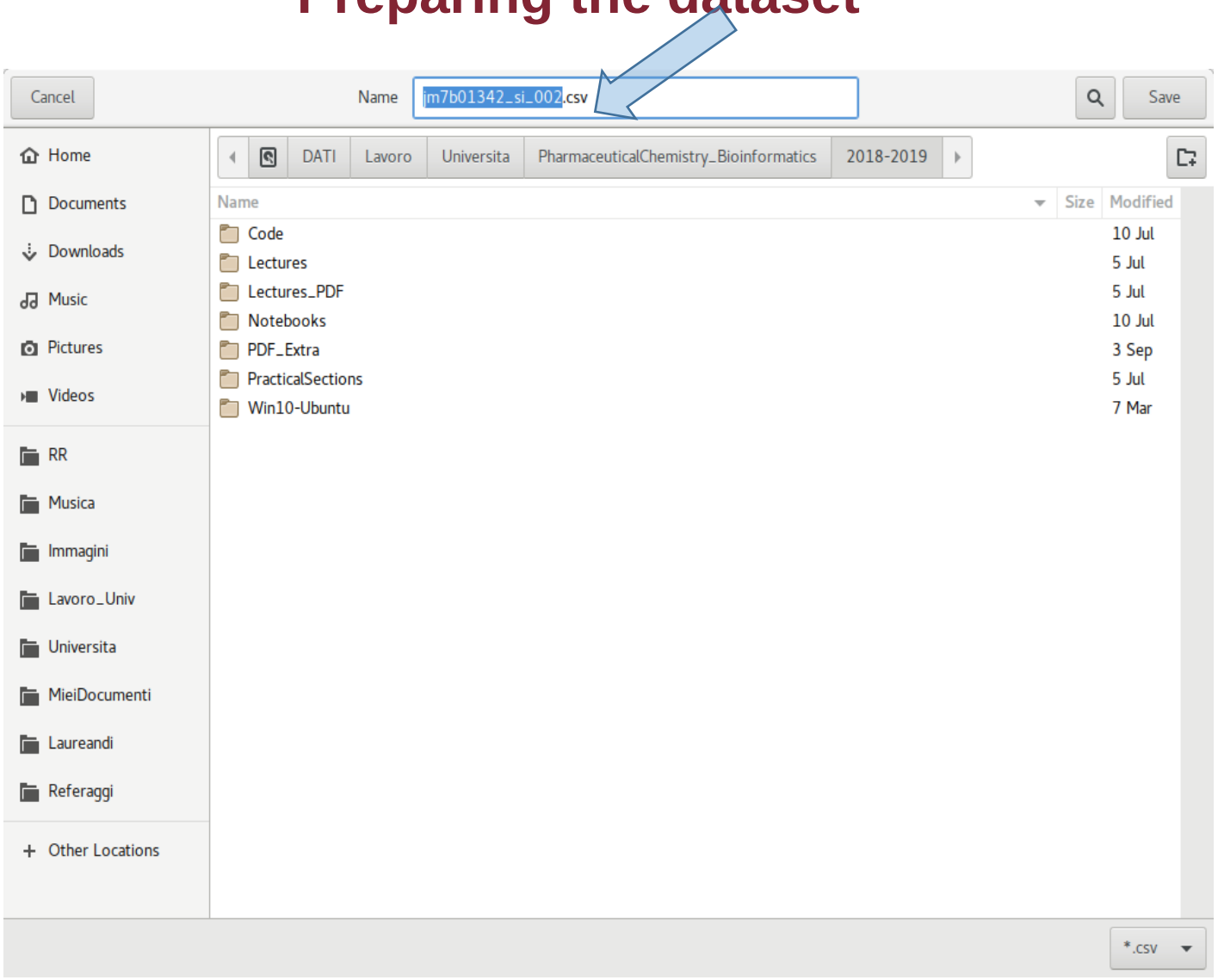
# Preparing the dataset



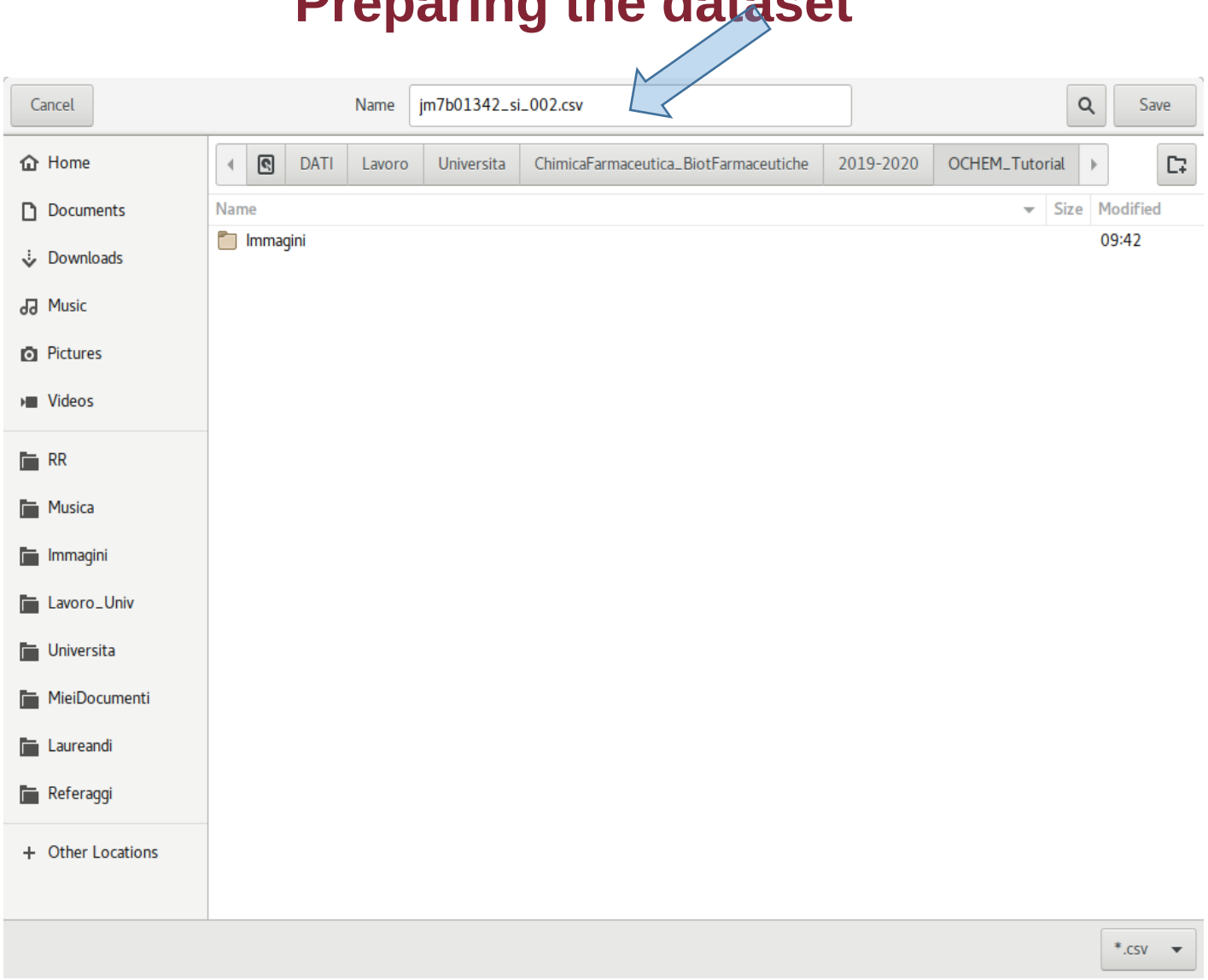
# Preparing the dataset



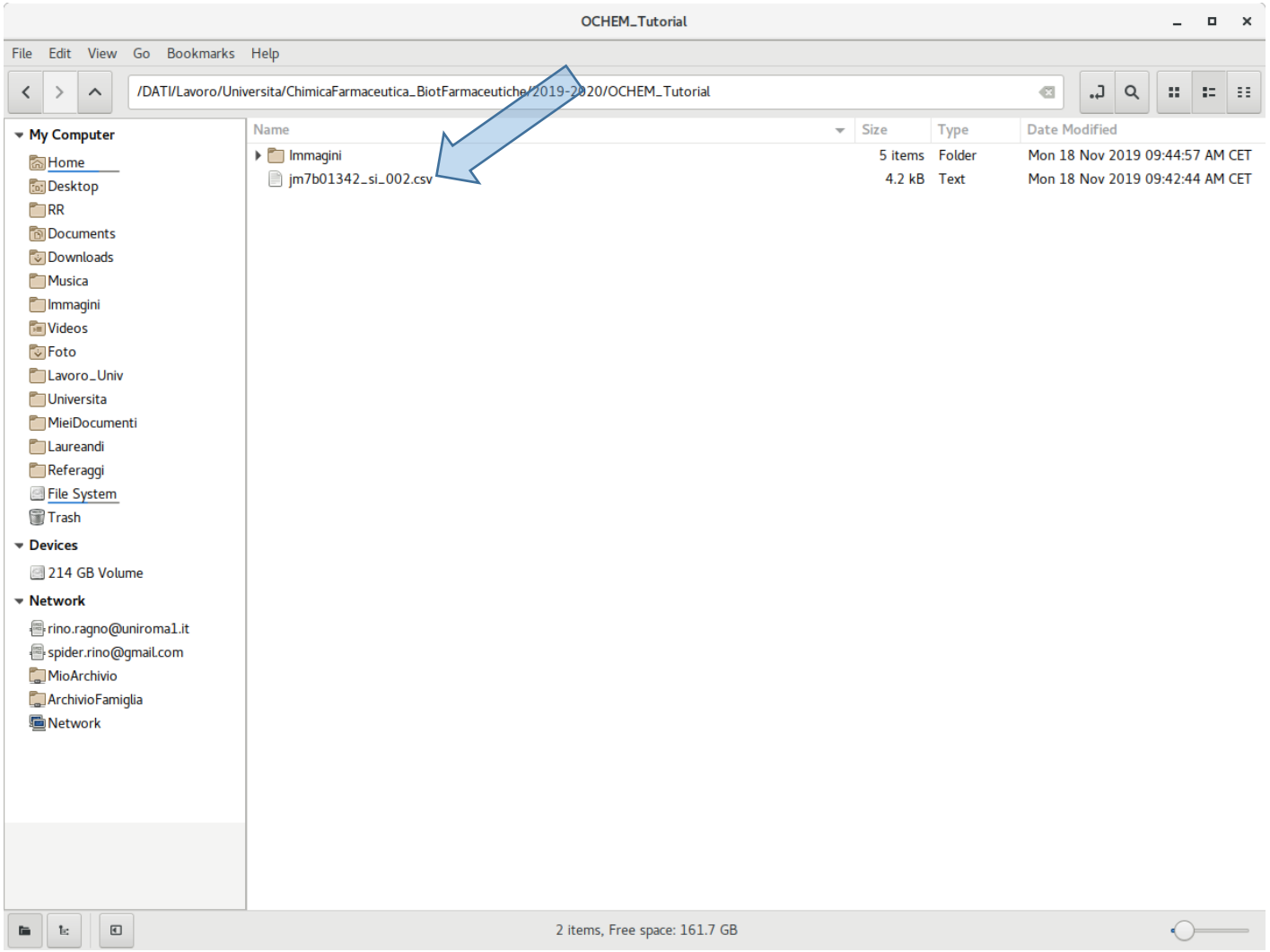
# Preparing the dataset



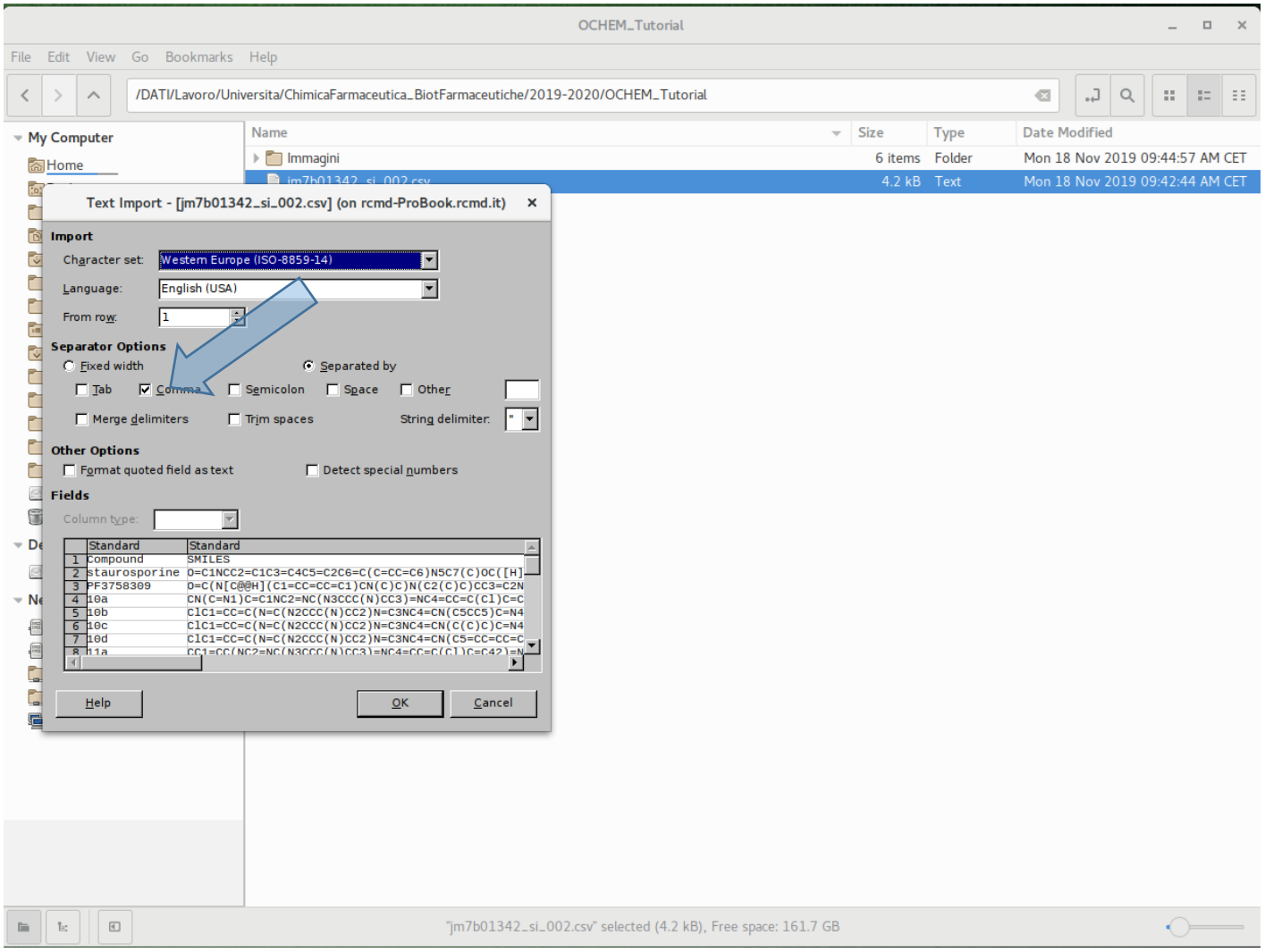
# Preparing the dataset



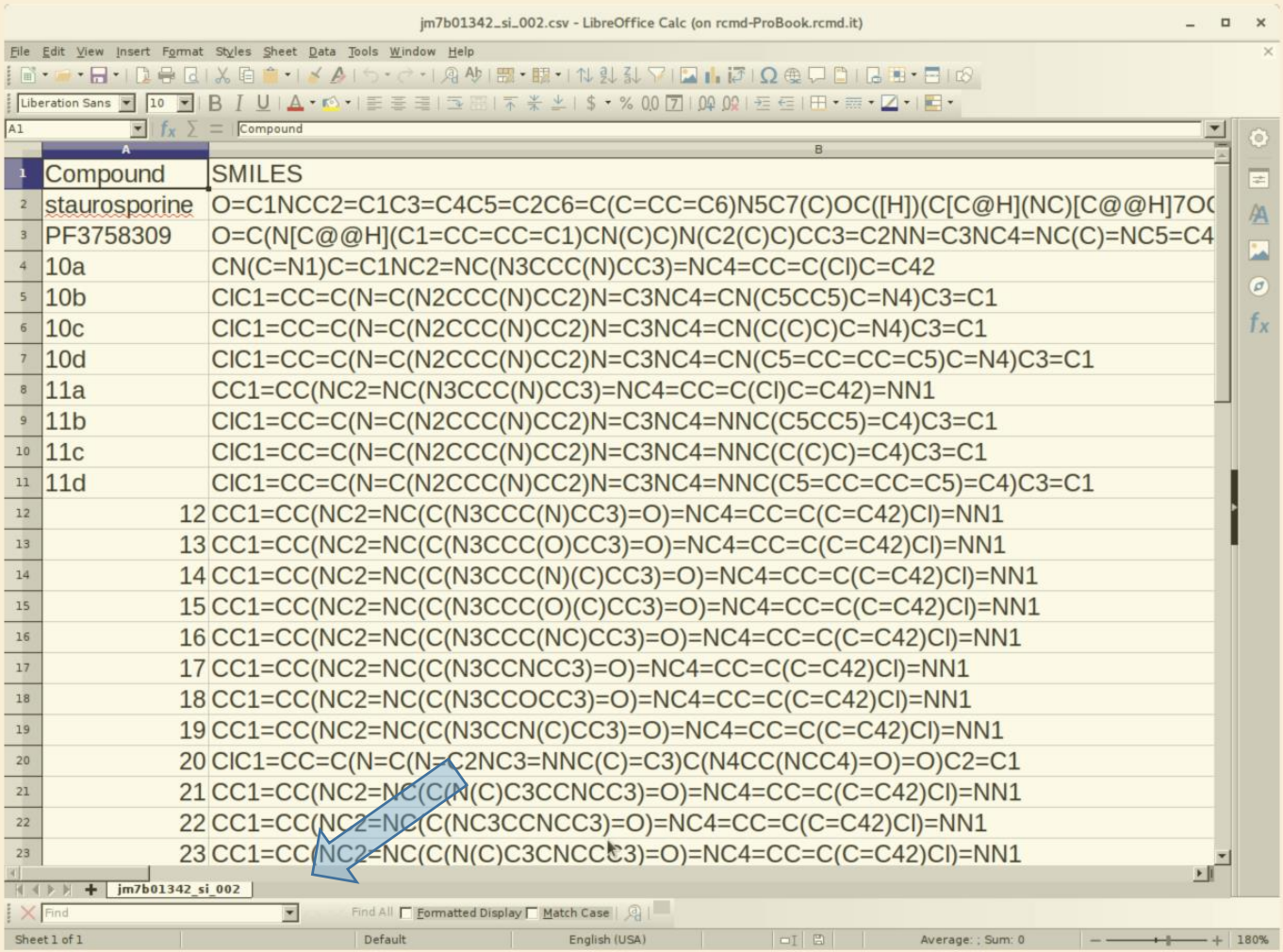
# Preparing the dataset



# Preparing the dataset

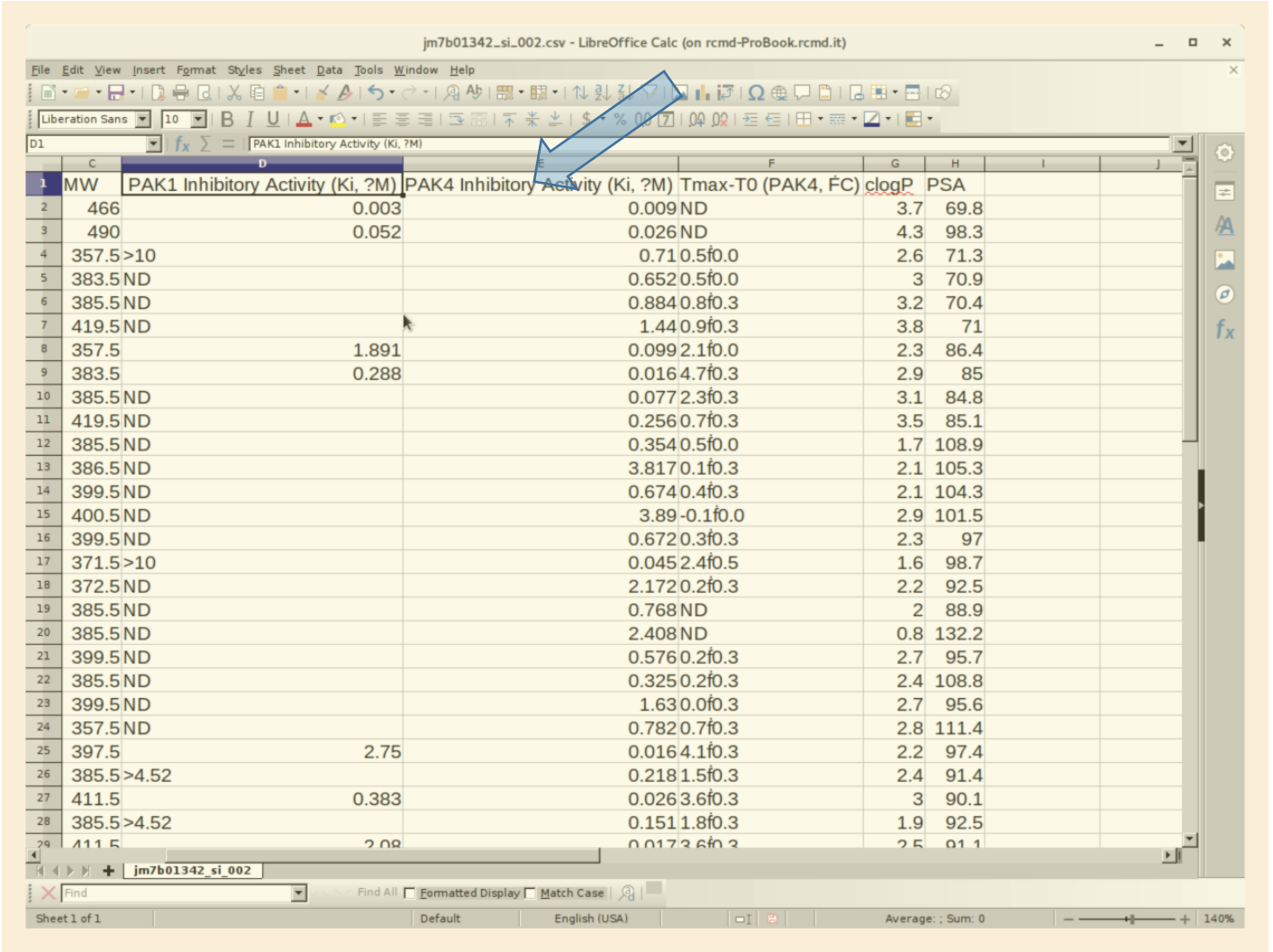


# Preparing the dataset



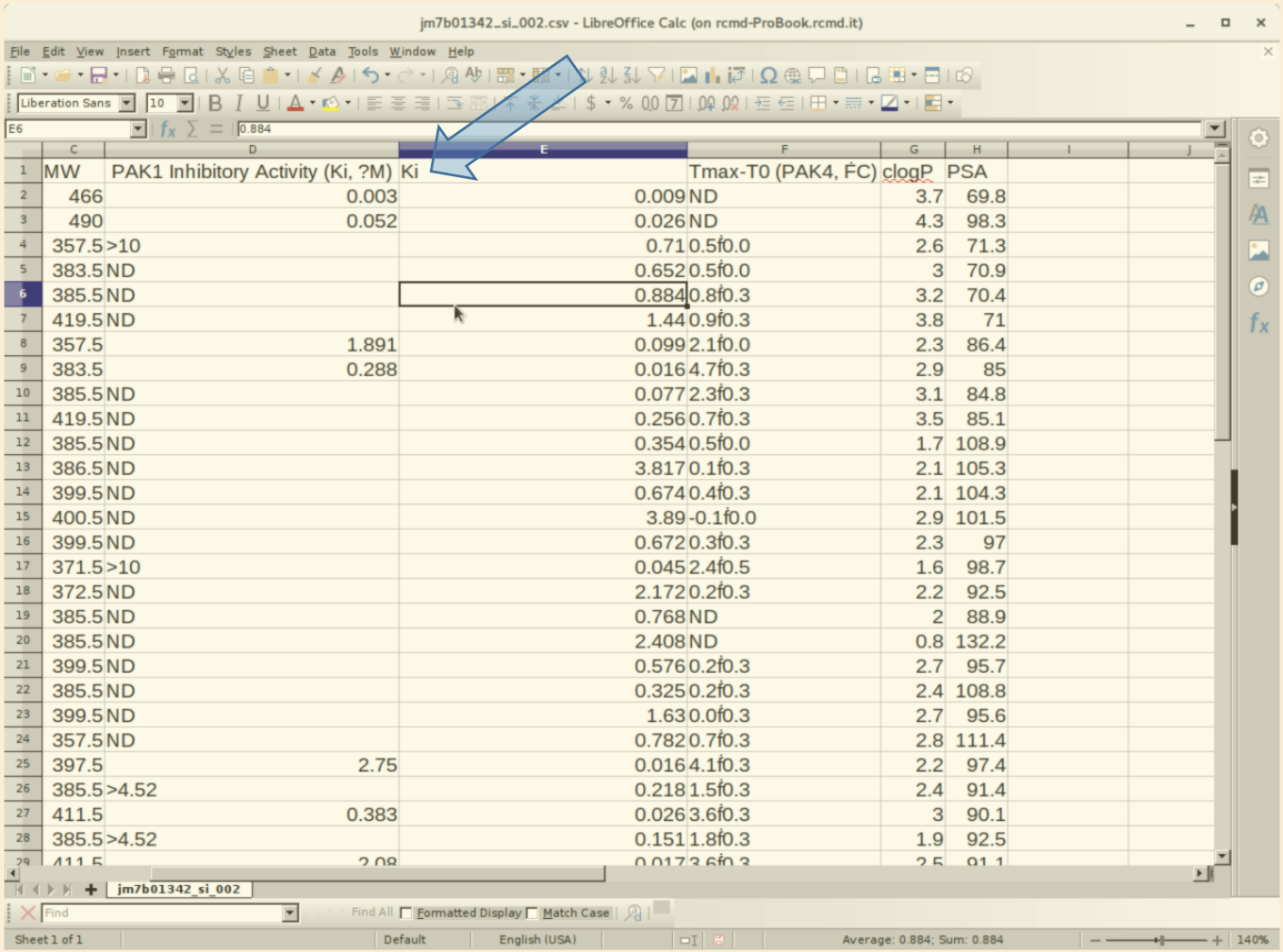

| | A | B |
|---|---|---|
| 1 | Compound | SMILES |
| 2 | staurosporine | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@H](NC)[C@@H]7OC |
| 3 | PF3758309 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3NC4=NC(C)=NC5=C4 |
| 4 | 10a | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42 |
| 5 | 10b | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C1 |
| 6 | 10c | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1 |
| 7 | 10d | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N4)C3=C1 |
| 8 | 11a | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42)=NN1 |
| 9 | 11b | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C1 |
| 10 | 11c | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1 |
| 11 | 11d | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C4)C3=C1 |
| 12 | 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 13 | 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 14 | 14 | CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 15 | 15 | CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 16 | 16 | CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 17 | 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 18 | 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 19 | 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 20 | 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2=C1 |
| 21 | 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 22 | 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |
| 23 | 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 |

# Preparing the dataset



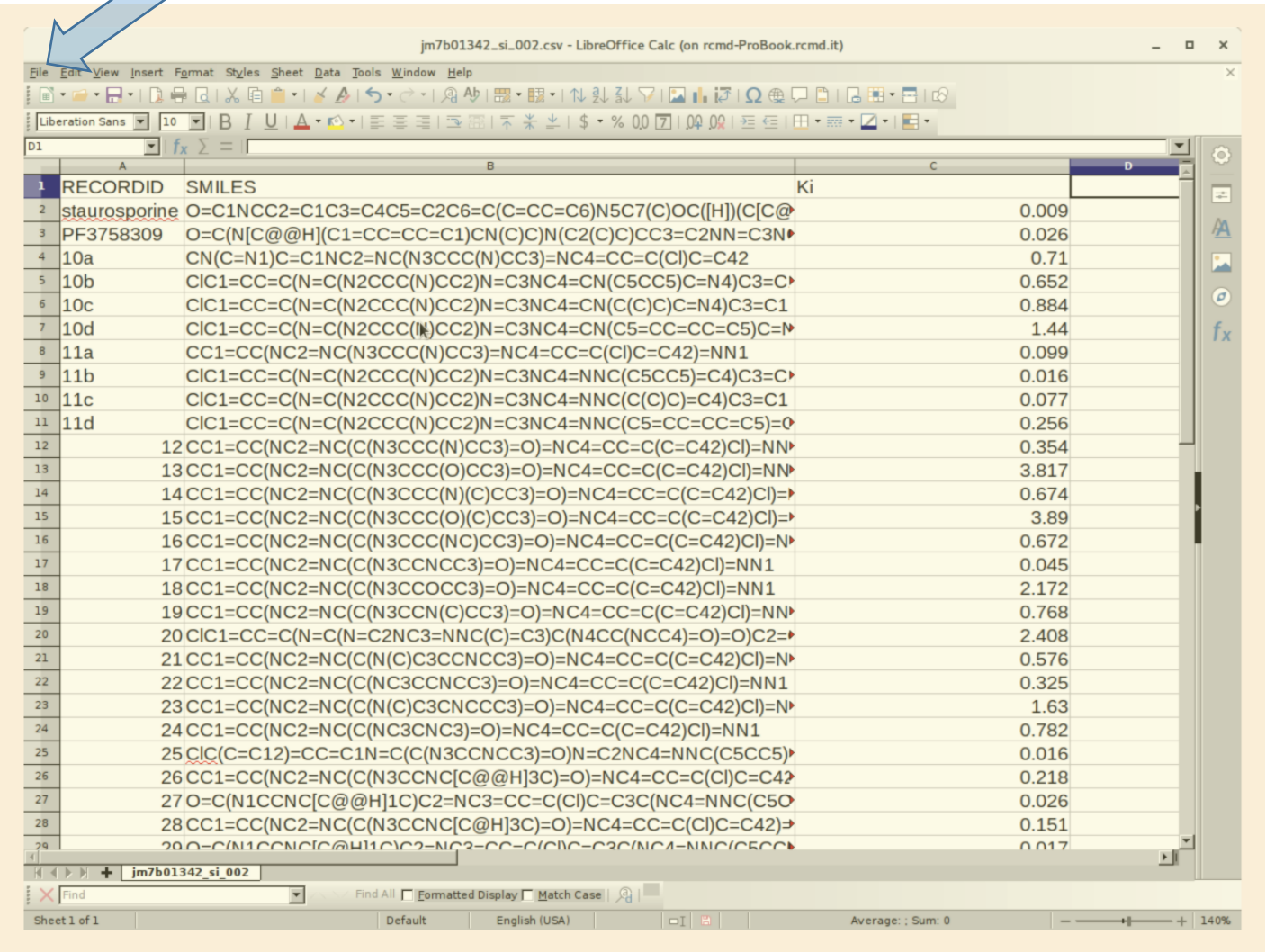

| | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|
| 1 | MW | PAK1 Inhibitory Activity (Ki, ?M) | PAK4 Inhibitory Activity (Ki, ?M) | Tmax-T0 (PAK4, FC) | clogP | PSA | | |
| 2 | 466 | 0.003 | 0.009 | ND | 3.7 | 69.8 | | |
| 3 | 490 | 0.052 | 0.026 | ND | 4.3 | 98.3 | | |
| 4 | 357.5 | >10 | 0.71 | 0.5±0.0 | 2.6 | 71.3 | | |
| 5 | 383.5 | ND | 0.652 | 0.5±0.0 | 3 | 70.9 | | |
| 6 | 385.5 | ND | 0.884 | 0.8±0.3 | 3.2 | 70.4 | | |
| 7 | 419.5 | ND | 1.44 | 0.9±0.3 | 3.8 | 71 | | |
| 8 | 357.5 | 1.891 | 0.099 | 2.1±0.0 | 2.3 | 86.4 | | |
| 9 | 383.5 | 0.288 | 0.016 | 4.7±0.3 | 2.9 | 85 | | |
| 10 | 385.5 | ND | 0.077 | 2.3±0.3 | 3.1 | 84.8 | | |
| 11 | 419.5 | ND | 0.256 | 0.7±0.3 | 3.5 | 85.1 | | |
| 12 | 385.5 | ND | 0.354 | 0.5±0.0 | 1.7 | 108.9 | | |
| 13 | 386.5 | ND | 3.817 | 0.1±0.3 | 2.1 | 105.3 | | |
| 14 | 399.5 | ND | 0.674 | 0.4±0.3 | 2.1 | 104.3 | | |
| 15 | 400.5 | ND | 3.89 | -0.1±0.0 | 2.9 | 101.5 | | |
| 16 | 399.5 | ND | 0.672 | 0.3±0.3 | 2.3 | 97 | | |
| 17 | 371.5 | >10 | 0.045 | 2.4±0.5 | 1.6 | 98.7 | | |
| 18 | 372.5 | ND | 2.172 | 0.2±0.3 | 2.2 | 92.5 | | |
| 19 | 385.5 | ND | 0.768 | ND | 2 | 88.9 | | |
| 20 | 385.5 | ND | 2.408 | ND | 0.8 | 132.2 | | |
| 21 | 399.5 | ND | 0.576 | 0.2±0.3 | 2.7 | 95.7 | | |
| 22 | 385.5 | ND | 0.325 | 0.2±0.3 | 2.4 | 108.8 | | |
| 23 | 399.5 | ND | 1.63 | 0.0±0.3 | 2.7 | 95.6 | | |
| 24 | 357.5 | ND | 0.782 | 0.7±0.3 | 2.8 | 111.4 | | |
| 25 | 397.5 | 2.75 | 0.016 | 4.1±0.3 | 2.2 | 97.4 | | |
| 26 | 385.5 | >4.52 | 0.218 | 1.5±0.3 | 2.4 | 91.4 | | |
| 27 | 411.5 | 0.383 | 0.026 | 3.6±0.3 | 3 | 90.1 | | |
| 28 | 385.5 | >4.52 | 0.151 | 1.8±0.3 | 1.9 | 92.5 | | |
| 29 | 411.5 | 2.08 | 0.017 | 3.6±0.3 | 2.5 | 91.1 | | |

# Preparing the dataset



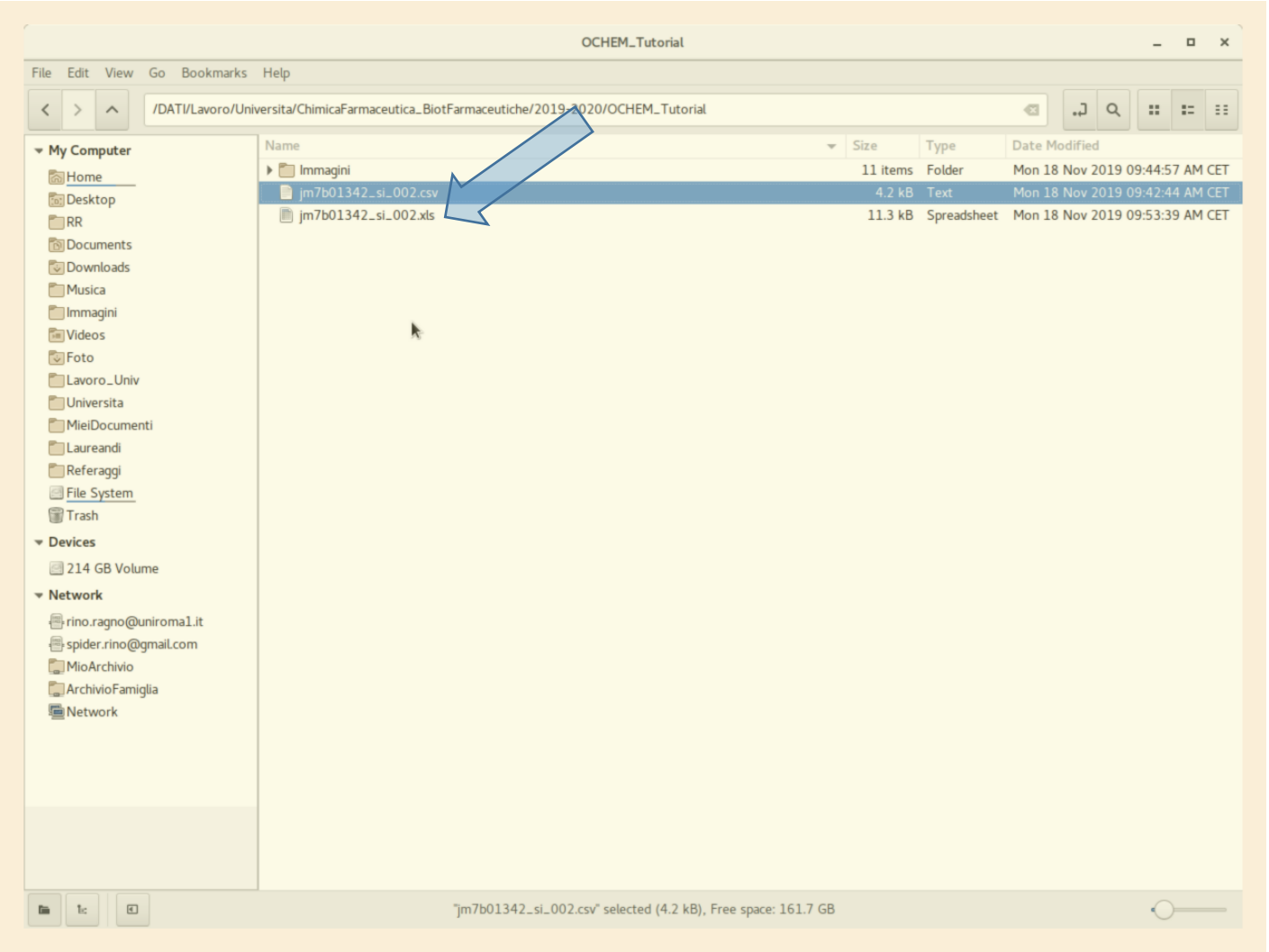

| | C | D | E | F | G | H |
|---|---|---|---|---|---|---|
| 1 | MW | PAK1 Inhibitory Activity (Ki, ?M) | Ki | Tmax-T0 (PAK4, ?C) | clogP | PSA |
| 2 | 466 | 0.003 | 0.009 | ND | 3.7 | 69.8 |
| 3 | 490 | 0.052 | 0.026 | ND | 4.3 | 98.3 |
| 4 | 357.5 | >10 | 0.71 | 0.5?0.0 | 2.6 | 71.3 |
| 5 | 383.5 | ND | 0.652 | 0.5?0.0 | 3 | 70.9 |
| 6 | 385.5 | ND | 0.884 | 0.8?0.3 | 3.2 | 70.4 |
| 7 | 419.5 | ND | 1.44 | 0.9?0.3 | 3.8 | 71 |
| 8 | 357.5 | 1.891 | 0.099 | 2.1?0.0 | 2.3 | 86.4 |
| 9 | 383.5 | 0.288 | 0.016 | 4.7?0.3 | 2.9 | 85 |
| 10 | 385.5 | ND | 0.077 | 2.3?0.3 | 3.1 | 84.8 |
| 11 | 419.5 | ND | 0.256 | 0.7?0.3 | 3.5 | 85.1 |
| 12 | 385.5 | ND | 0.354 | 0.5?0.0 | 1.7 | 108.9 |
| 13 | 386.5 | ND | 3.817 | 0.1?0.3 | 2.1 | 105.3 |
| 14 | 399.5 | ND | 0.674 | 0.4?0.3 | 2.1 | 104.3 |
| 15 | 400.5 | ND | 3.89 | -0.1?0.0 | 2.9 | 101.5 |
| 16 | 399.5 | ND | 0.672 | 0.3?0.3 | 2.3 | 97 |
| 17 | 371.5 | >10 | 0.045 | 2.4?0.5 | 1.6 | 98.7 |
| 18 | 372.5 | ND | 2.172 | 0.2?0.3 | 2.2 | 92.5 |
| 19 | 385.5 | ND | 0.768 | ND | 2 | 88.9 |
| 20 | 385.5 | ND | 2.408 | ND | 0.8 | 132.2 |
| 21 | 399.5 | ND | 0.576 | 0.2?0.3 | 2.7 | 95.7 |
| 22 | 385.5 | ND | 0.325 | 0.2?0.3 | 2.4 | 108.8 |
| 23 | 399.5 | ND | 1.63 | 0.0?0.3 | 2.7 | 95.6 |
| 24 | 357.5 | ND | 0.782 | 0.7?0.3 | 2.8 | 111.4 |
| 25 | 397.5 | 2.75 | 0.016 | 4.1?0.3 | 2.2 | 97.4 |
| 26 | 385.5 | >4.52 | 0.218 | 1.5?0.3 | 2.4 | 91.4 |
| 27 | 411.5 | 0.383 | 0.026 | 3.6?0.3 | 3 | 90.1 |
| 28 | 385.5 | >4.52 | 0.151 | 1.8?0.3 | 1.9 | 92.5 |
| 29 | 411.5 | 2.08 | 0.017 | 3.6?0.3 | 2.5 | 91.1 |

# Preparing the dataset

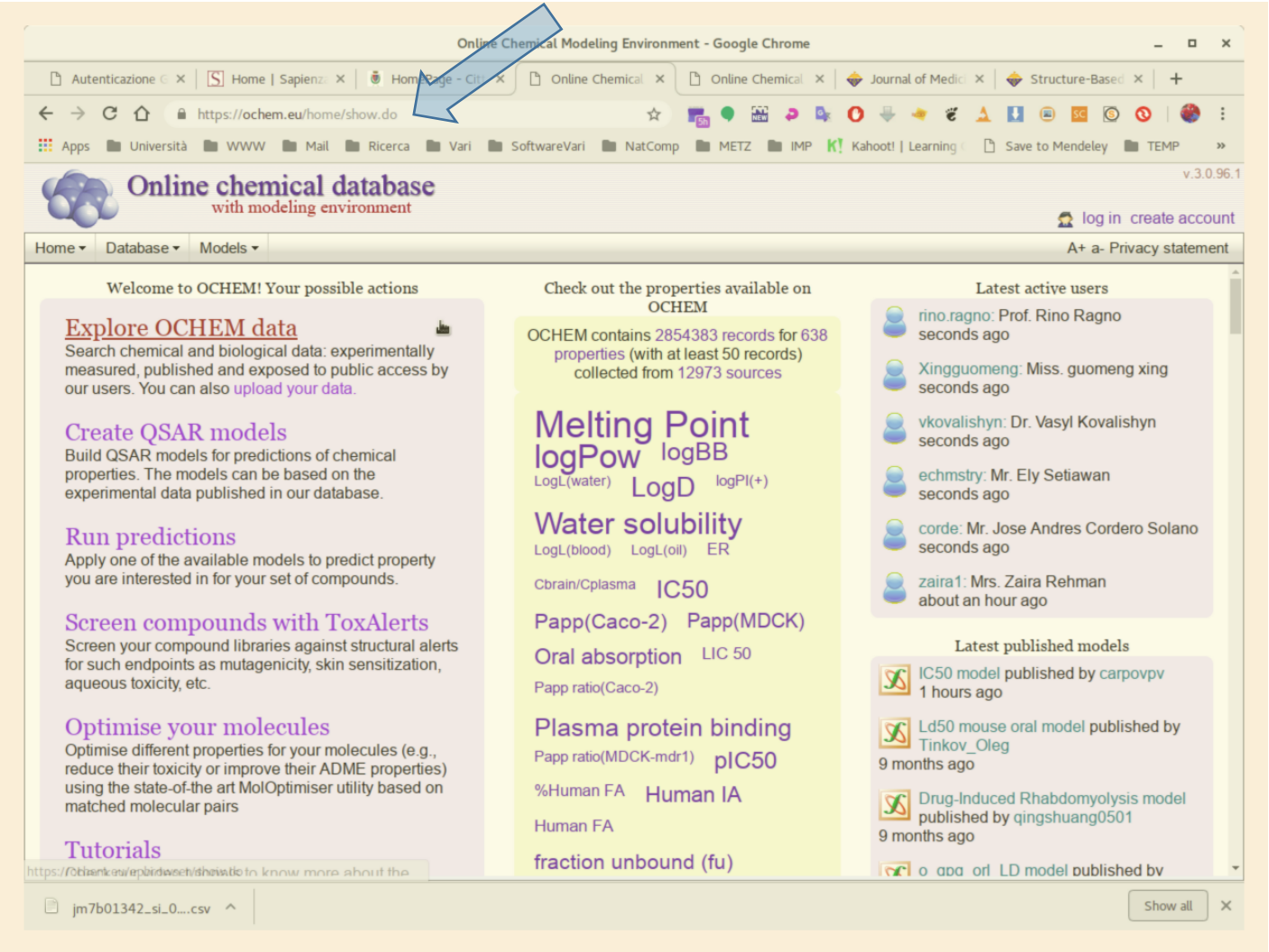



| | A | B | C | D |
|---|---|---|---|---|
| 1 | RECORDID | SMILES | Ki | |
| 2 | staurosporine | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@ | 0.009 | |
| 3 | PF3758309 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N | 0.026 | |
| 4 | 10a | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42 | 0.71 | |
| 5 | 10b | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C | 0.652 | |
| 6 | 10c | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1 | 0.884 | |
| 7 | 10d | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N | 1.44 | |
| 8 | 11a | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42)=NN1 | 0.099 | |
| 9 | 11b | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C | 0.016 | |
| 10 | 11c | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1 | 0.077 | |
| 11 | 11d | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C | 0.256 | |
| 12 | 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.354 | |
| 13 | 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 3.817 | |
| 14 | 14 | CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 0.674 | |
| 15 | 15 | CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 3.89 | |
| 16 | 16 | CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.672 | |
| 17 | 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.045 | |
| 18 | 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 2.172 | |
| 19 | 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.768 | |
| 20 | 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2= | 2.408 | |
| 21 | 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.576 | |
| 22 | 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.325 | |
| 23 | 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 1.63 | |
| 24 | 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.782 | |
| 25 | 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5) | 0.016 | |
| 26 | 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(Cl)C=C42 | 0.218 | |
| 27 | 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5O | 0.026 | |
| 28 | 28 | CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(Cl)C=C42) | 0.151 | |
| 29 | | O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5CC | 0.017 | |

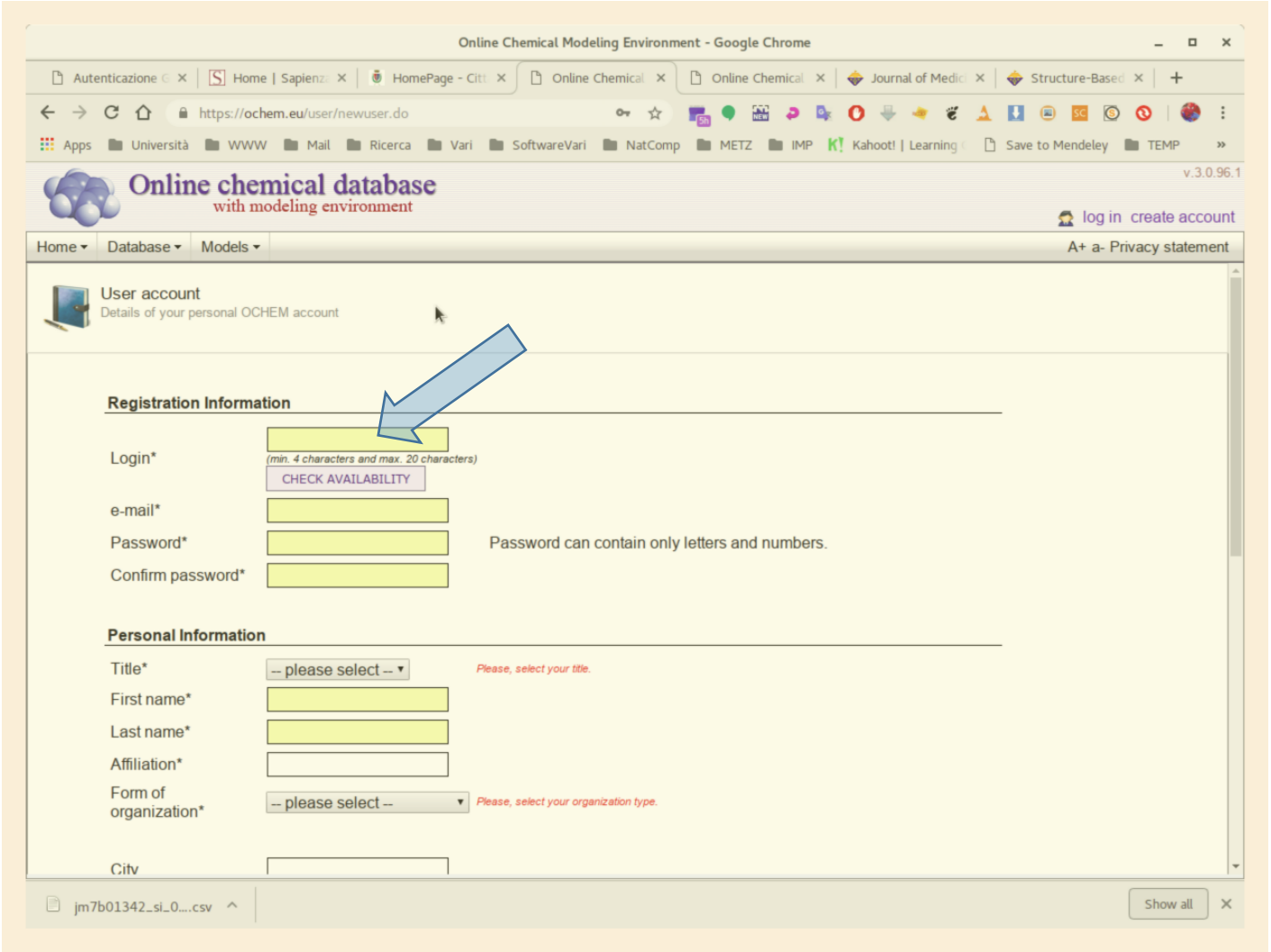
# Preparing the dataset



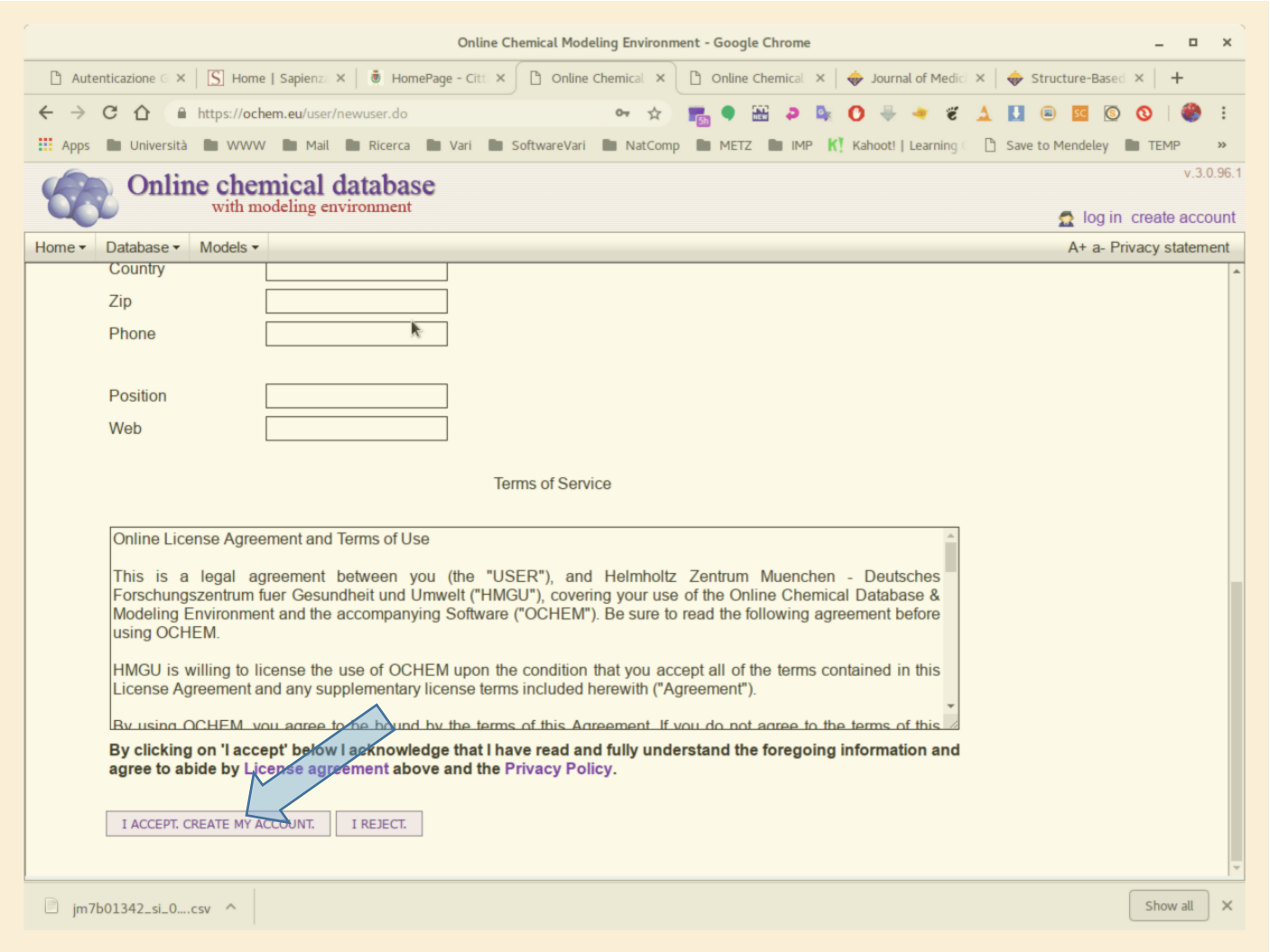
# Preparing the dataset



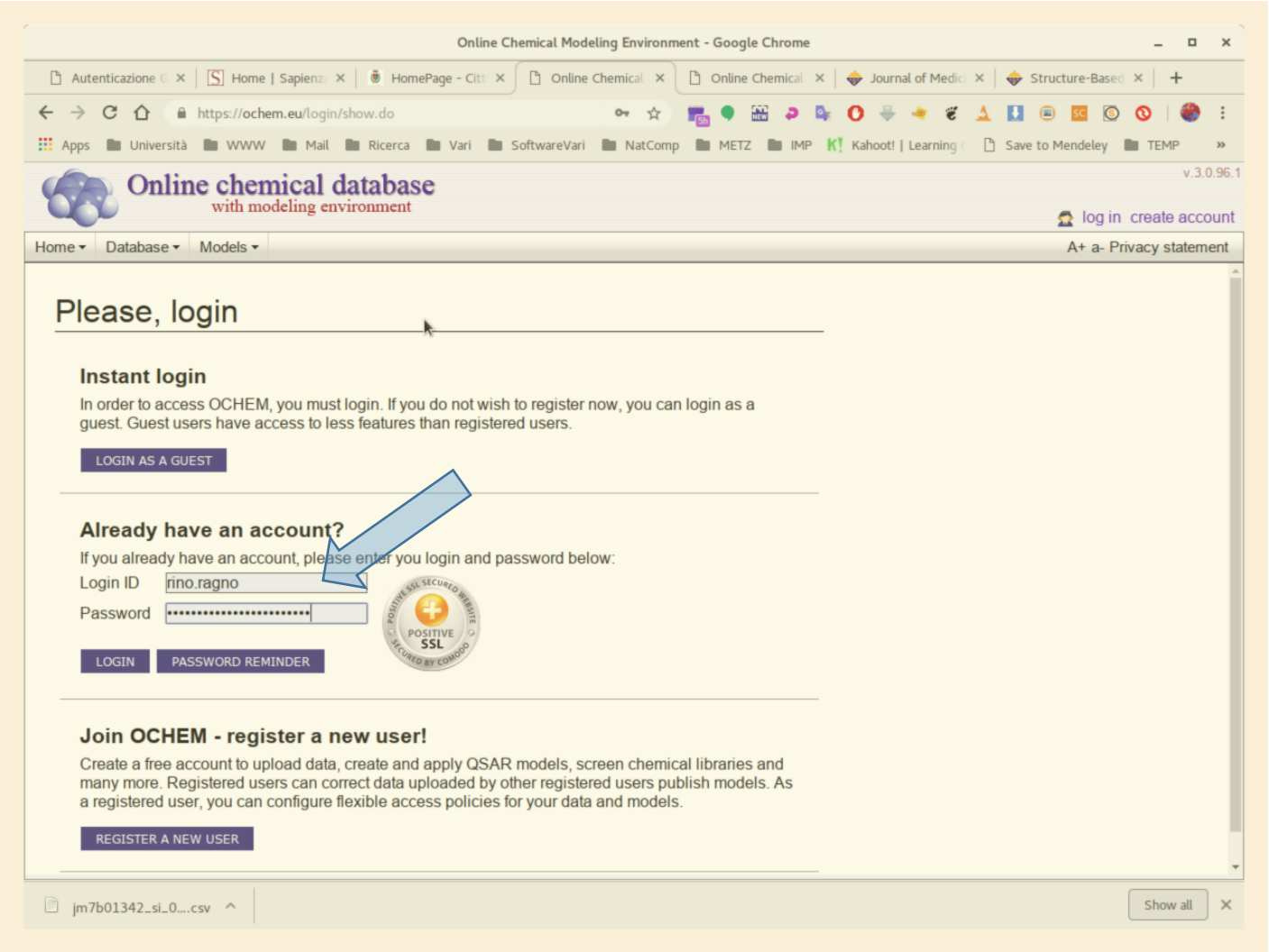
# Preparing the dataset



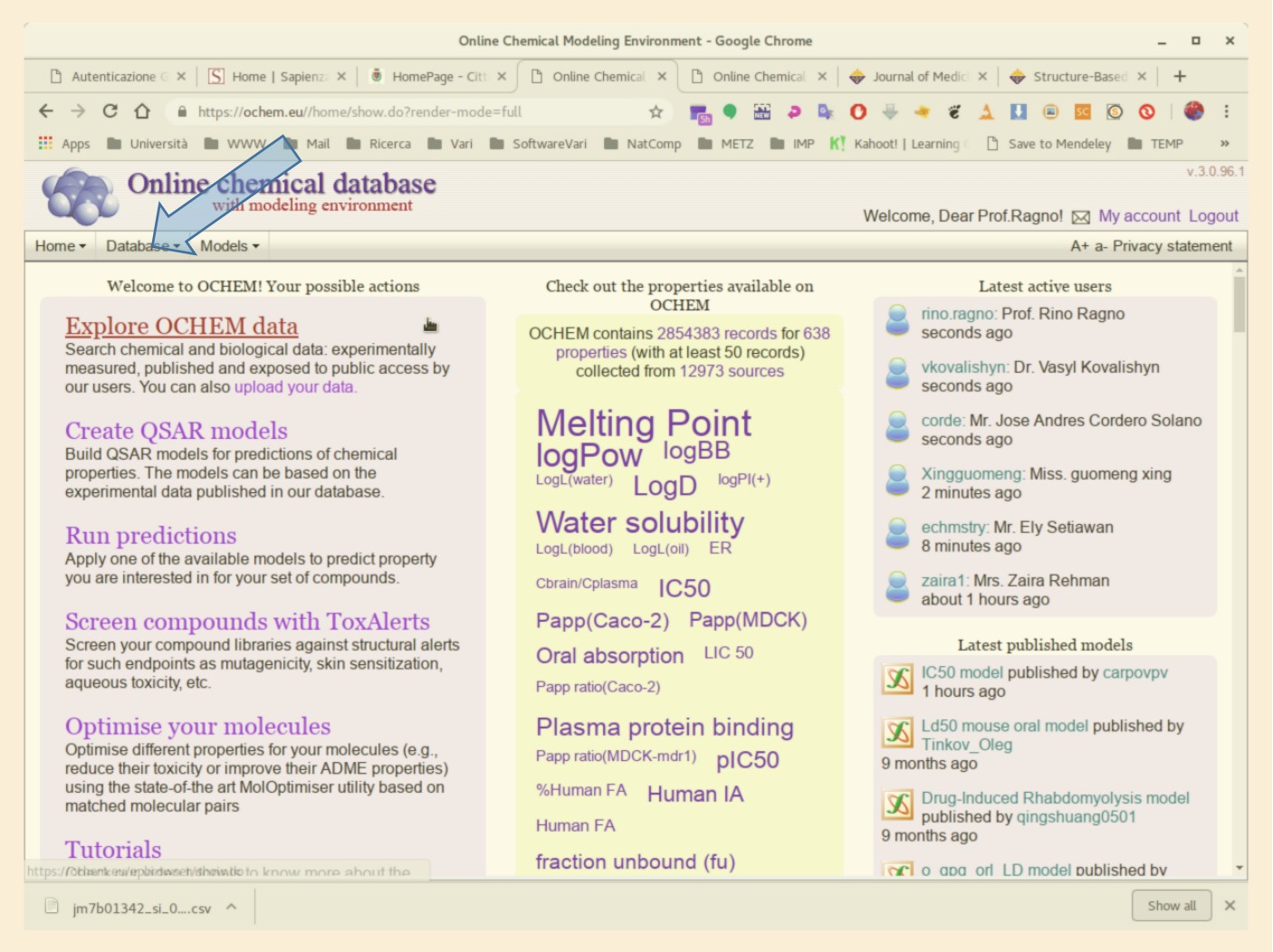
# Preparing the dataset



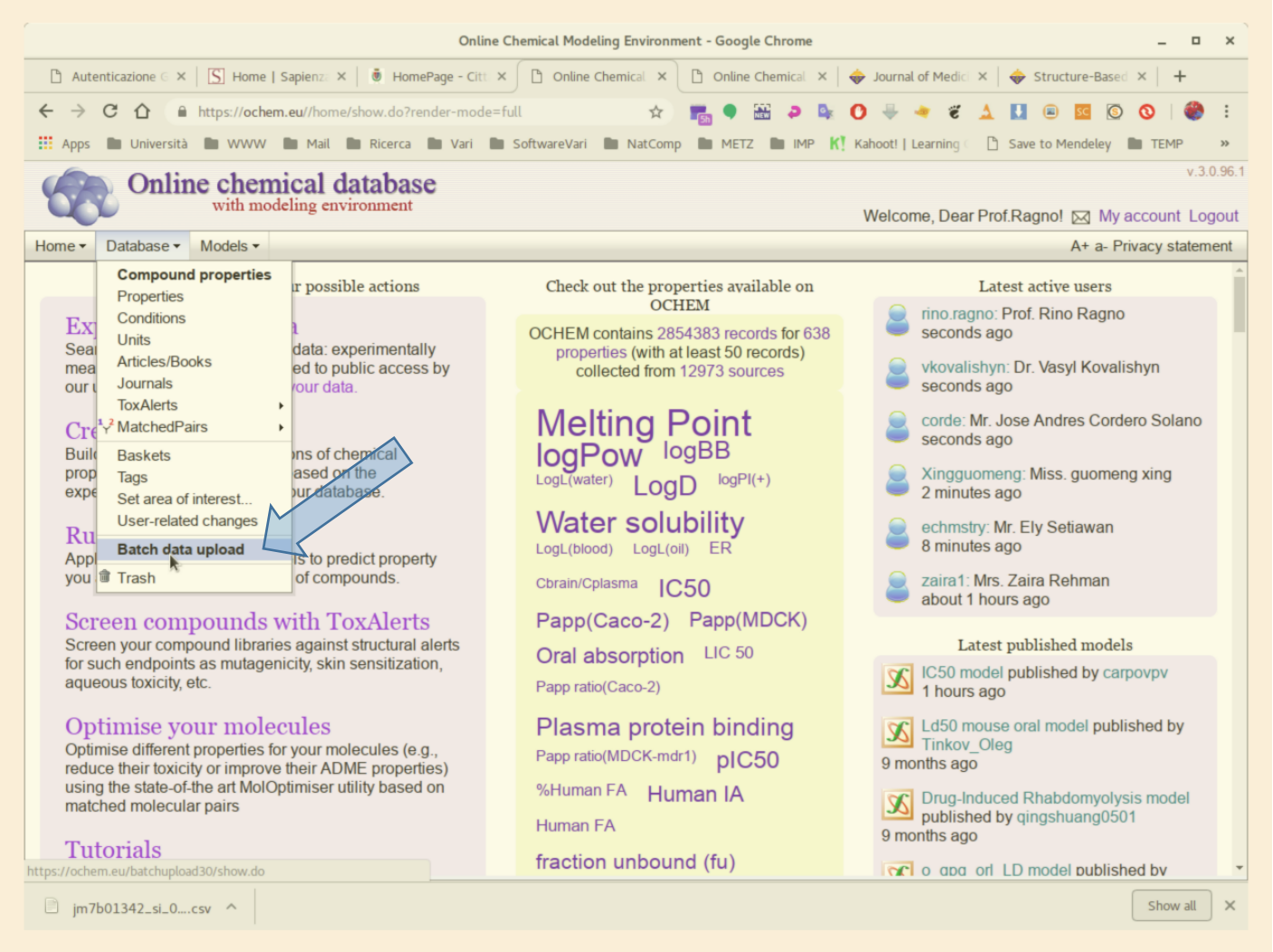
# Preparing the dataset



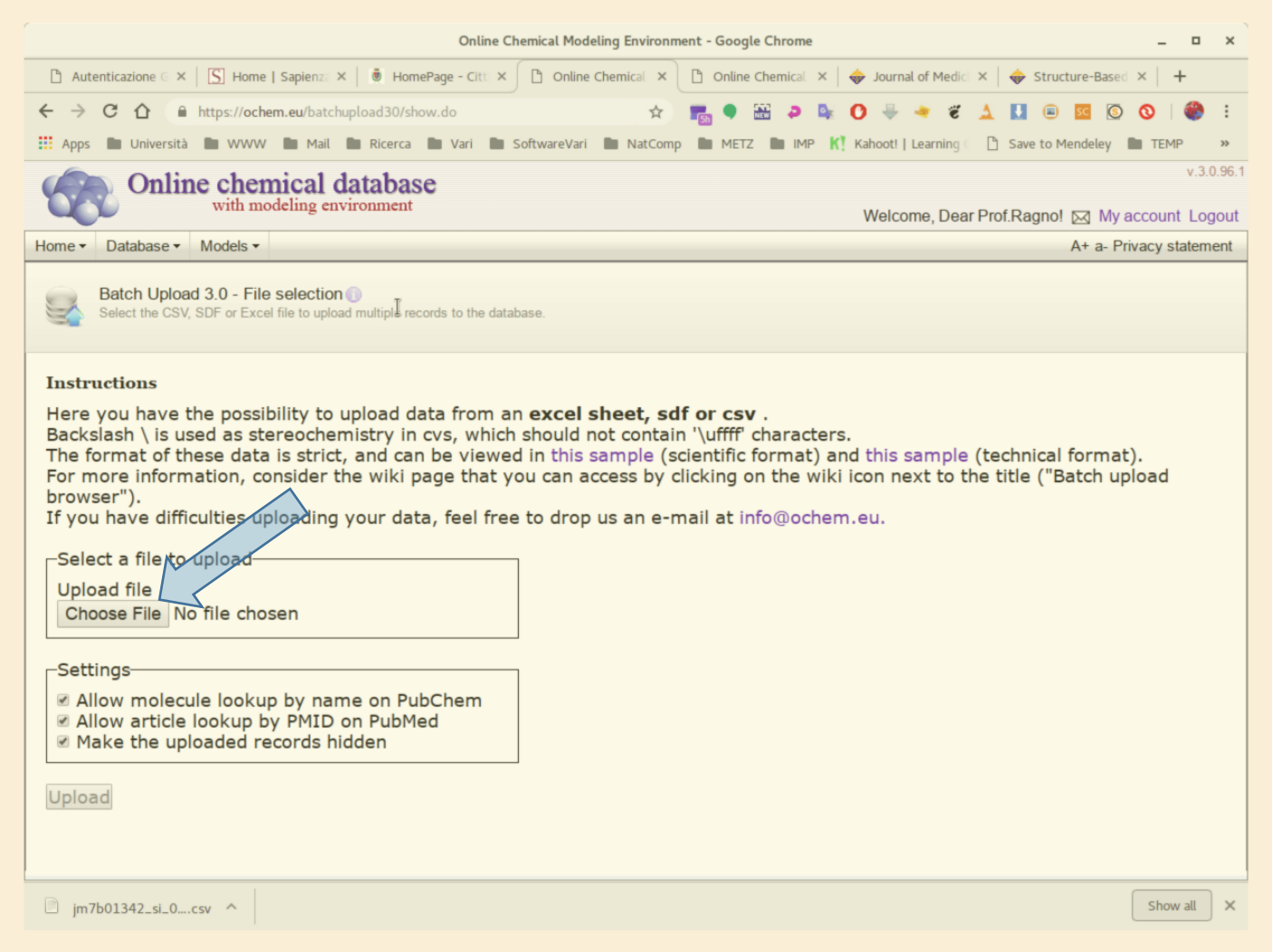
# Preparing the dataset



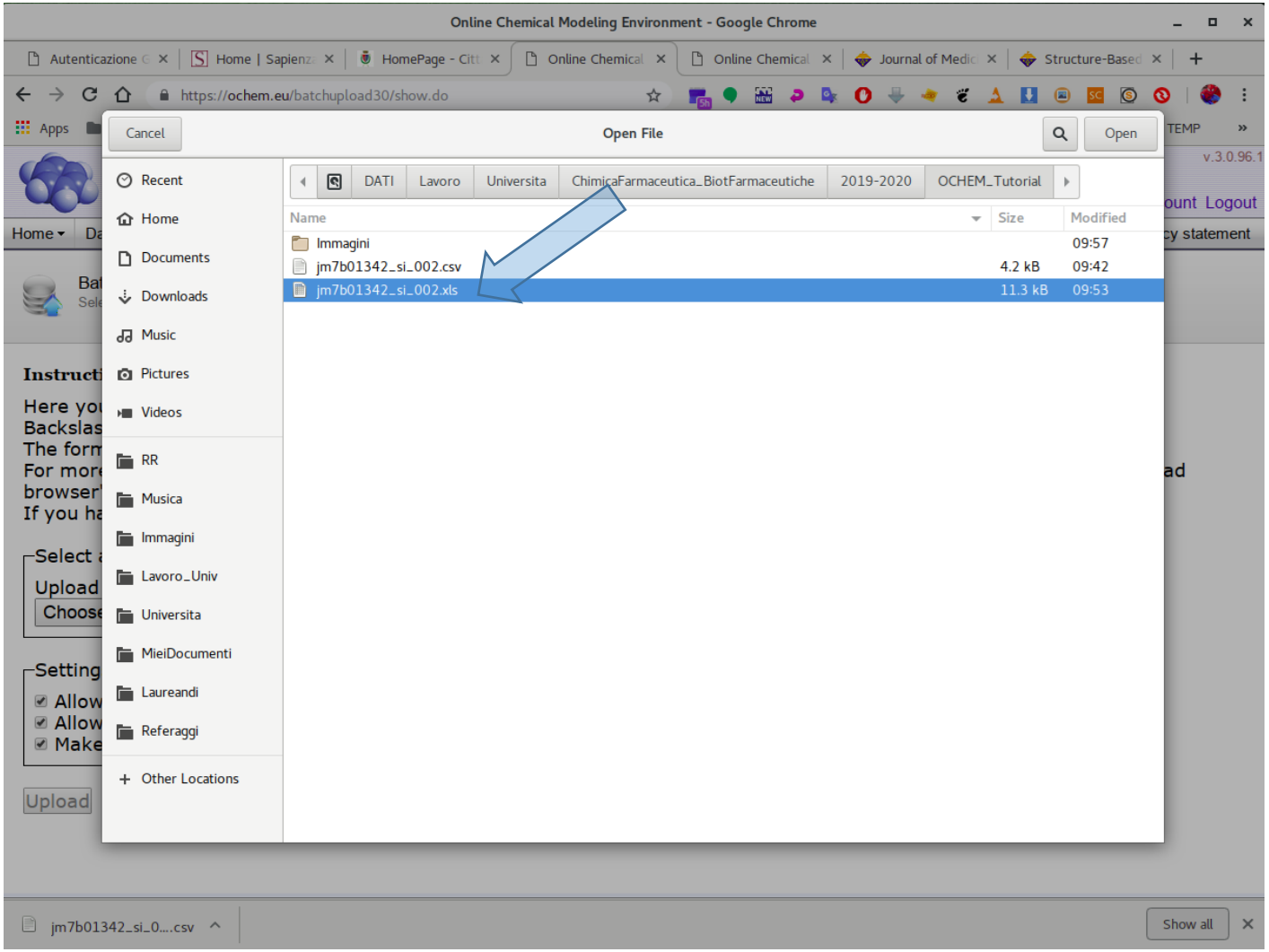
# Preparing the dataset



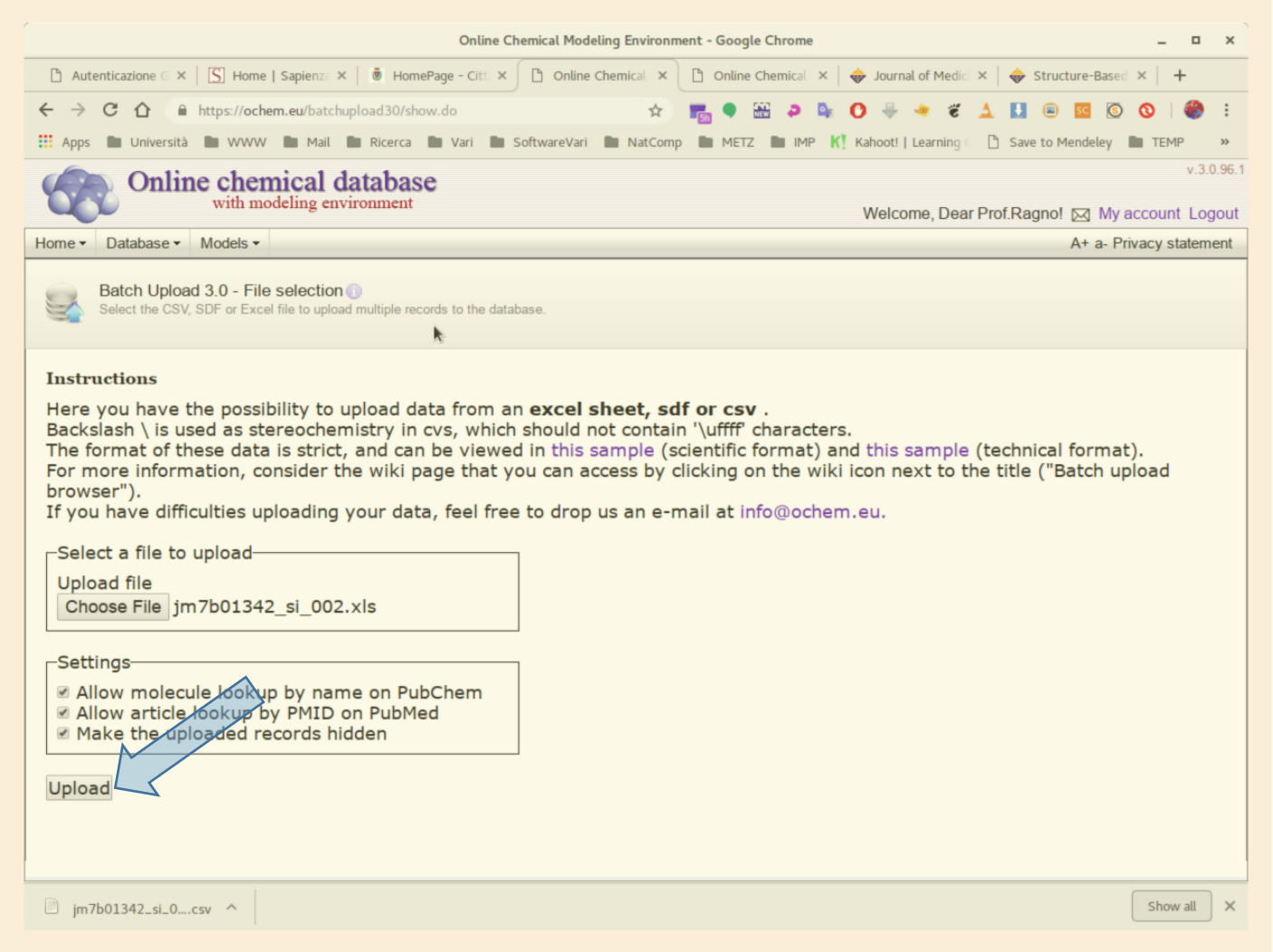
# Preparing the dataset



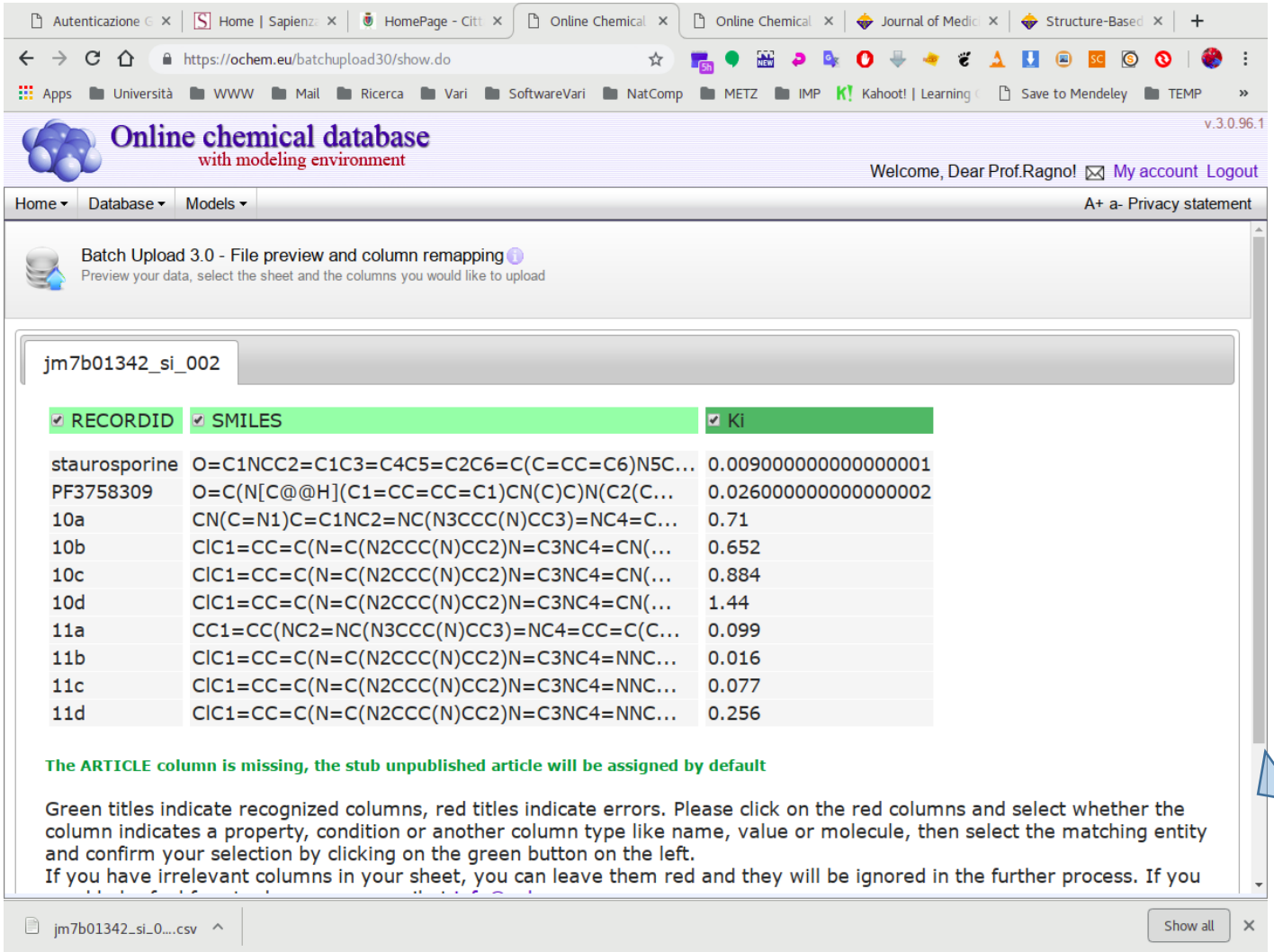
# Preparing the dataset



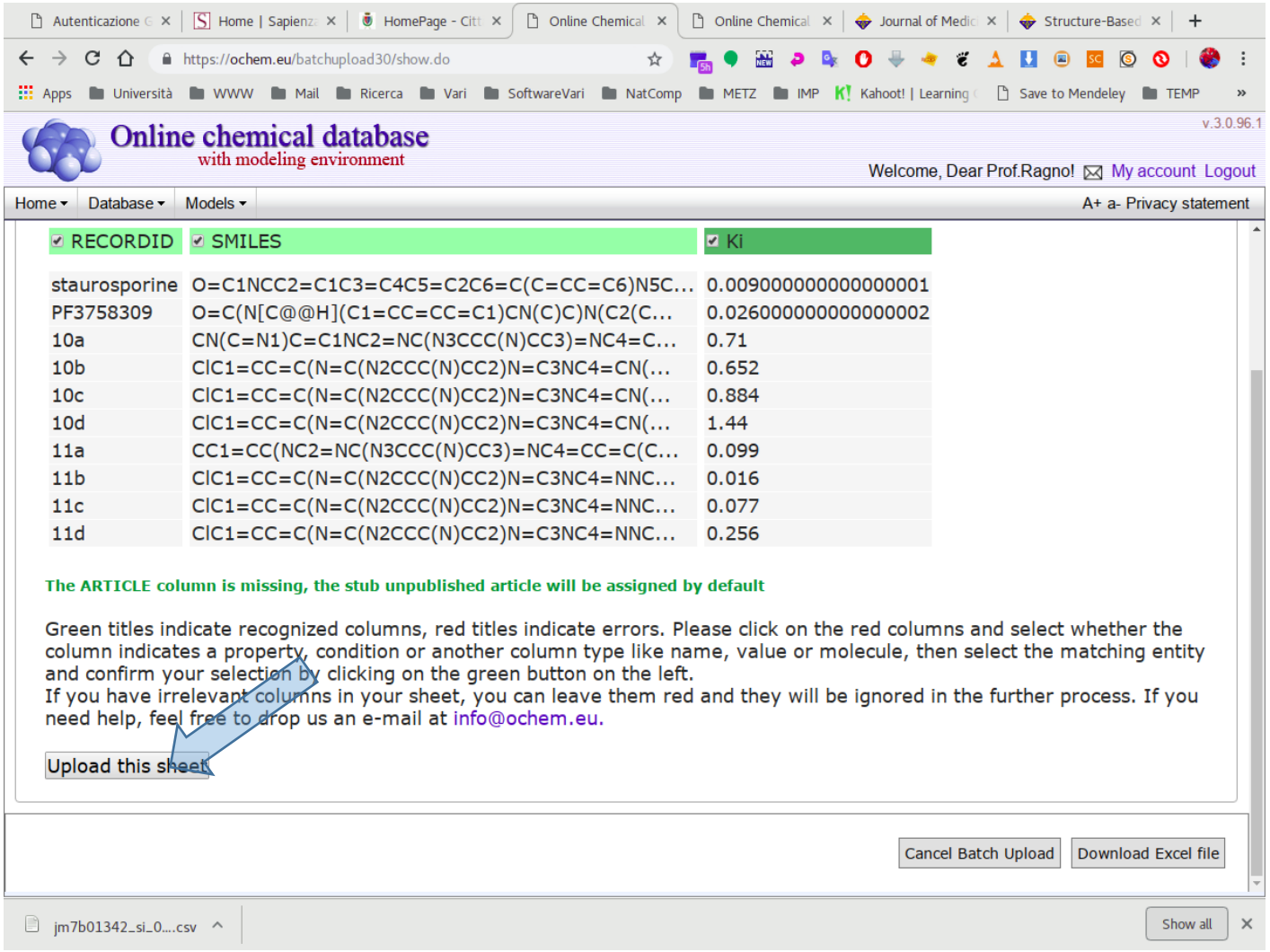
# Preparing the dataset



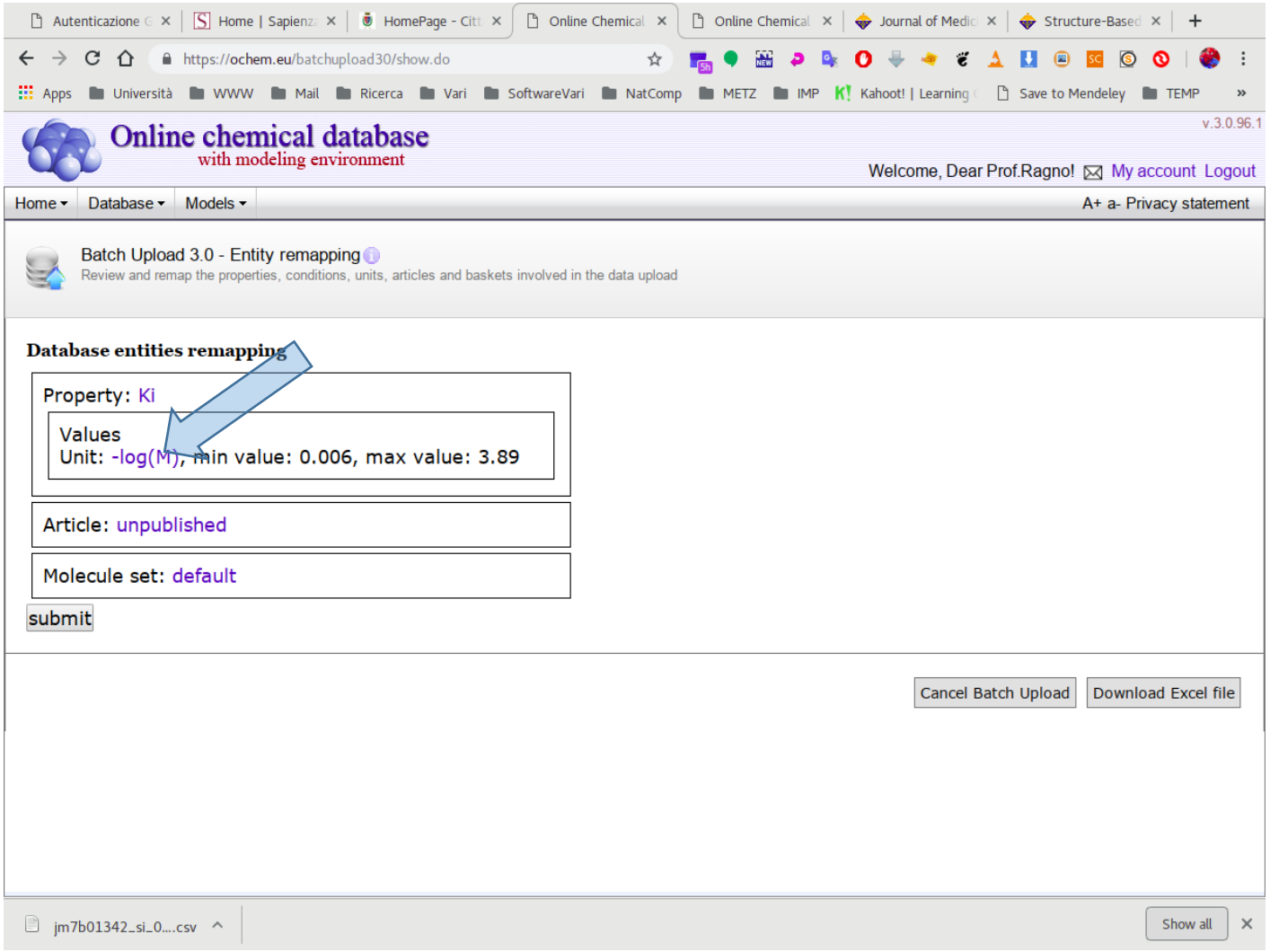
# Preparing the dataset



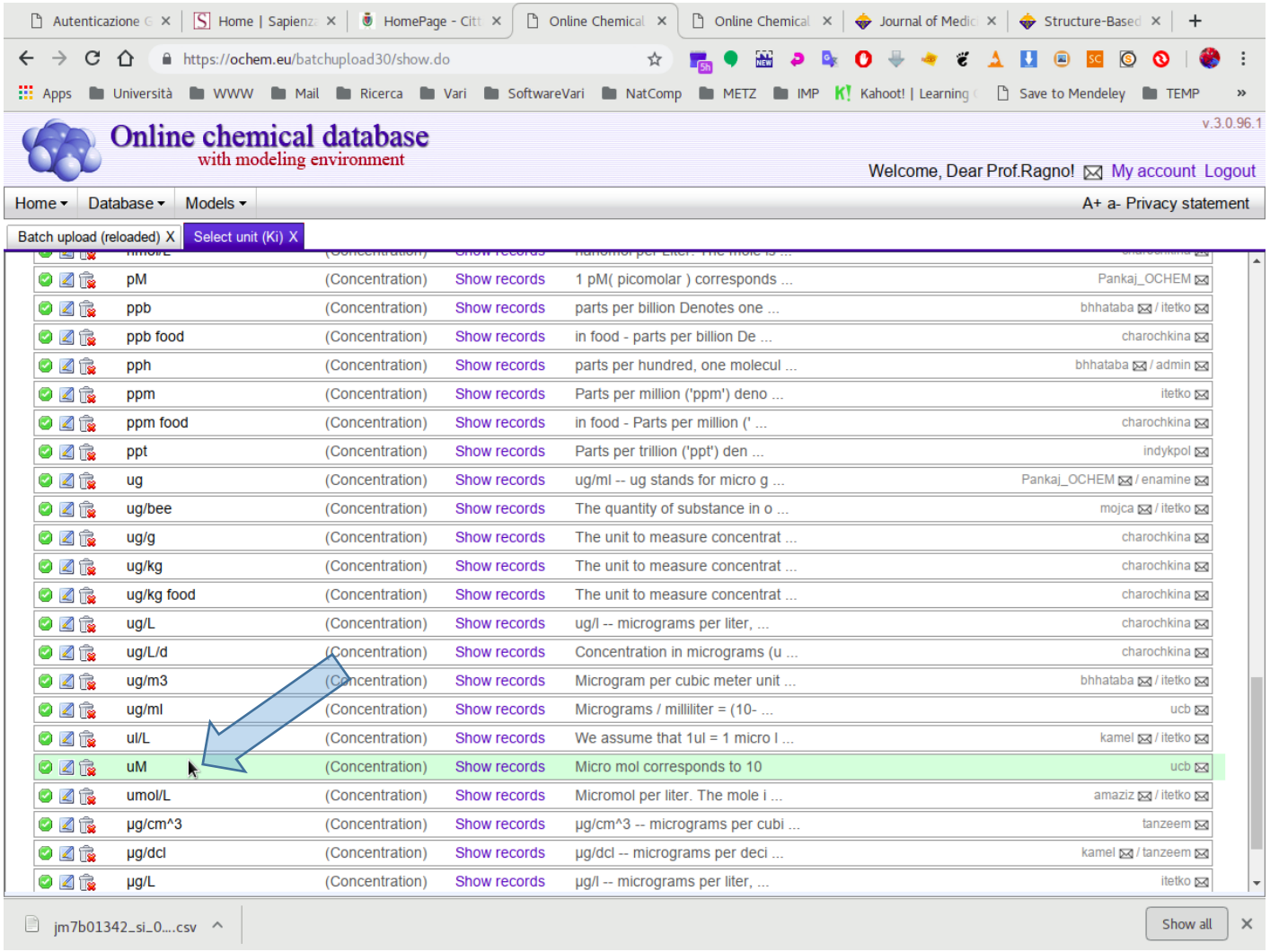
# Preparing the dataset



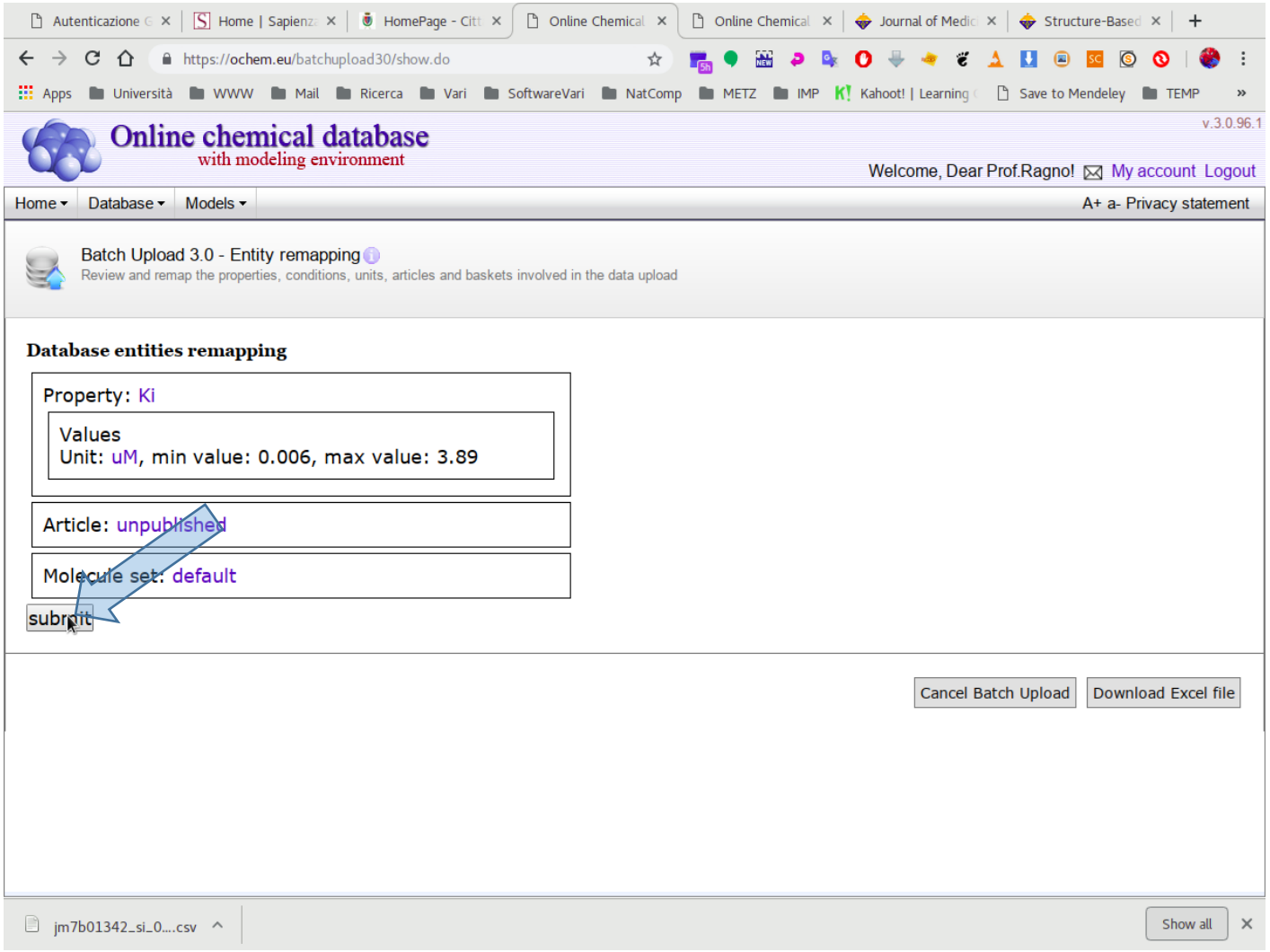
# Preparing the dataset



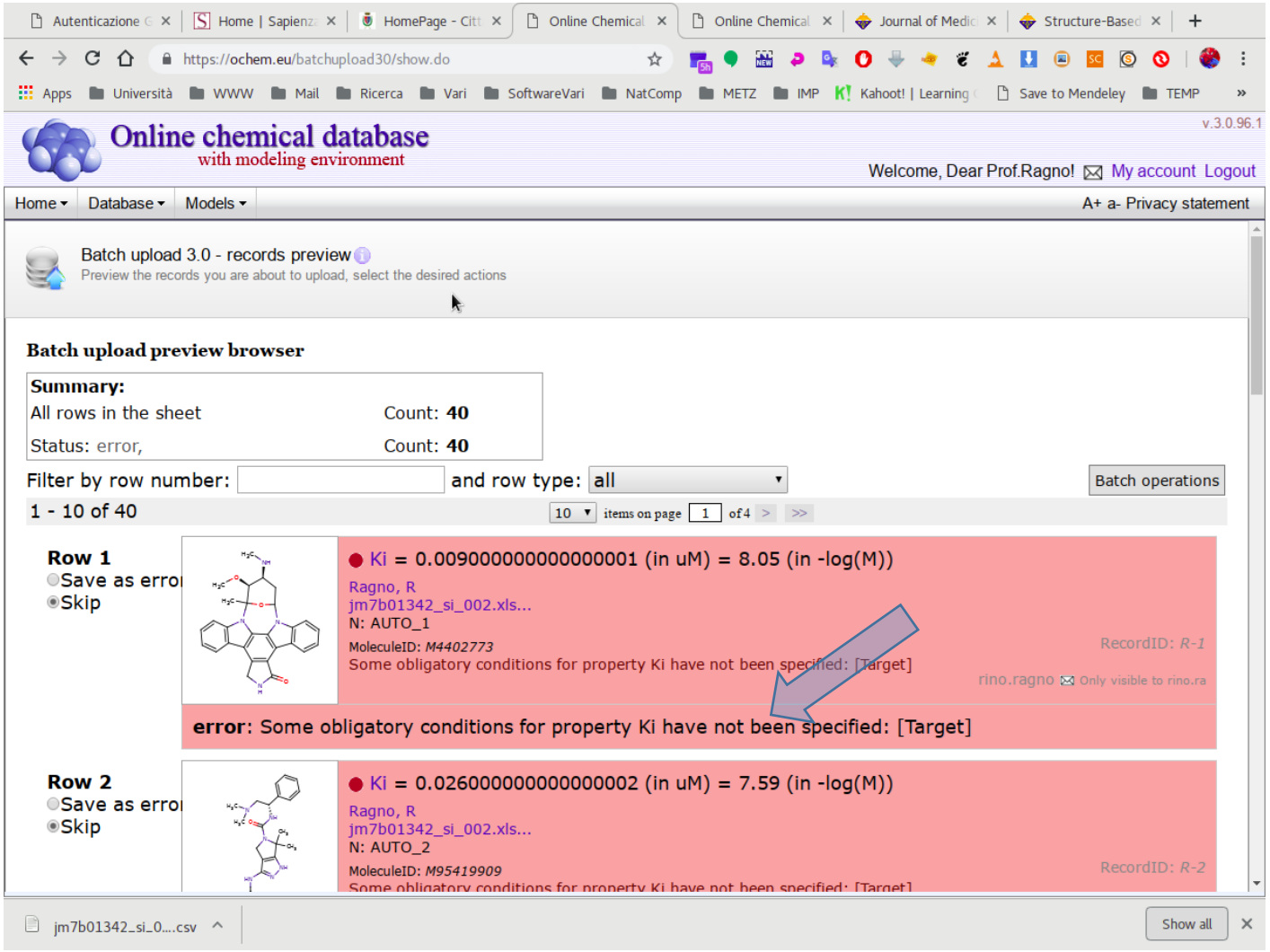
# Preparing the dataset



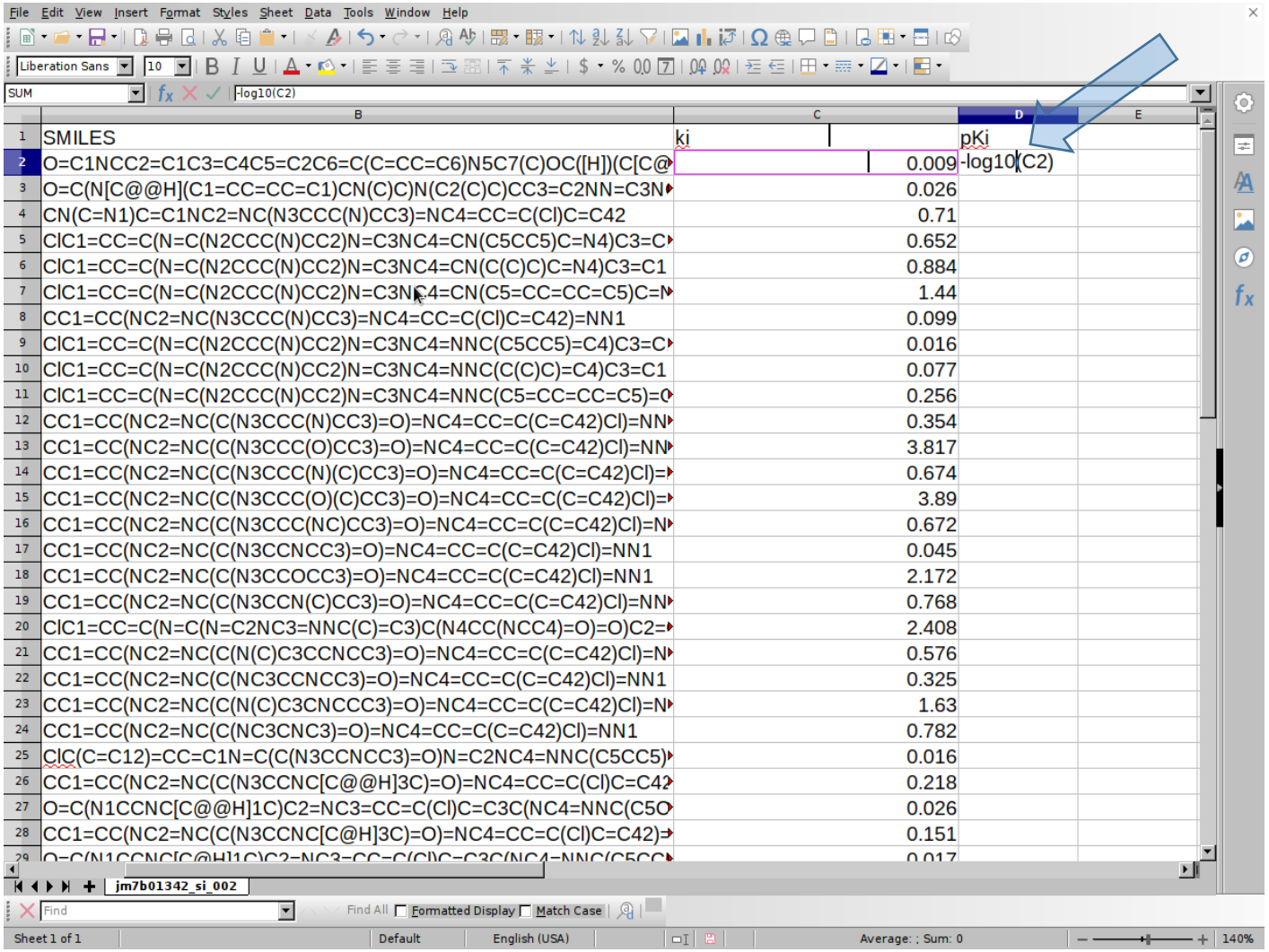
# Preparing the dataset



# Preparing the dataset

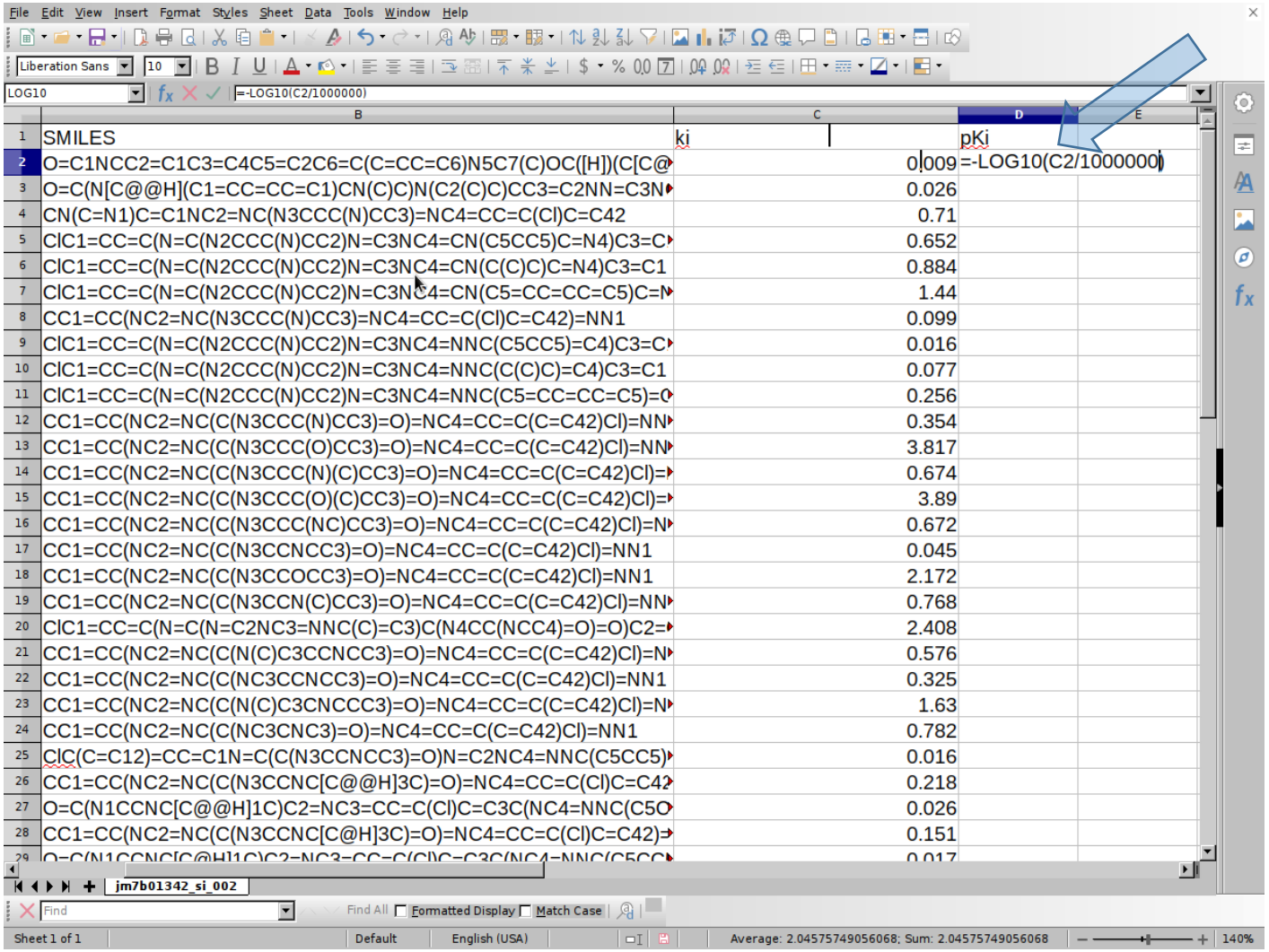


# Preparing the dataset



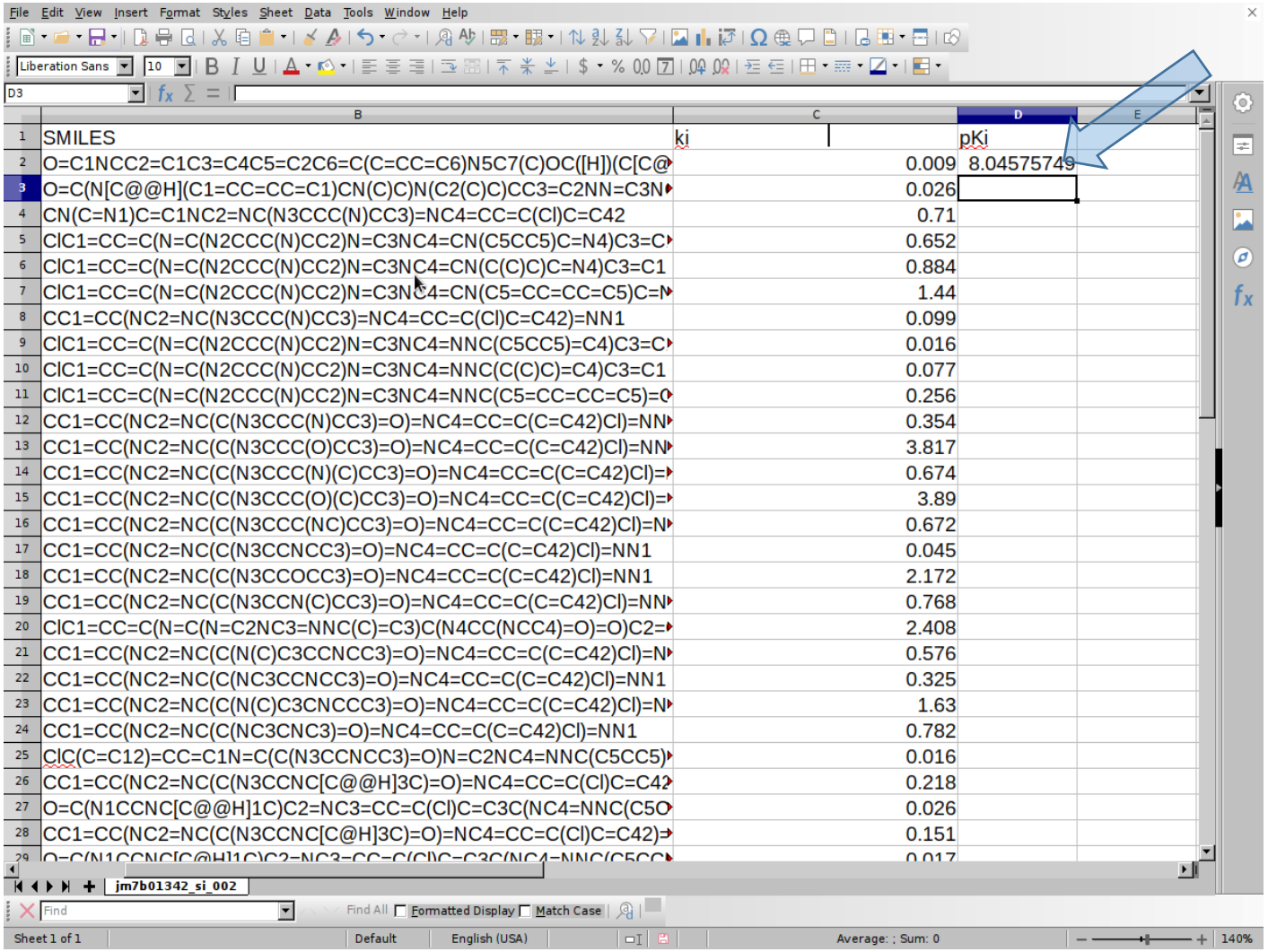

| | B | C | D | E |
|---|---|---|---|---|
| 1 | SMILES | ki | pKi | |
| 2 | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@ | 0.009 | -log10(C2) | |
| 3 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N | 0.026 | | |
| 4 | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42 | 0.71 | | |
| 5 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C | 0.652 | | |
| 6 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1 | 0.884 | | |
| 7 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N | 1.44 | | |
| 8 | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42)=NN1 | 0.099 | | |
| 9 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C | 0.016 | | |
| 10 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1 | 0.077 | | |
| 11 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C | 0.256 | | |
| 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.354 | | |
| 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 3.817 | | |
| 14 | CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 0.674 | | |
| 15 | CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 3.89 | | |
| 16 | CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.672 | | |
| 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.045 | | |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 2.172 | | |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.768 | | |
| 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2= | 2.408 | | |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.576 | | |
| 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.325 | | |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 1.63 | | |
| 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.782 | | |
| 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5) | 0.016 | | |
| 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(Cl)C=C42 | 0.218 | | |
| 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5O | 0.026 | | |
| 28 | CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(Cl)C=C42) | 0.151 | | |
| 29 | O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5CC | 0.017 | | |

# Preparing the dataset



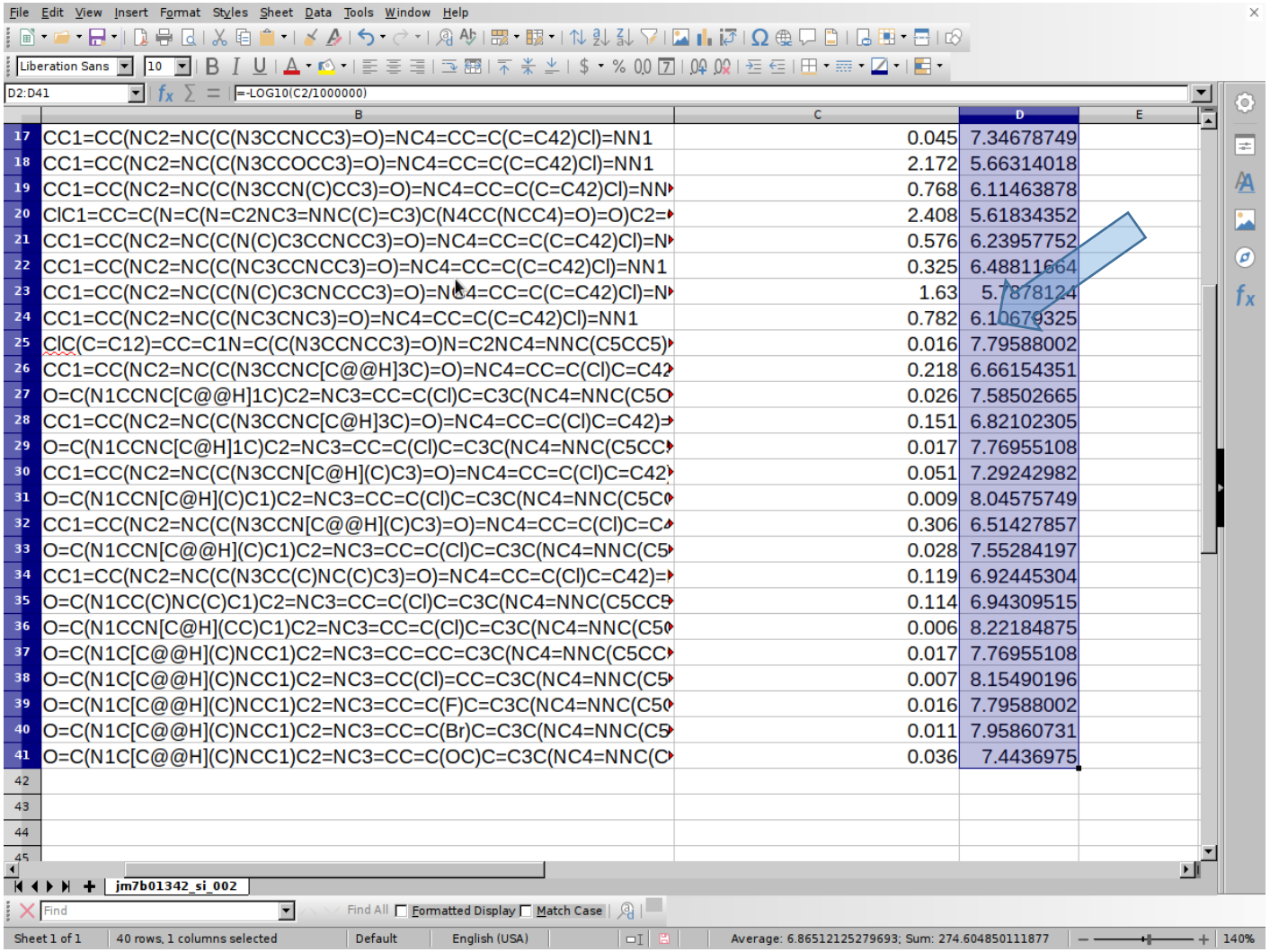

| | B | C | D | E |
|---|---|---|---|---|
| 1 | SMILES | ki | pKi | |
| 2 | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@ | 0.009 | =-LOG10(C2/1000000) | |
| 3 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N | 0.026 | | |
| 4 | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42 | 0.71 | | |
| 5 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C | 0.652 | | |
| 6 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1 | 0.884 | | |
| 7 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N | 1.44 | | |
| 8 | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42)=NN1 | 0.099 | | |
| 9 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C | 0.016 | | |
| 10 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1 | 0.077 | | |
| 11 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C | 0.256 | | |
| 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.354 | | |
| 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 3.817 | | |
| 14 | CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 0.674 | | |
| 15 | CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 3.89 | | |
| 16 | CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.672 | | |
| 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.045 | | |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 2.172 | | |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.768 | | |
| 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2= | 2.408 | | |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.576 | | |
| 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.325 | | |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 1.63 | | |
| 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.782 | | |
| 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5) | 0.016 | | |
| 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(Cl)C=C42 | 0.218 | | |
| 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5O | 0.026 | | |
| 28 | CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(Cl)C=C42) | 0.151 | | |
| 29 | O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5CC | 0.017 | | |

# Preparing the dataset



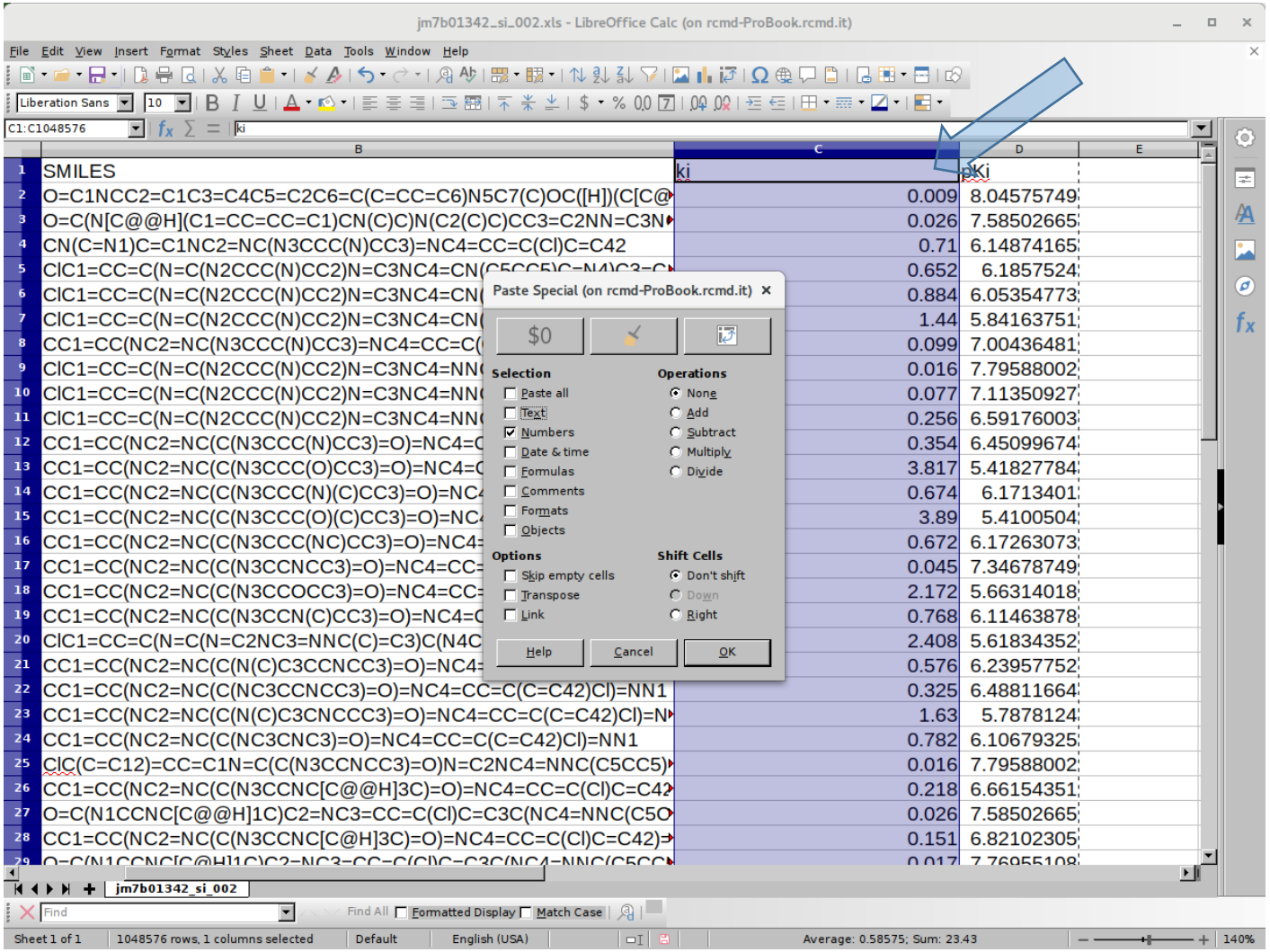
| | B | C | D | E |
|---|---|---|---|---|
| 1 | SMILES | ki | pKi | |
| 2 | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@ | 0.009 | 8.04575749 | |
| 3 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N | 0.026 | | |
| 4 | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42 | 0.71 | | |
| 5 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C | 0.652 | | |
| 6 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1 | 0.884 | | |
| 7 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N | 1.44 | | |
| 8 | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42)=NN1 | 0.099 | | |
| 9 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C | 0.016 | | |
| 10 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1 | 0.077 | | |
| 11 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=O | 0.256 | | |
| 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.354 | | |
| 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 3.817 | | |
| 14 | CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 0.674 | | |
| 15 | CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 3.89 | | |
| 16 | CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.672 | | |
| 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.045 | | |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 2.172 | | |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.768 | | |
| 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2= | 2.408 | | |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.576 | | |
| 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.325 | | |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 1.63 | | |
| 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.782 | | |
| 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)= | 0.016 | | |
| 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(Cl)C=C42 | 0.218 | | |
| 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5C | 0.026 | | |
| 28 | CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(Cl)C=C42)= | 0.151 | | |
| 29 | O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5CC | 0.017 | | |

# Preparing the dataset

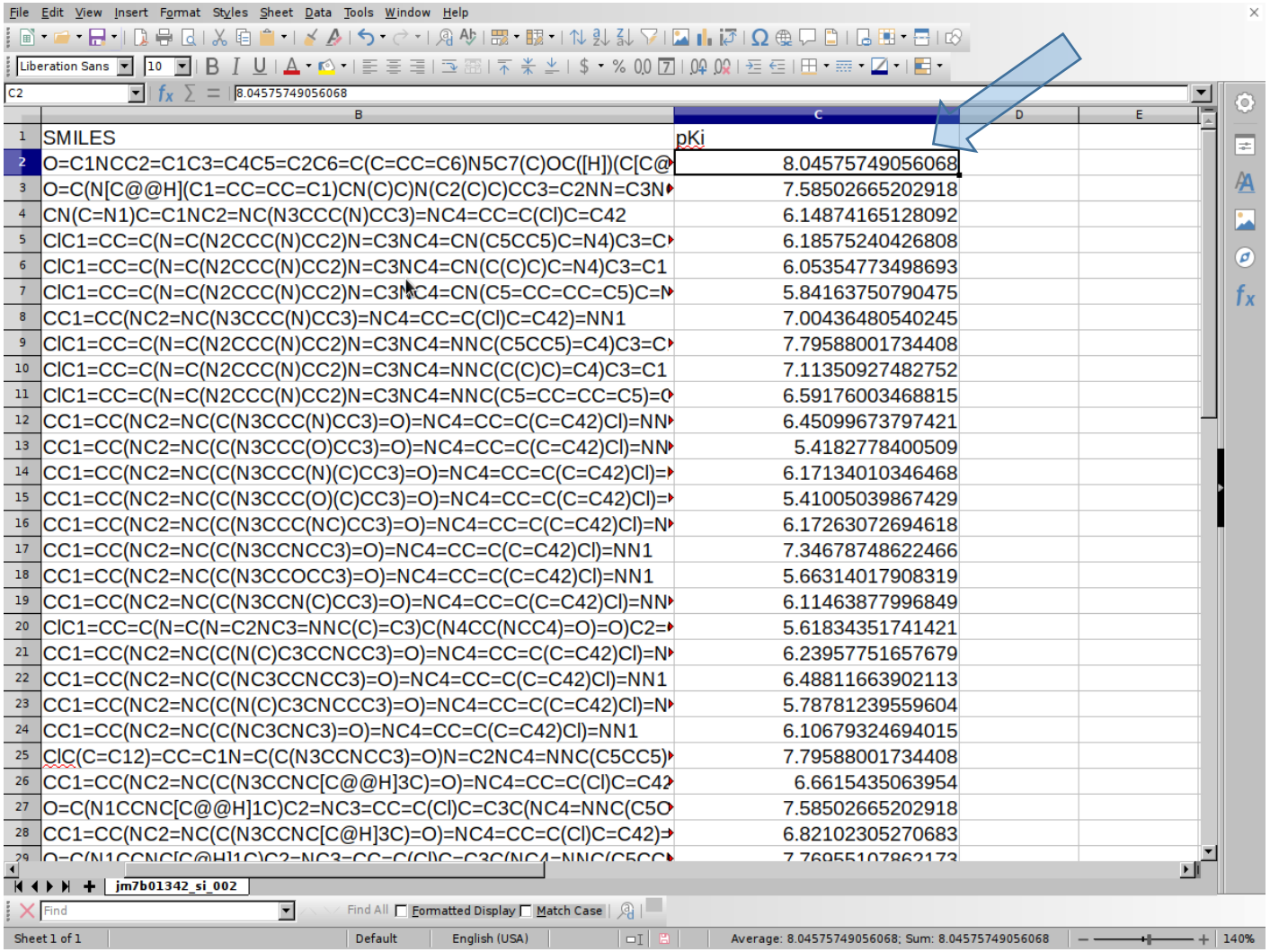



| | B | C | D | E |
|---|---|---|---|---|
| 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.045 | 7.34678749 | |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 2.172 | 5.66314018 | |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 0.768 | 6.11463878 | |
| 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2= | 2.408 | 5.61834352 | |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 0.576 | 6.23957752 | |
| 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.325 | 6.48811664 | |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 1.63 | 5.7878124 | |
| 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 0.782 | 6.10679325 | |
| 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5) | 0.016 | 7.79588002 | |
| 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(Cl)C=C42 | 0.218 | 6.66154351 | |
| 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5O | 0.026 | 7.58502665 | |
| 28 | CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(Cl)C=C42)= | 0.151 | 6.82102305 | |
| 29 | O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5CC | 0.017 | 7.76955108 | |
| 30 | CC1=CC(NC2=NC(C(N3CCN[C@H](C)C3)=O)=NC4=CC=C(Cl)C=C42 | 0.051 | 7.29242982 | |
| 31 | O=C(N1CCN[C@H](C)C1)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5C | 0.009 | 8.04575749 | |
| 32 | CC1=CC(NC2=NC(C(N3CCN[C@@H](C)C3)=O)=NC4=CC=C(Cl)C=C4 | 0.306 | 6.51427857 | |
| 33 | O=C(N1CCN[C@@H](C)C1)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5 | 0.028 | 7.55284197 | |
| 34 | CC1=CC(NC2=NC(C(N3CC(C)NC(C)C3)=O)=NC4=CC=C(Cl)C=C42)= | 0.119 | 6.92445304 | |
| 35 | O=C(N1CC(C)NC(C)C1)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5CC5 | 0.114 | 6.94309515 | |
| 36 | O=C(N1CCN[C@H](CC)C1)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5( | 0.006 | 8.22184875 | |
| 37 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=CC=C3C(NC4=NNC(C5CC | 0.017 | 7.76955108 | |
| 38 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC(Cl)=CC=C3C(NC4=NNC(C5 | 0.007 | 8.15490196 | |
| 39 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=C(F)C=C3C(NC4=NNC(C5( | 0.016 | 7.79588002 | |
| 40 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=C(Br)C=C3C(NC4=NNC(C5 | 0.011 | 7.95860731 | |
| 41 | O=C(N1C[C@@H](C)NCC1)C2=NC3=CC=C(OC)C=C3C(NC4=NNC(C | 0.036 | 7.4436975 | |
| 42 | | | | |
| 43 | | | | |
| 44 | | | | |
| 45 | | | | |

# Preparing the dataset

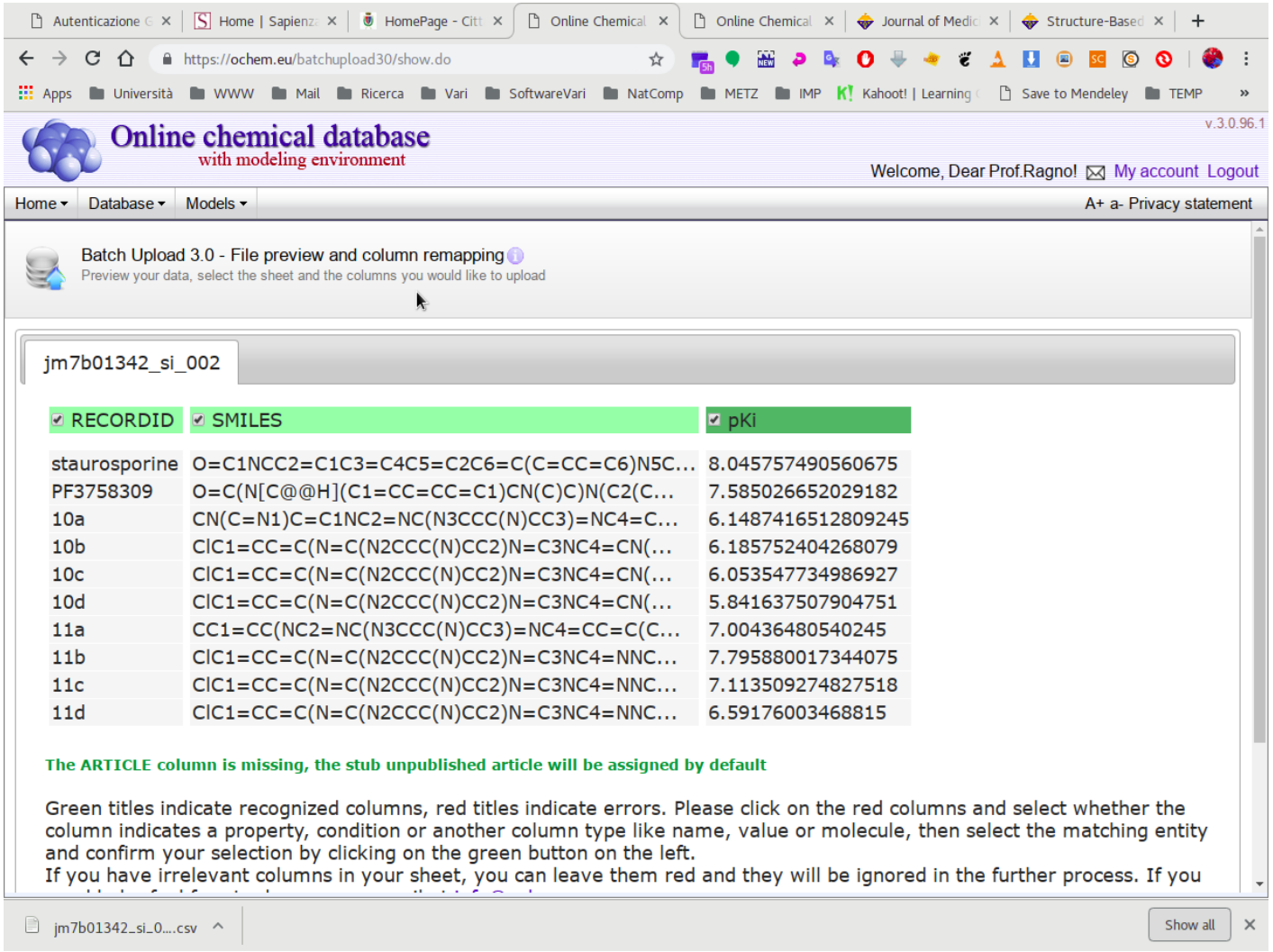


# Preparing the dataset

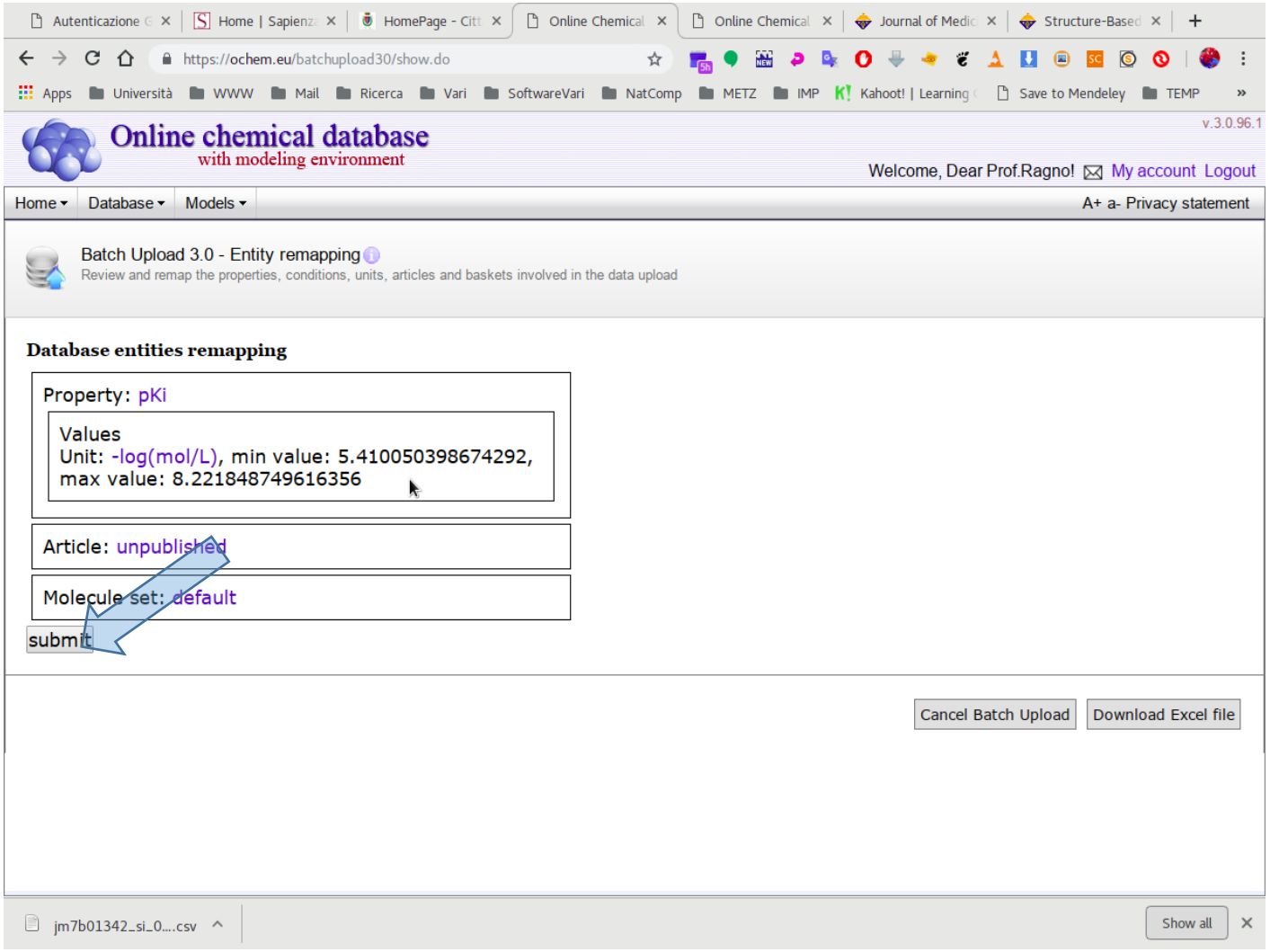



| | B | C | D | E |
|---|---|---|---|---|
| 1 | SMILES | pKi | | |
| 2 | O=C1NCC2=C1C3=C4C5=C2C6=C(C=CC=C6)N5C7(C)OC([H])(C[C@ | 8.04575749056068 | | |
| 3 | O=C(N[C@@H](C1=CC=CC=C1)CN(C)C)N(C2(C)C)CC3=C2NN=C3N | 7.58502665202918 | | |
| 4 | CN(C=N1)C=C1NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42 | 6.14874165128092 | | |
| 5 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5CC5)C=N4)C3=C | 6.18575240426808 | | |
| 6 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C(C)C)C=N4)C3=C1 | 6.05354773498693 | | |
| 7 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=CN(C5=CC=CC=C5)C=N | 5.84163750790475 | | |
| 8 | CC1=CC(NC2=NC(N3CCC(N)CC3)=NC4=CC=C(Cl)C=C42)=NN1 | 7.00436480540245 | | |
| 9 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5CC5)=C4)C3=C | 7.79588001734408 | | |
| 10 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C(C)C)=C4)C3=C1 | 7.11350927482752 | | |
| 11 | ClC1=CC=C(N=C(N2CCC(N)CC2)N=C3NC4=NNC(C5=CC=CC=C5)=C | 6.59176003468815 | | |
| 12 | CC1=CC(NC2=NC(C(N3CCC(N)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 6.45099673797421 | | |
| 13 | CC1=CC(NC2=NC(C(N3CCC(O)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 5.4182778400509 | | |
| 14 | CC1=CC(NC2=NC(C(N3CCC(N)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 6.17134010346468 | | |
| 15 | CC1=CC(NC2=NC(C(N3CCC(O)(C)CC3)=O)=NC4=CC=C(C=C42)Cl)= | 5.41005039867429 | | |
| 16 | CC1=CC(NC2=NC(C(N3CCC(NC)CC3)=O)=NC4=CC=C(C=C42)Cl)=N | 6.17263072694618 | | |
| 17 | CC1=CC(NC2=NC(C(N3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 7.34678748622466 | | |
| 18 | CC1=CC(NC2=NC(C(N3CCOCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 5.66314017908319 | | |
| 19 | CC1=CC(NC2=NC(C(N3CCN(C)CC3)=O)=NC4=CC=C(C=C42)Cl)=NN | 6.11463877996849 | | |
| 20 | ClC1=CC=C(N=C(N=C2NC3=NNC(C)=C3)C(N4CC(NCC4)=O)=O)C2= | 5.61834351741421 | | |
| 21 | CC1=CC(NC2=NC(C(N(C)C3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 6.23957751657679 | | |
| 22 | CC1=CC(NC2=NC(C(NC3CCNCC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 6.48811663902113 | | |
| 23 | CC1=CC(NC2=NC(C(N(C)C3CNCCC3)=O)=NC4=CC=C(C=C42)Cl)=N | 5.78781239559604 | | |
| 24 | CC1=CC(NC2=NC(C(NC3CNC3)=O)=NC4=CC=C(C=C42)Cl)=NN1 | 6.10679324694015 | | |
| 25 | ClC(C=C12)=CC=C1N=C(C(N3CCNCC3)=O)N=C2NC4=NNC(C5CC5)= | 7.79588001734408 | | |
| 26 | CC1=CC(NC2=NC(C(N3CCNC[C@@H]3C)=O)=NC4=CC=C(Cl)C=C42 | 6.6615435063954 | | |
| 27 | O=C(N1CCNC[C@@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5O | 7.58502665202918 | | |
| 28 | CC1=CC(NC2=NC(C(N3CCNC[C@H]3C)=O)=NC4=CC=C(Cl)C=C42)= | 6.82102305270683 | | |
| 29 | O=C(N1CCNC[C@H]1C)C2=NC3=CC=C(Cl)C=C3C(NC4=NNC(C5CC= | 7.76955107862173 | | |

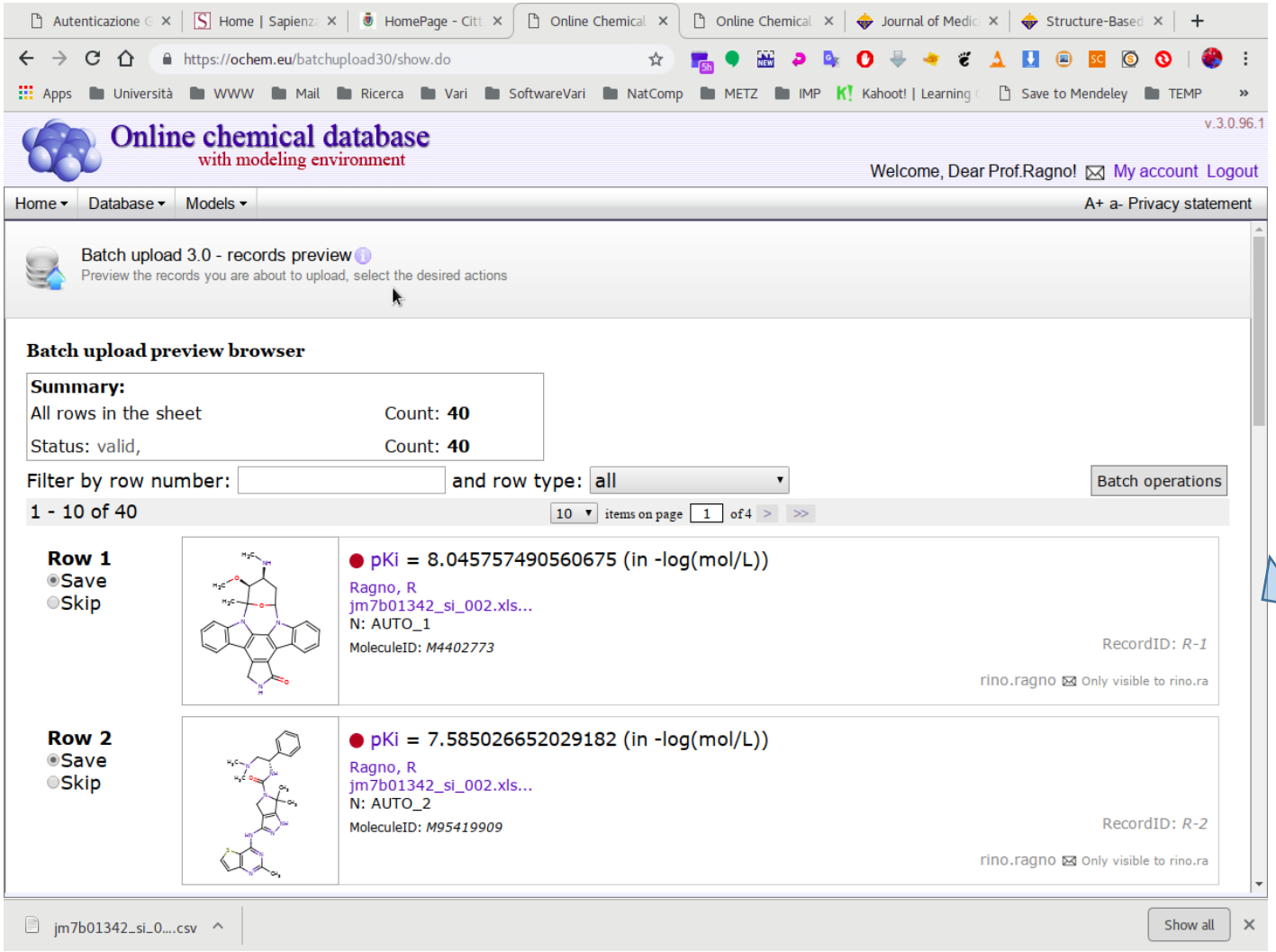
# Preparing the dataset



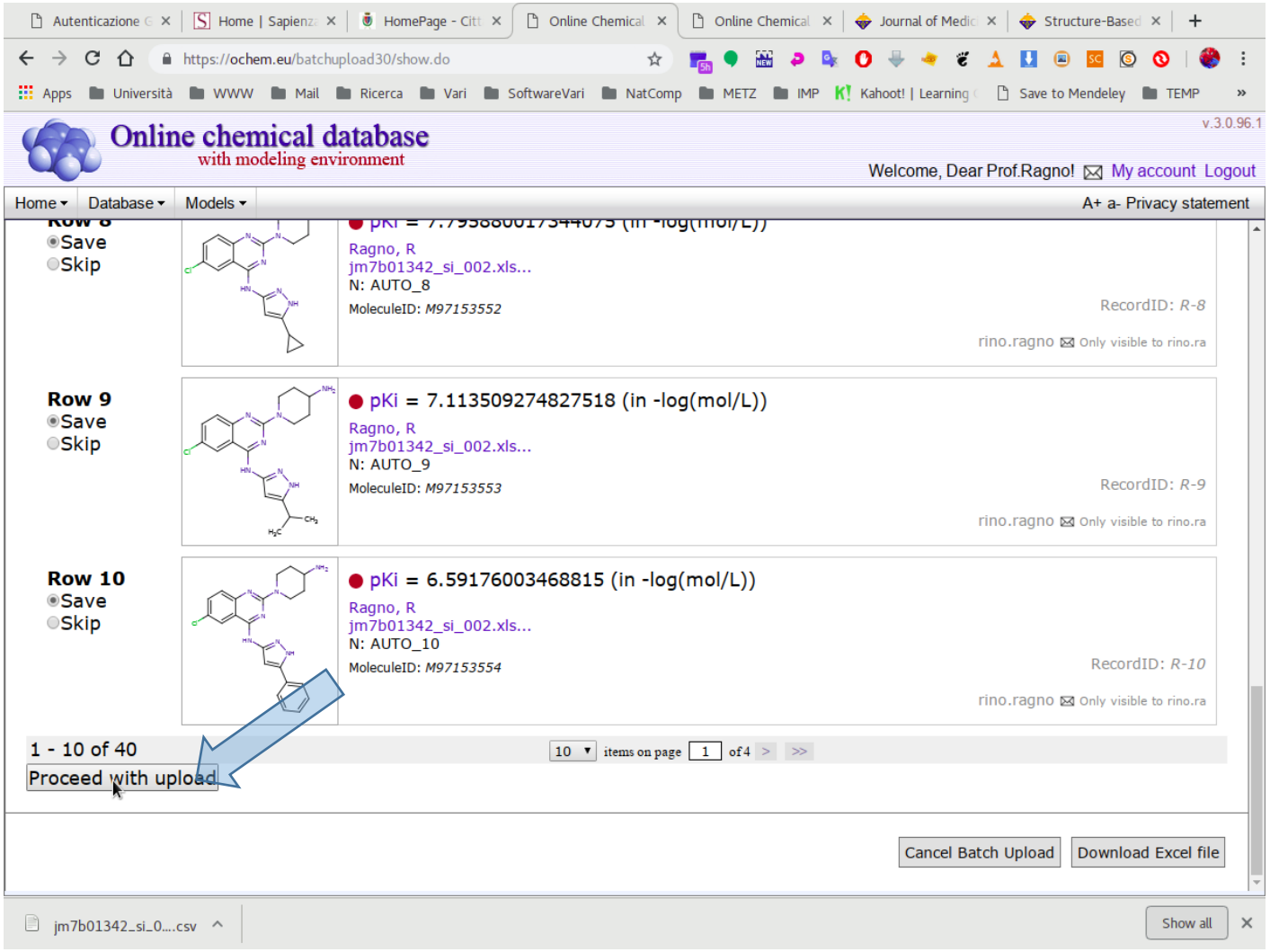
# Preparing the dataset



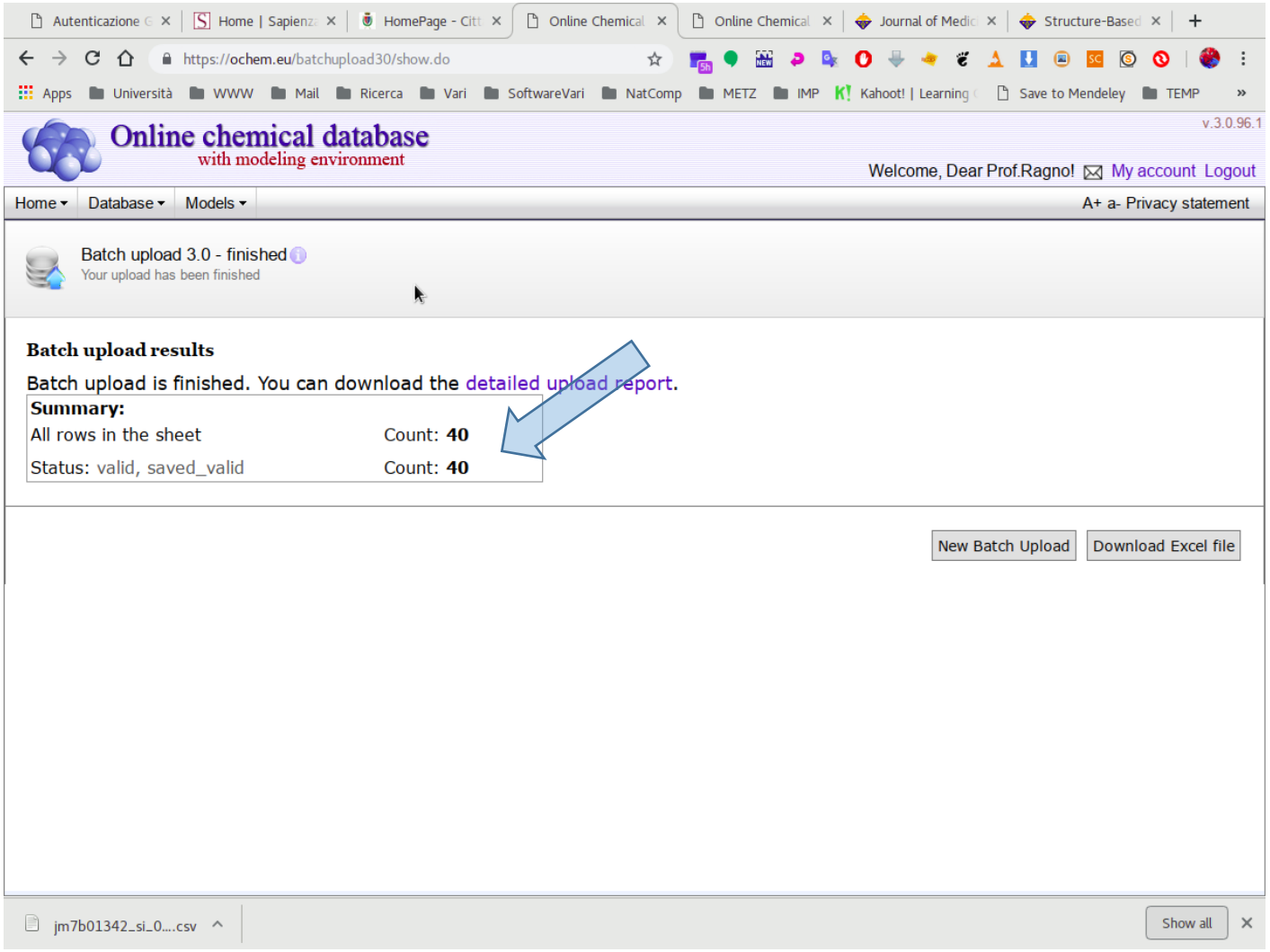
# Preparing the dataset



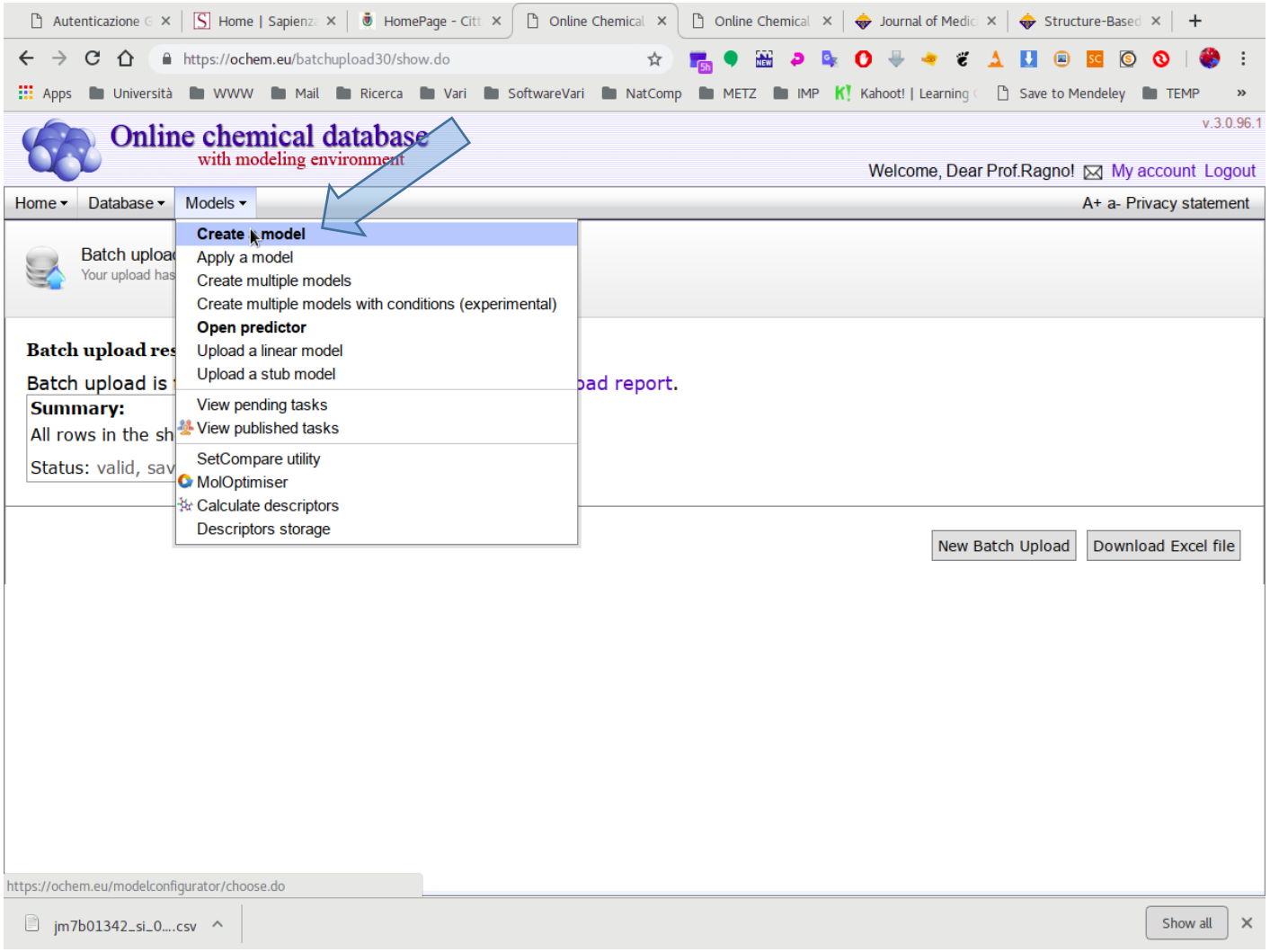
# Preparing the dataset



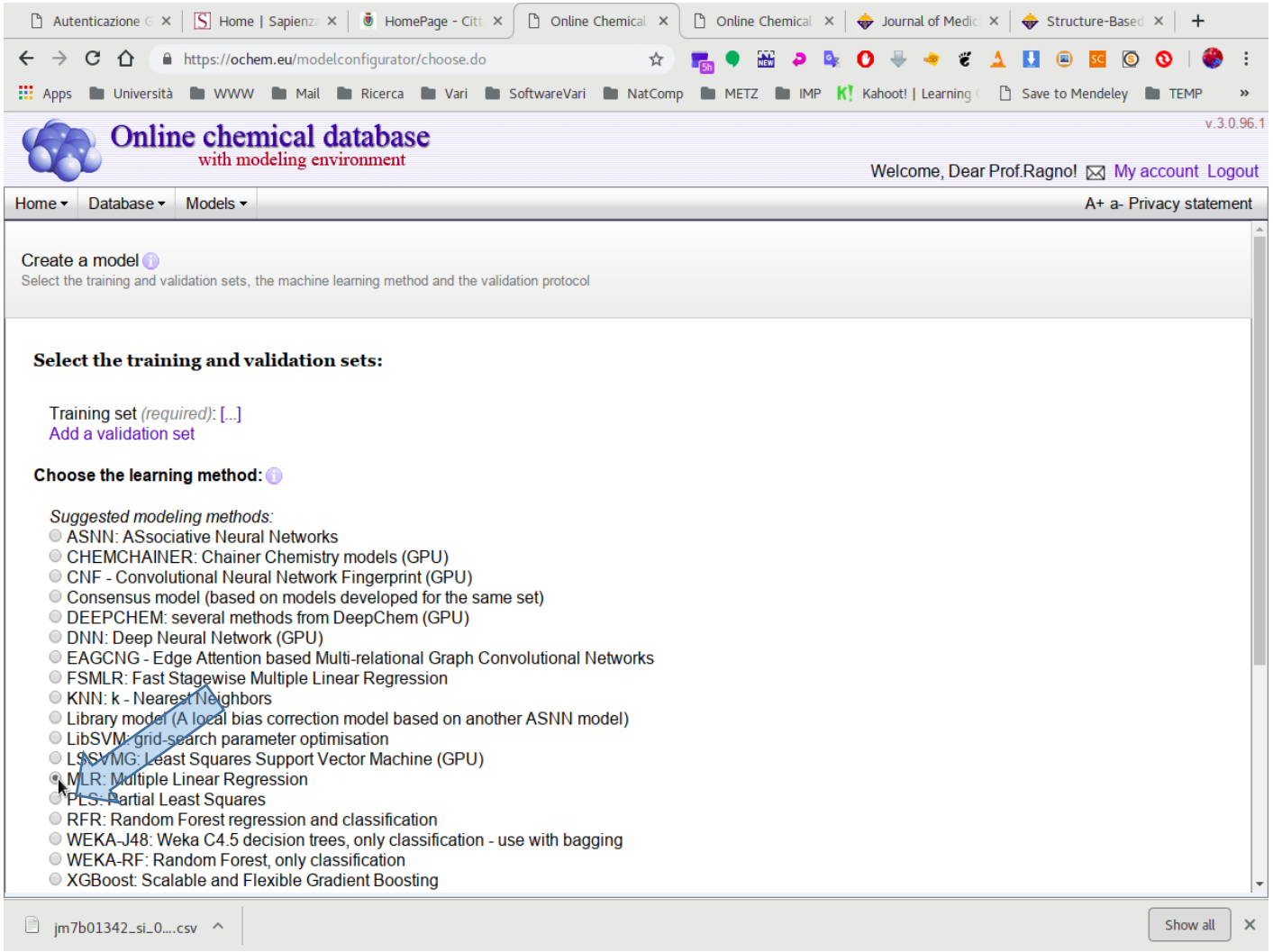
# Preparing the dataset



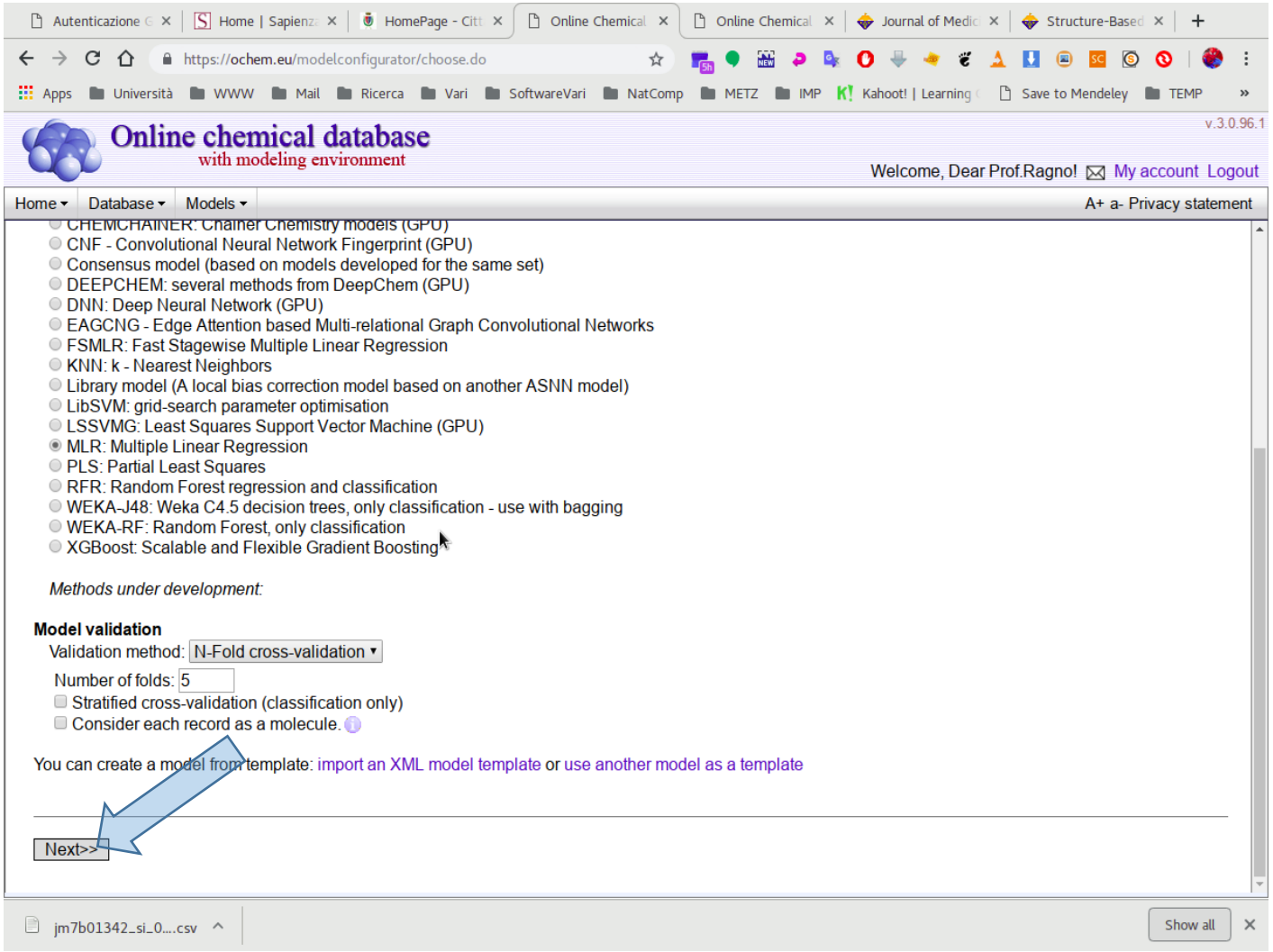
# Preparing the dataset



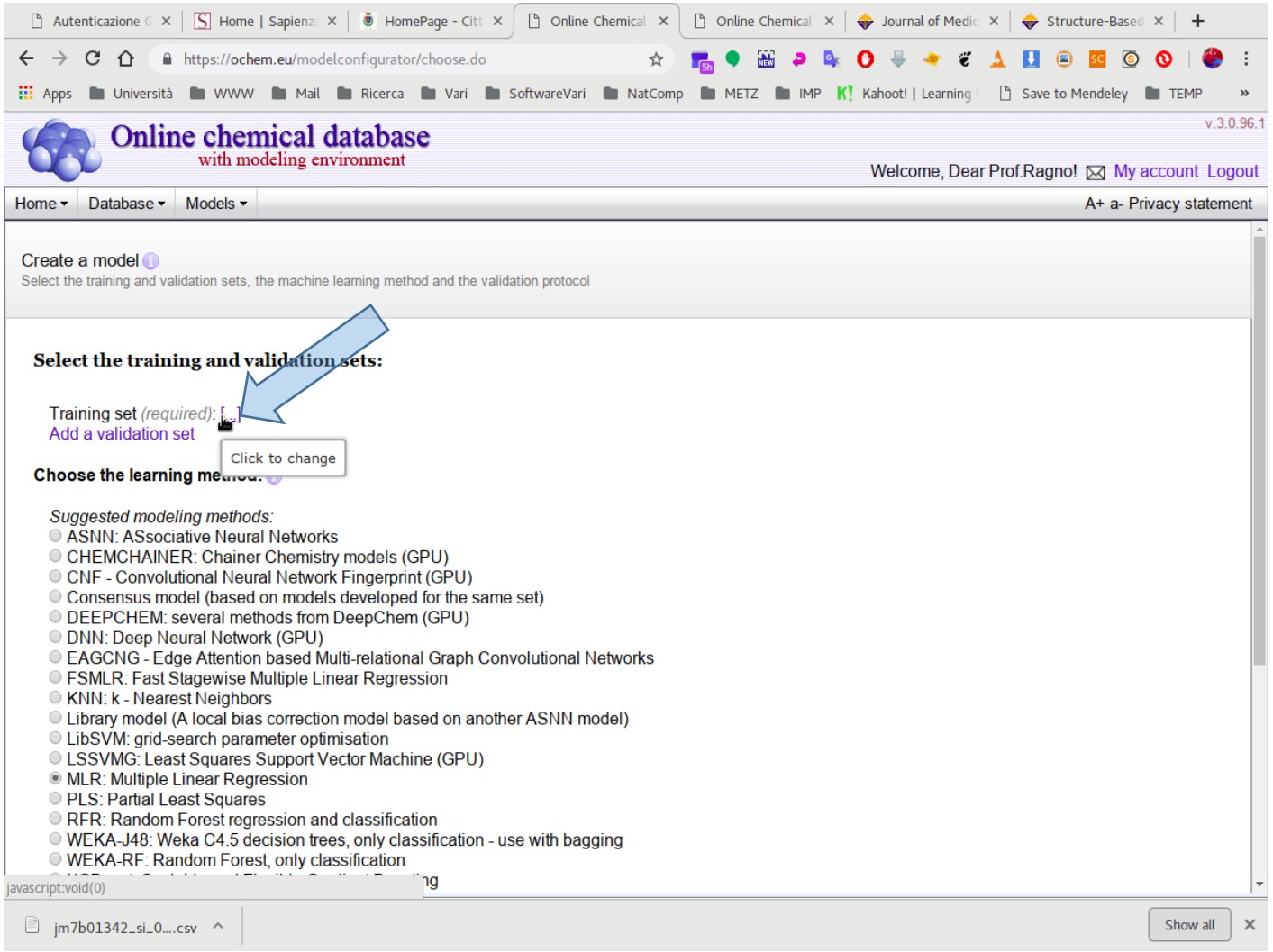
# Preparing the dataset



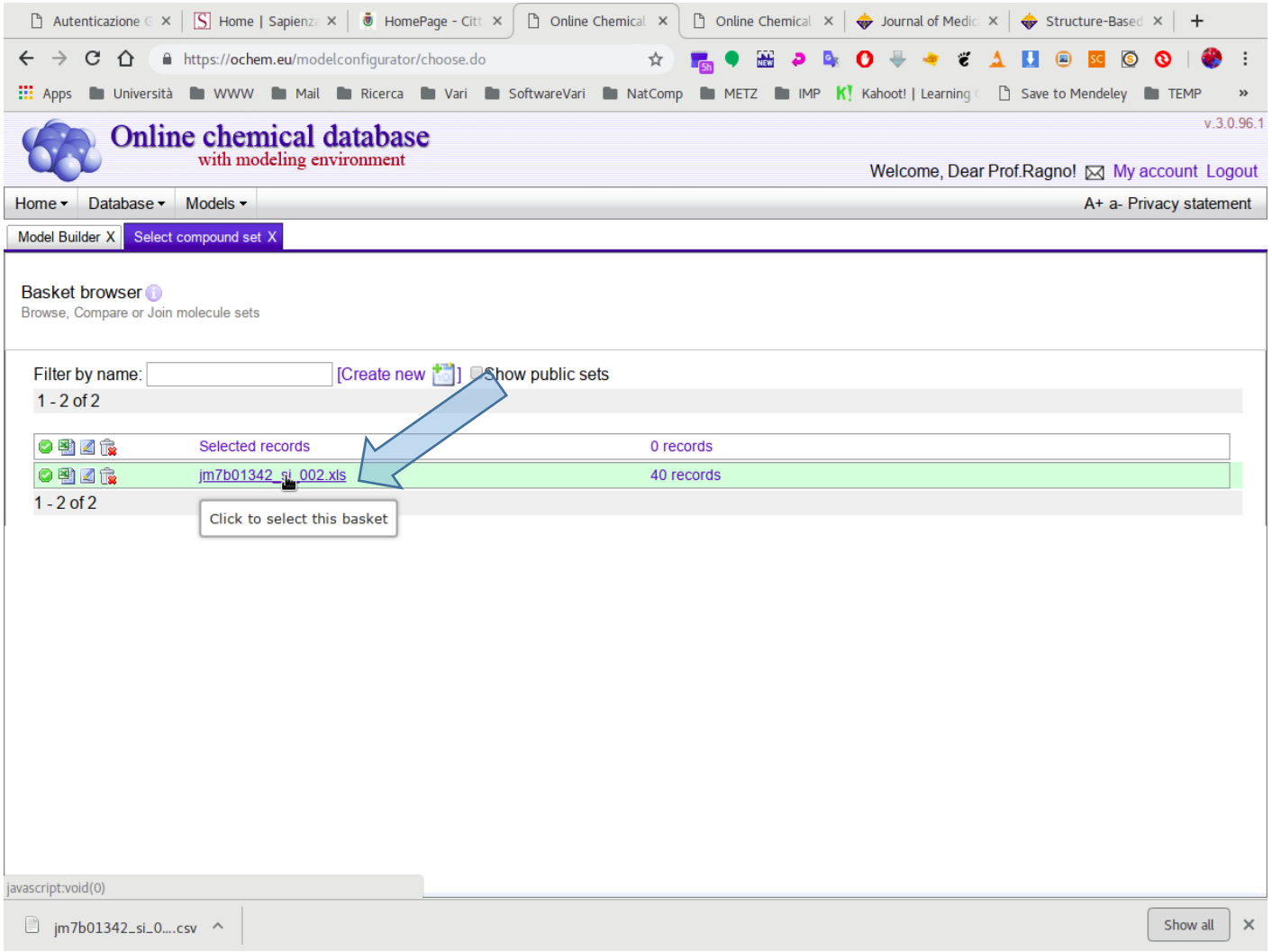
# Preparing the dataset



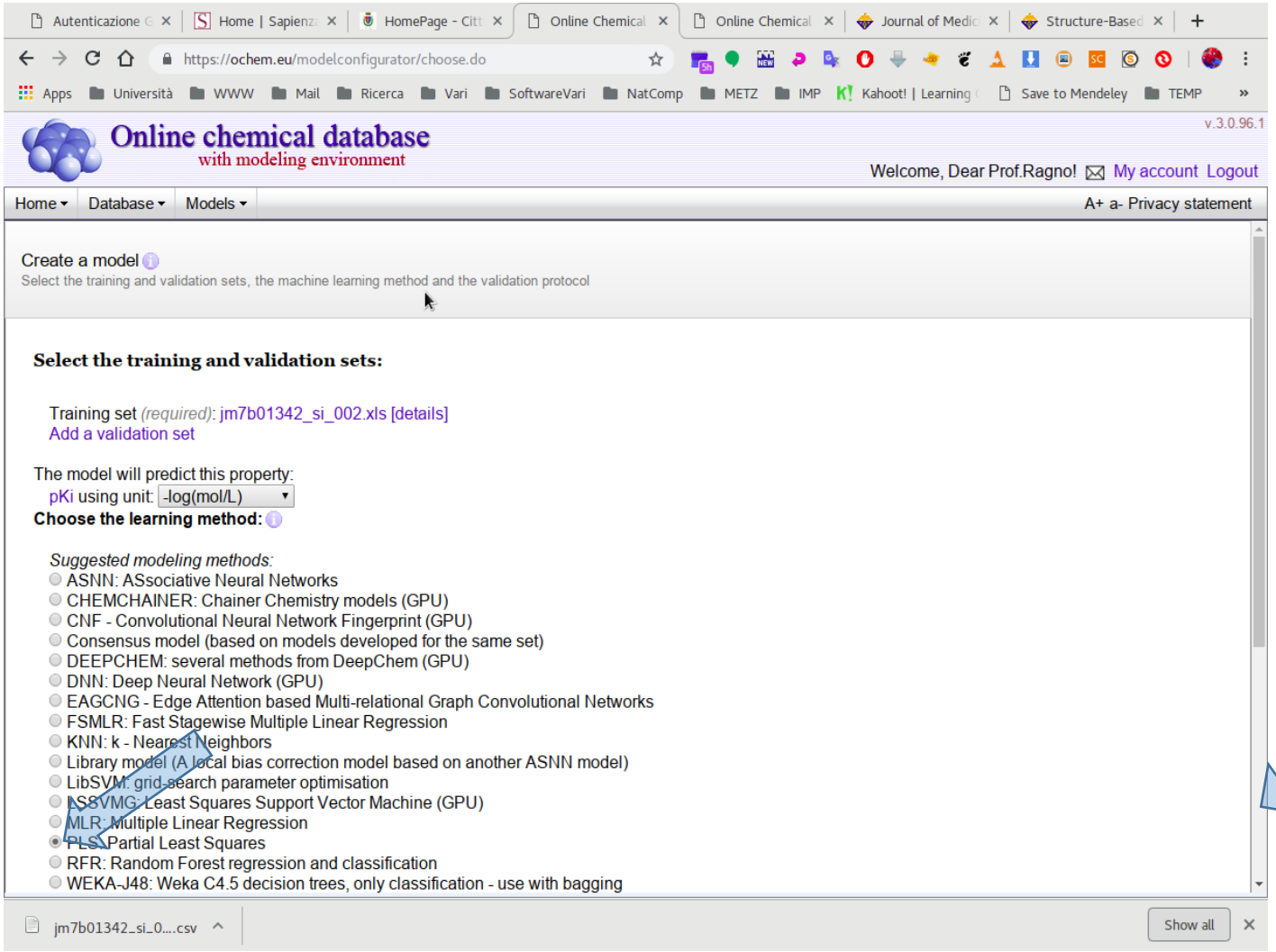
# Preparing the dataset



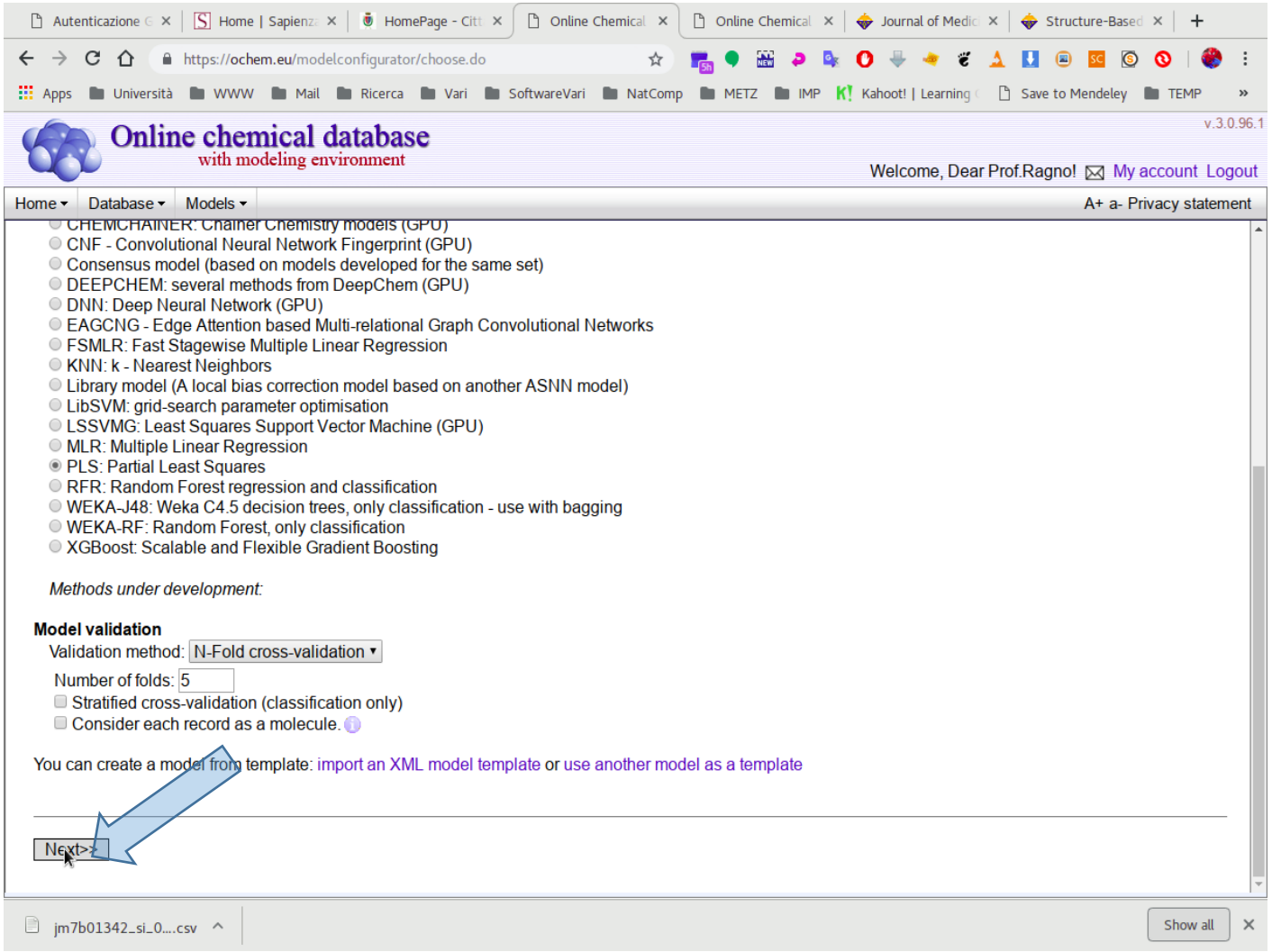
# Preparing the dataset



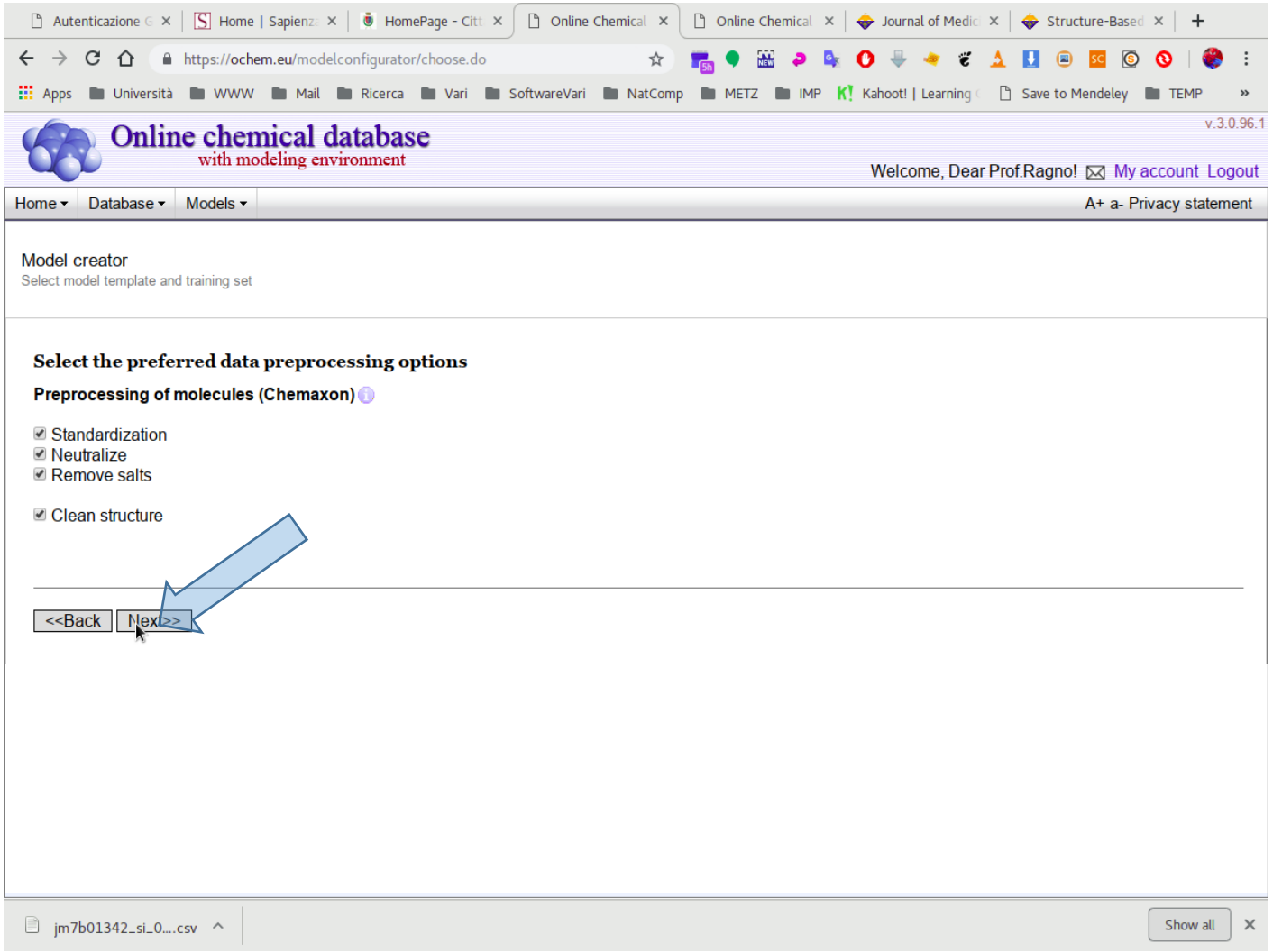
# Preparing the dataset



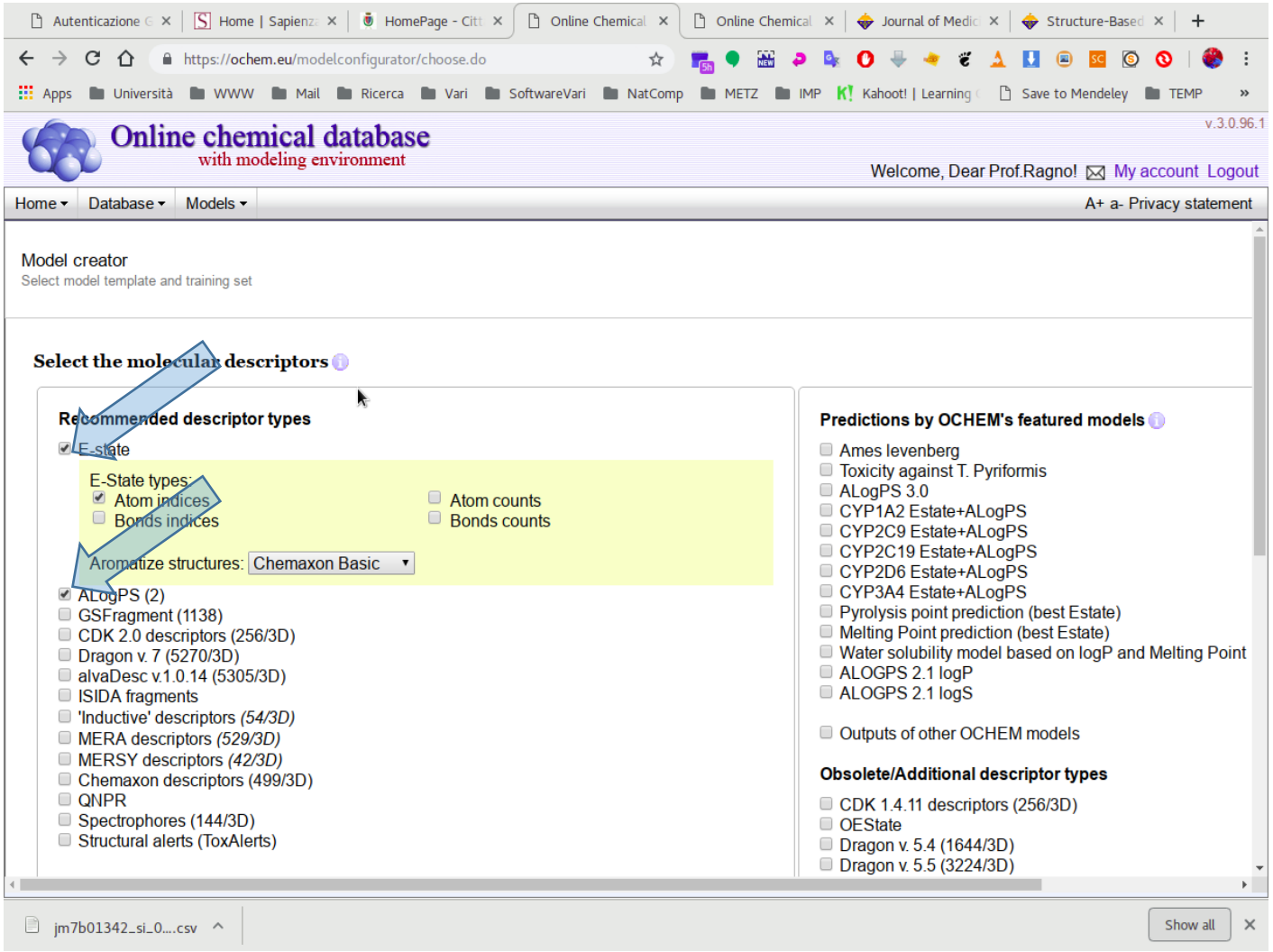
# Preparing the dataset



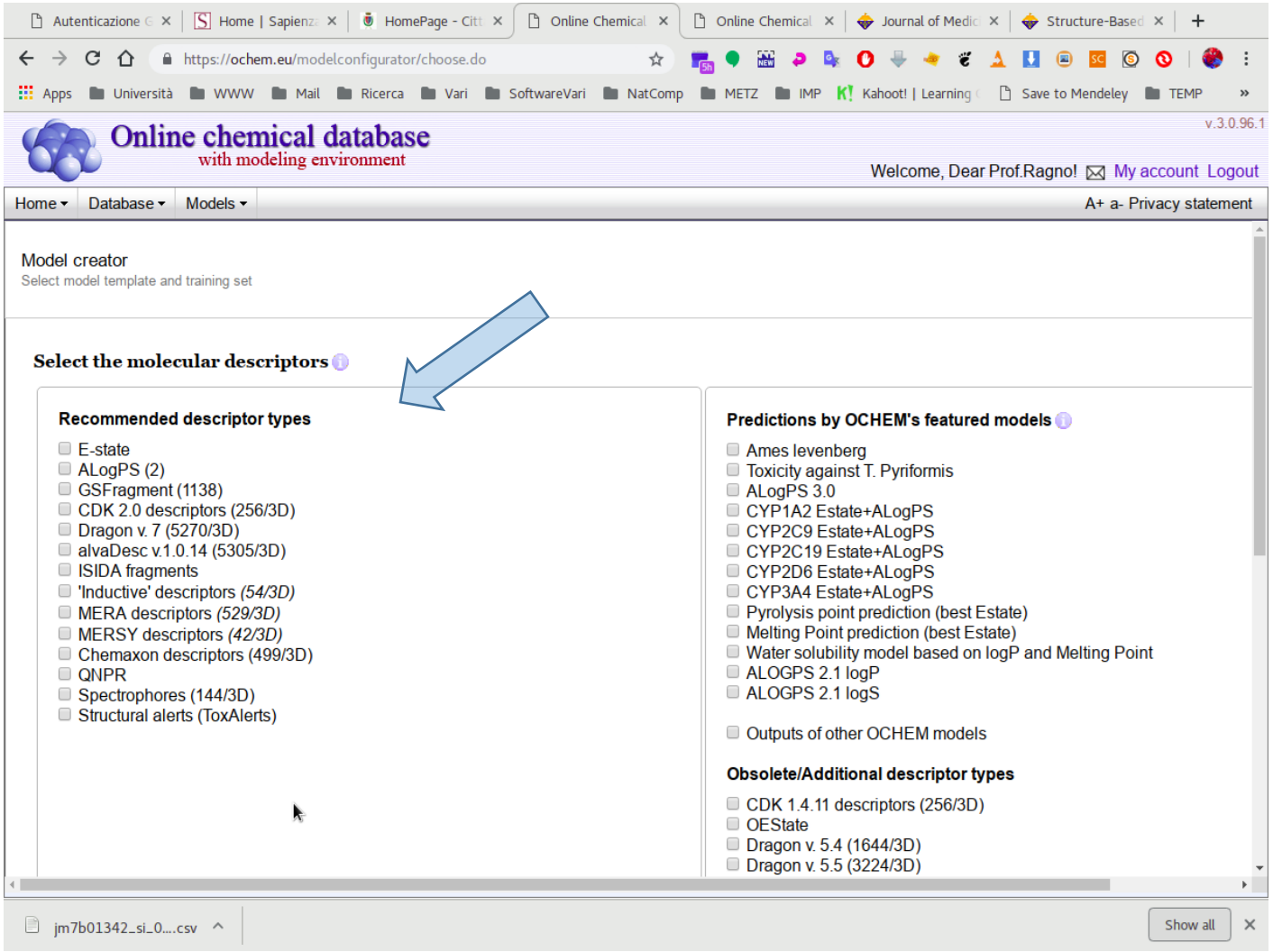
# Preparing the dataset



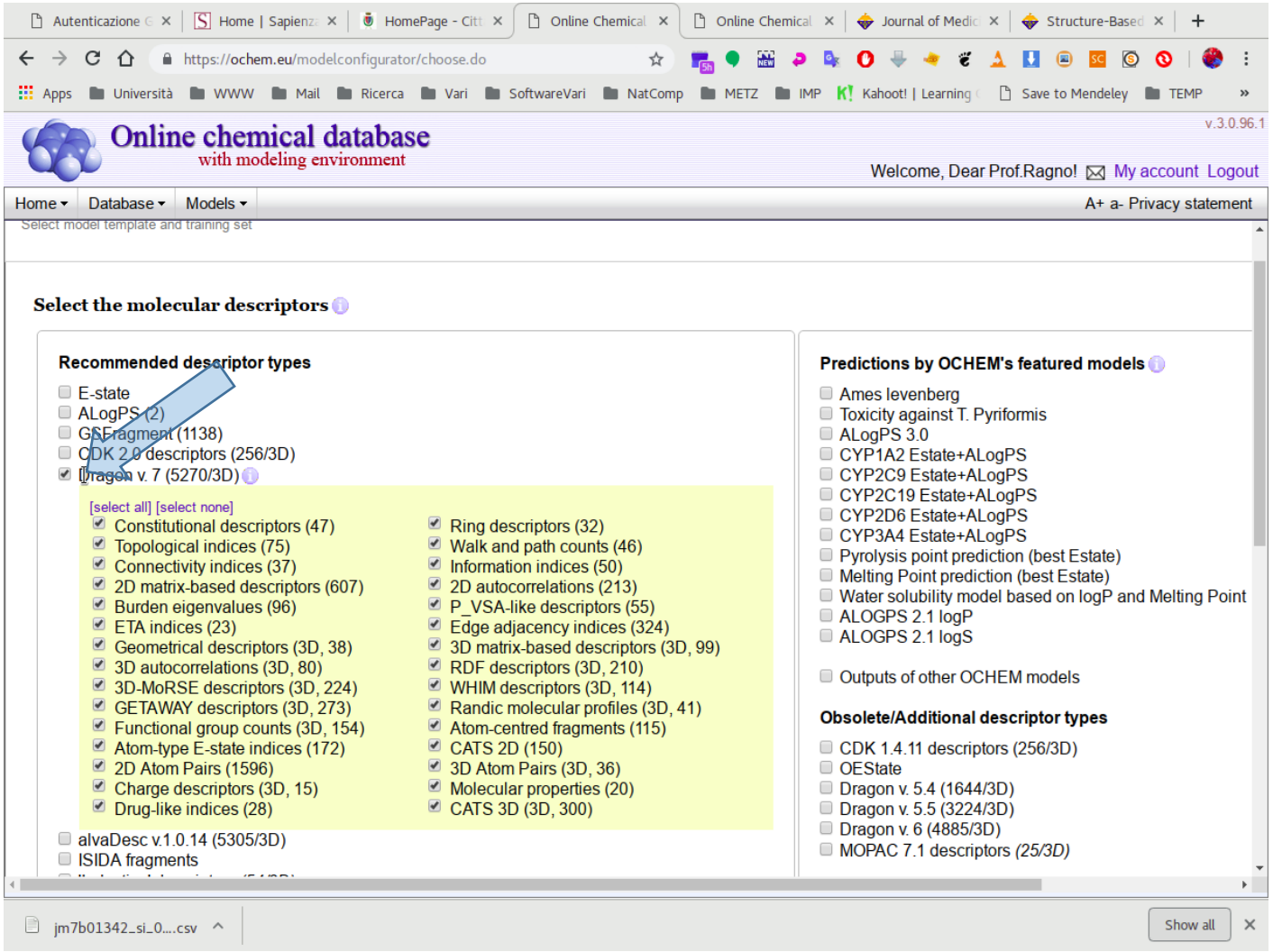
# Preparing the dataset



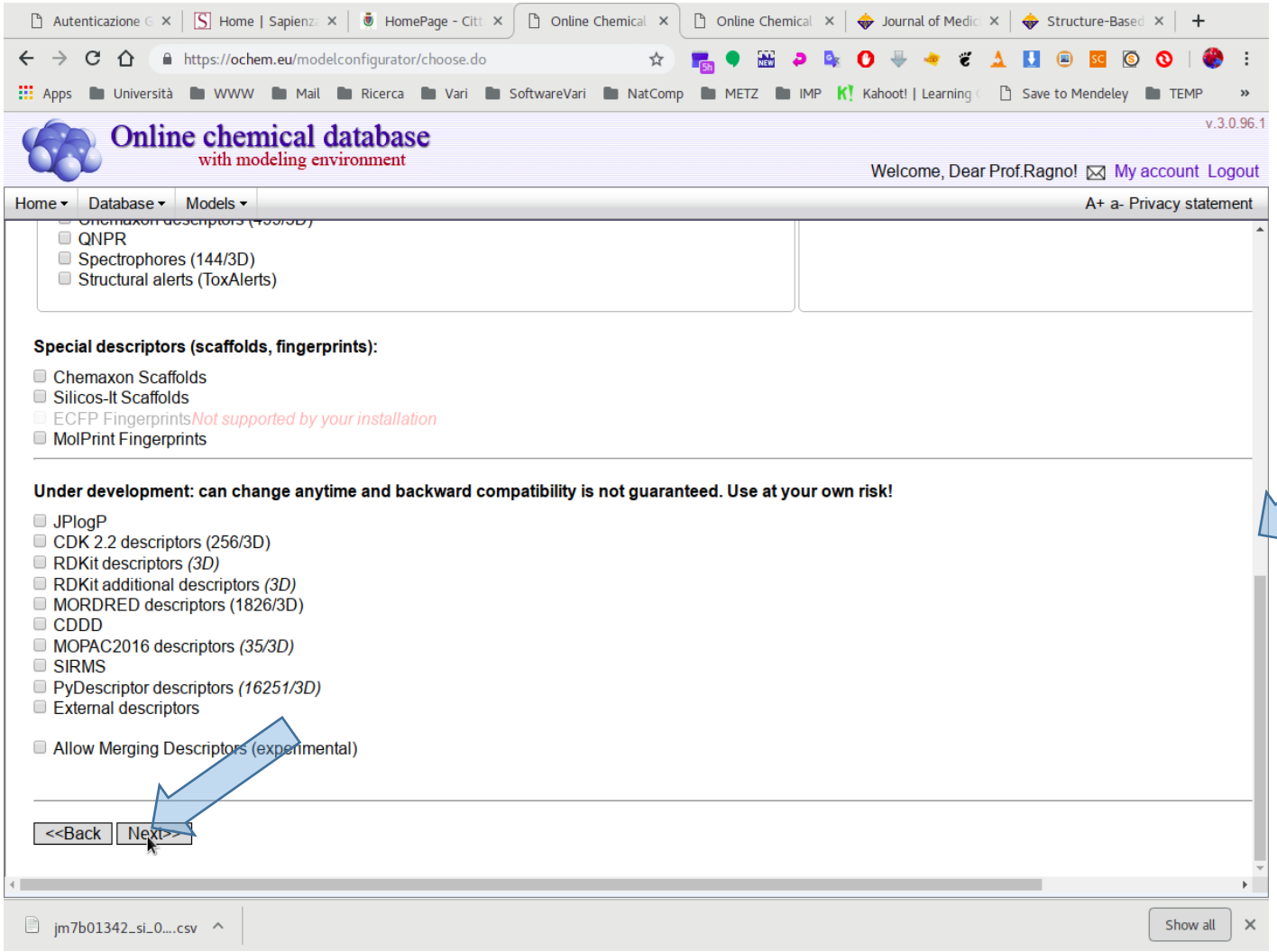
# Preparing the dataset



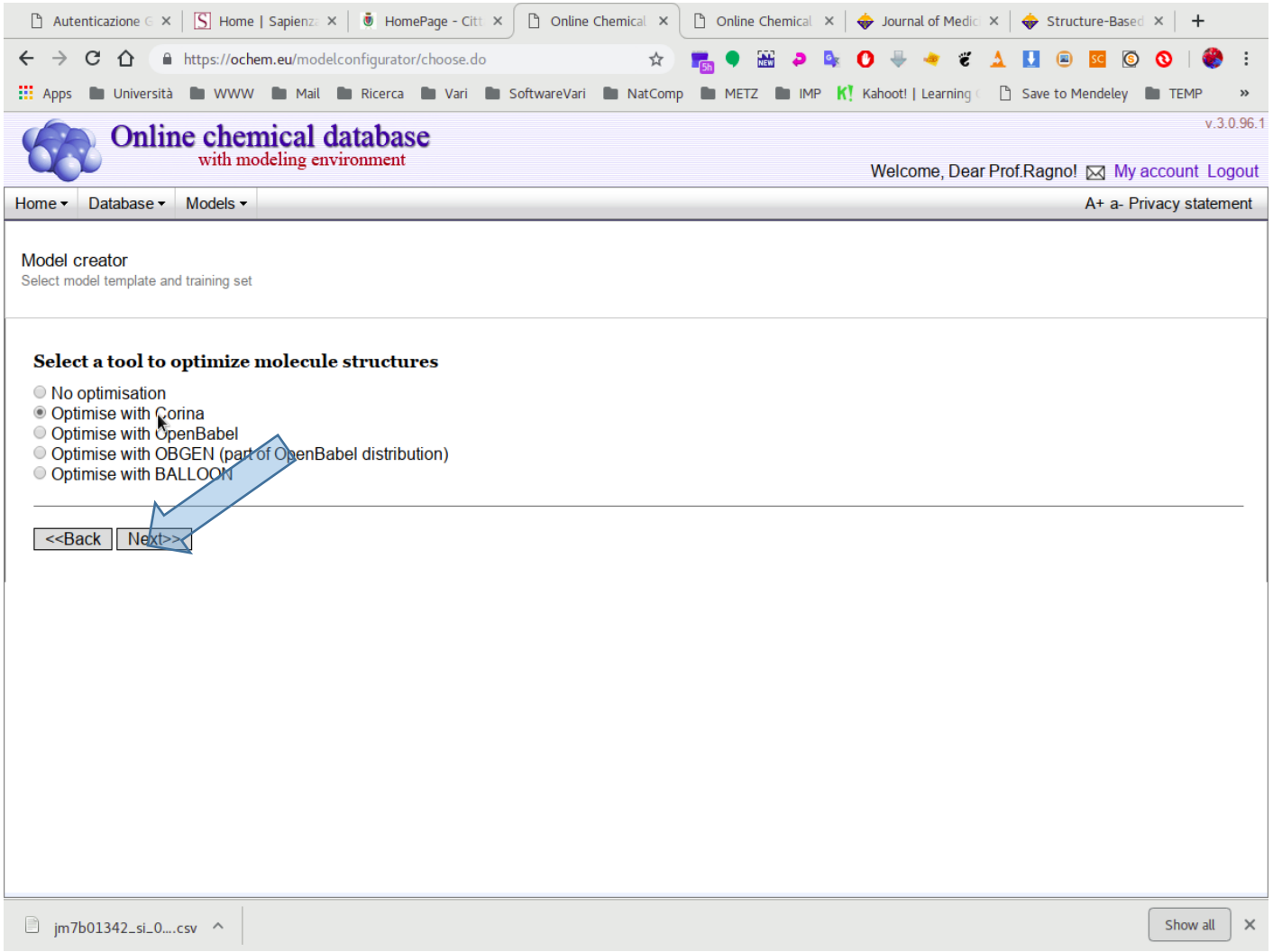
# Preparing the dataset



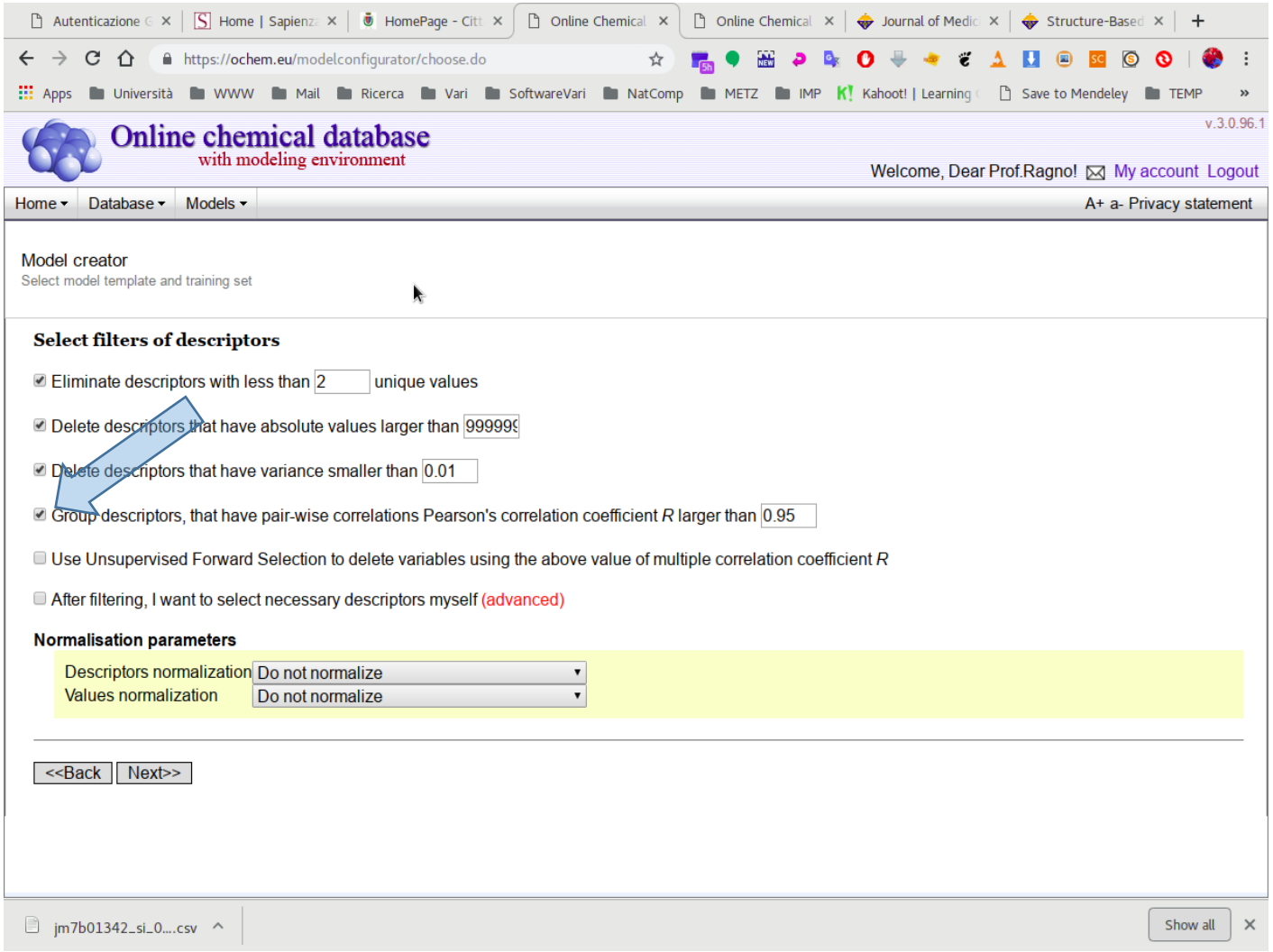
# Preparing the dataset



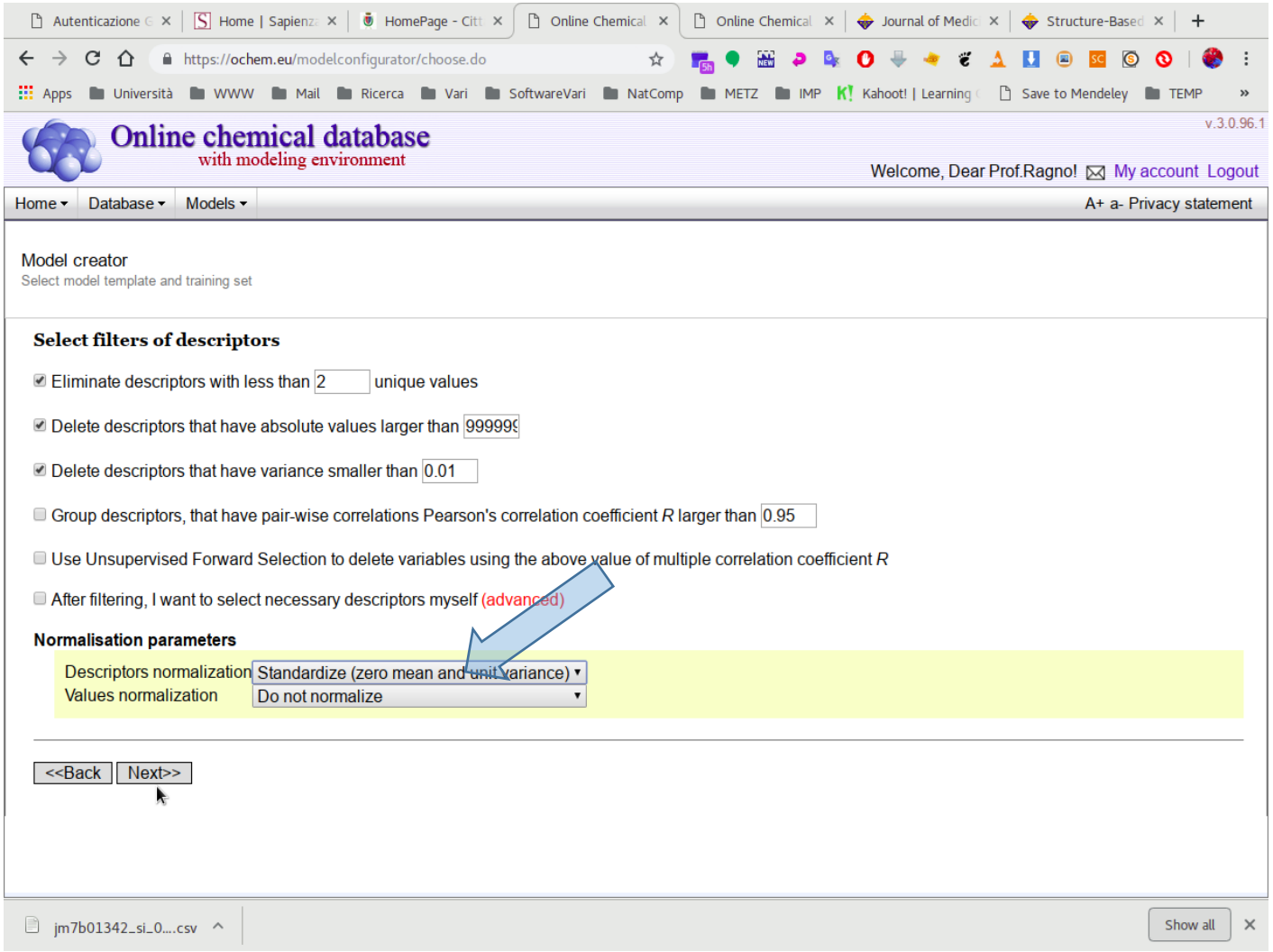
# Preparing the dataset



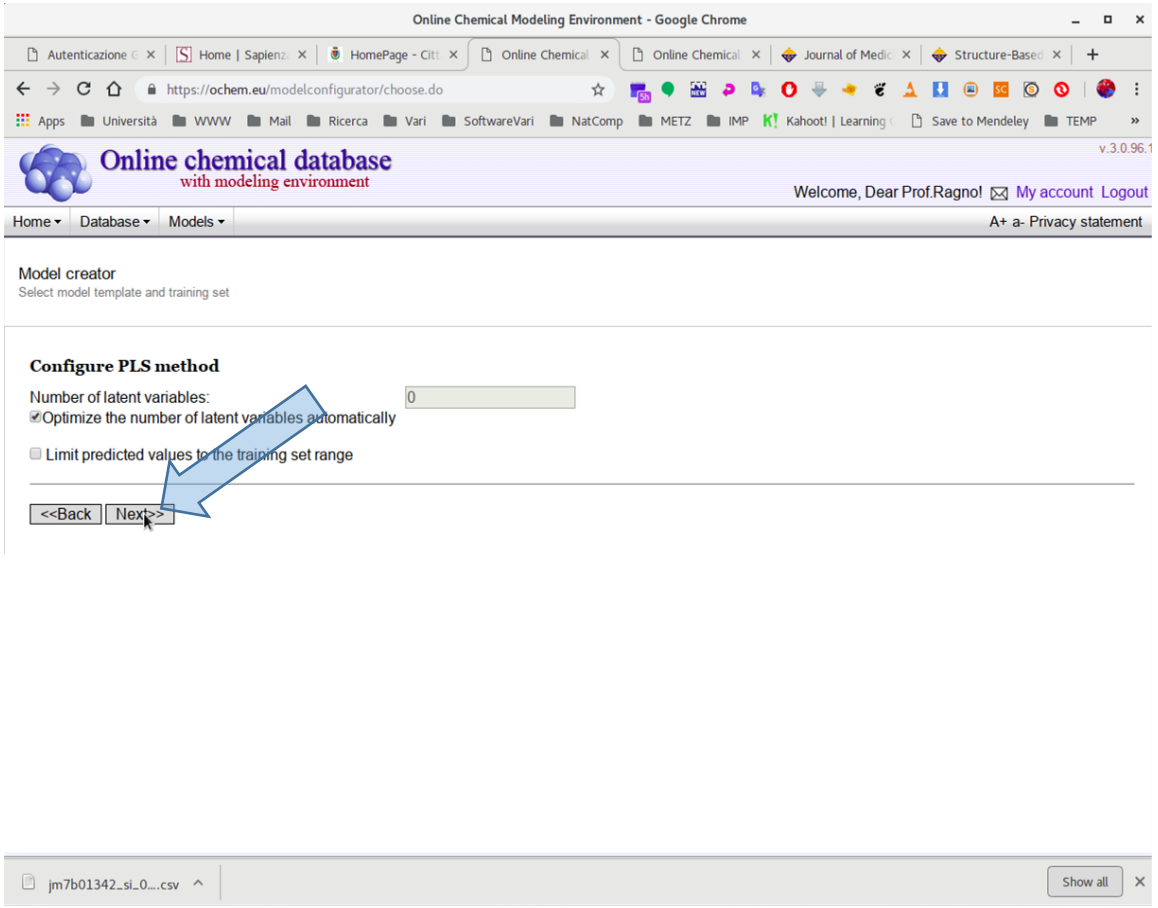
# Preparing the dataset



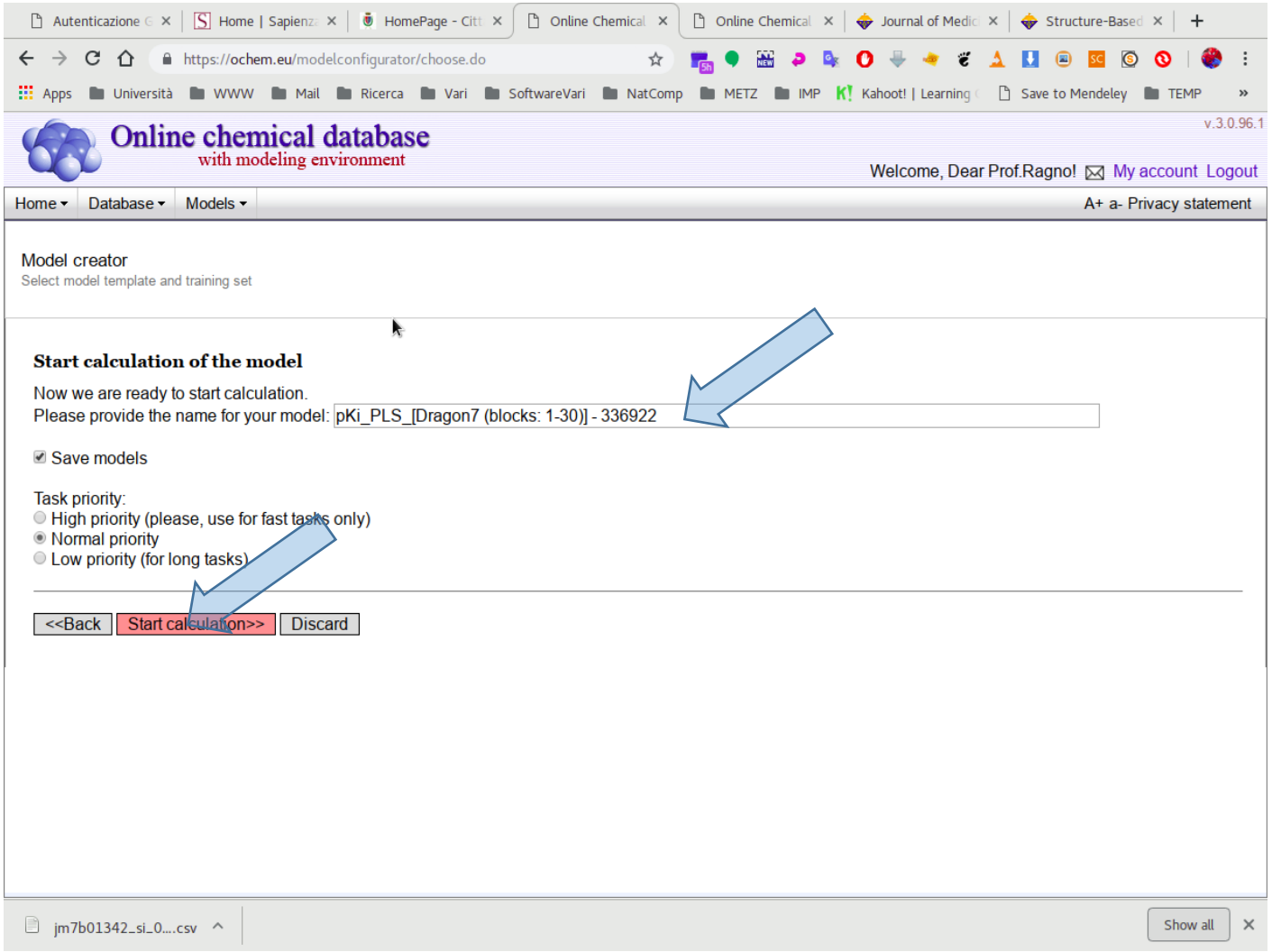
# Preparing the dataset



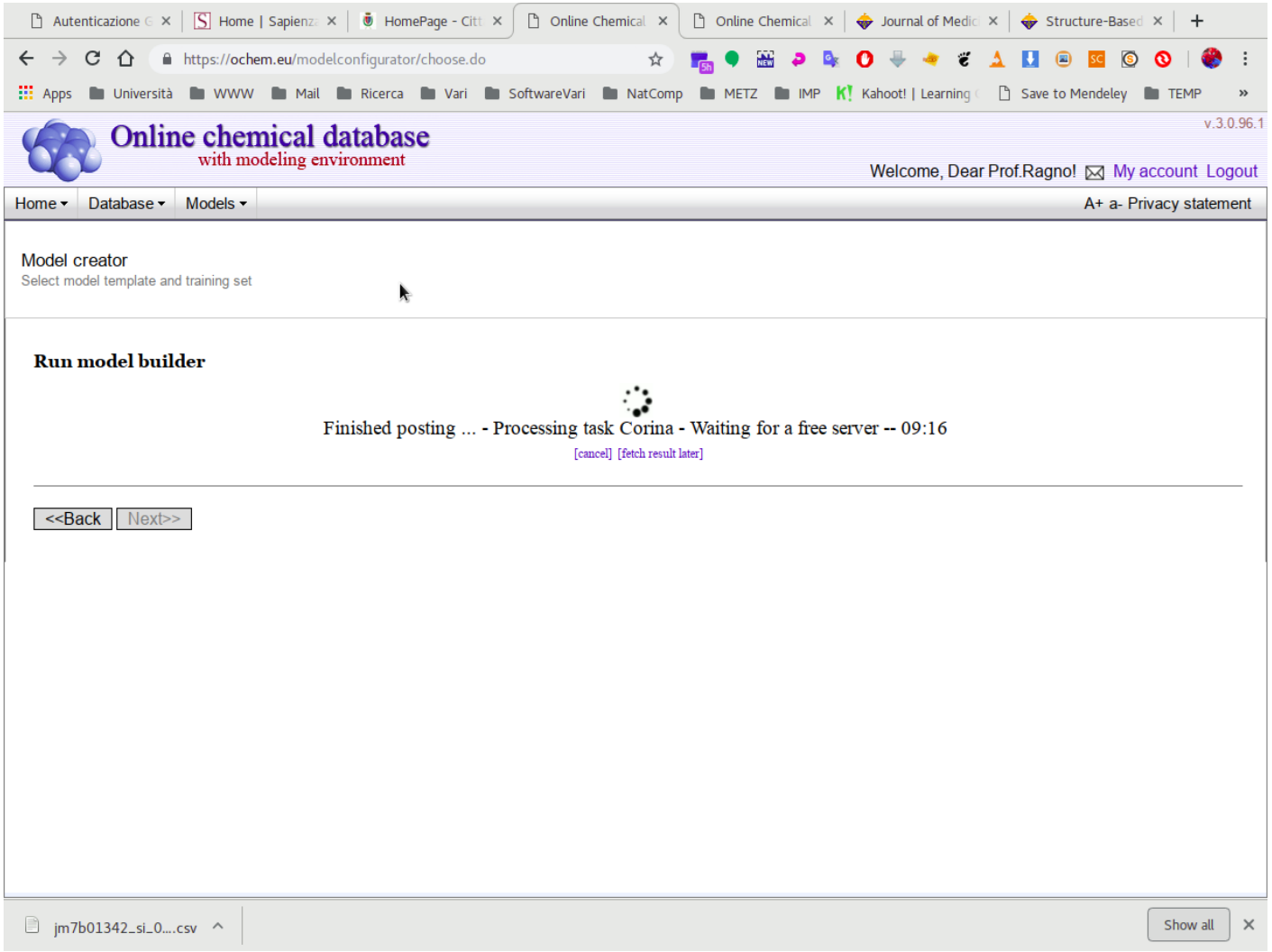
# Preparing the dataset



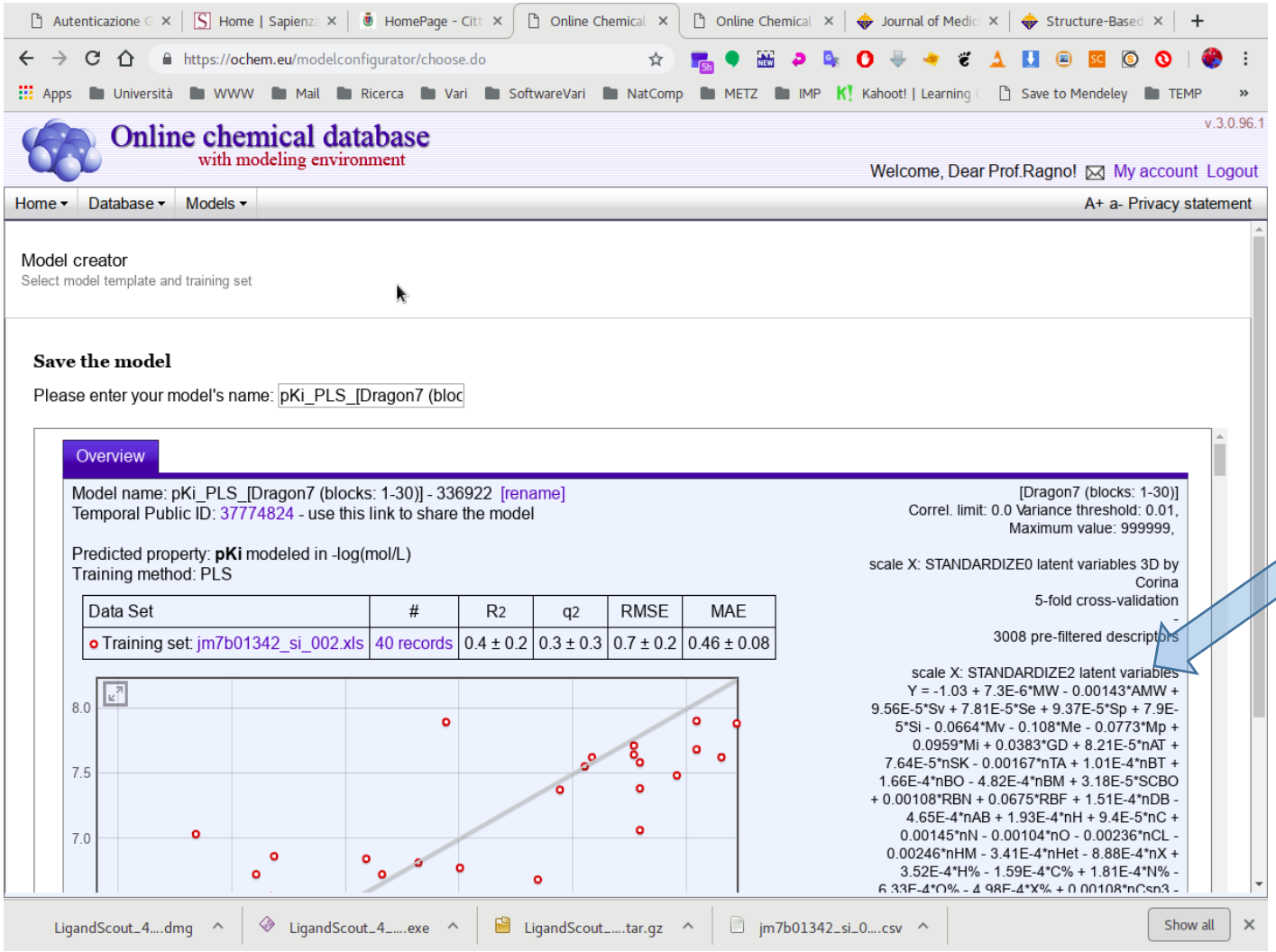
# Preparing the dataset



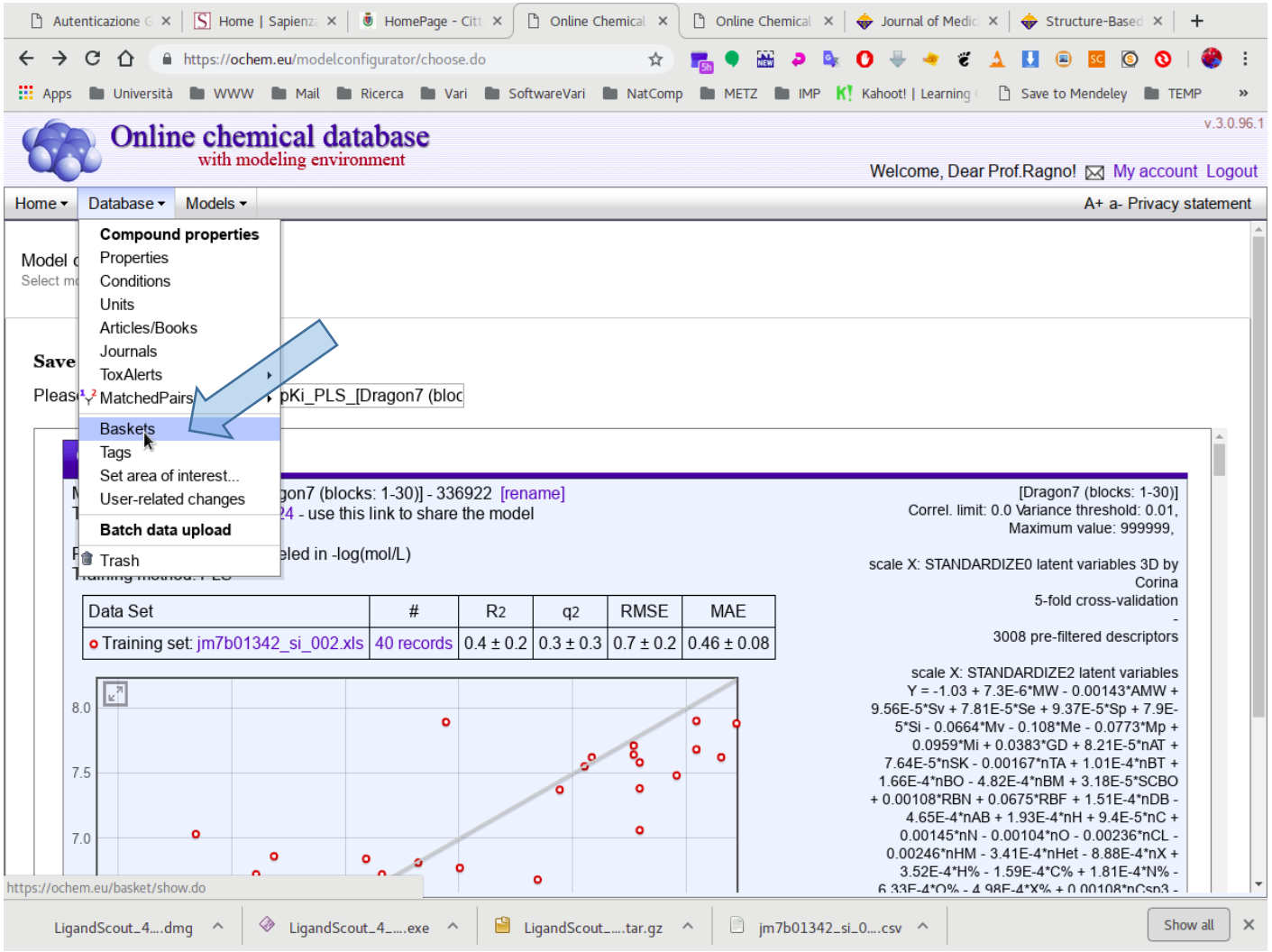
# Preparing the dataset



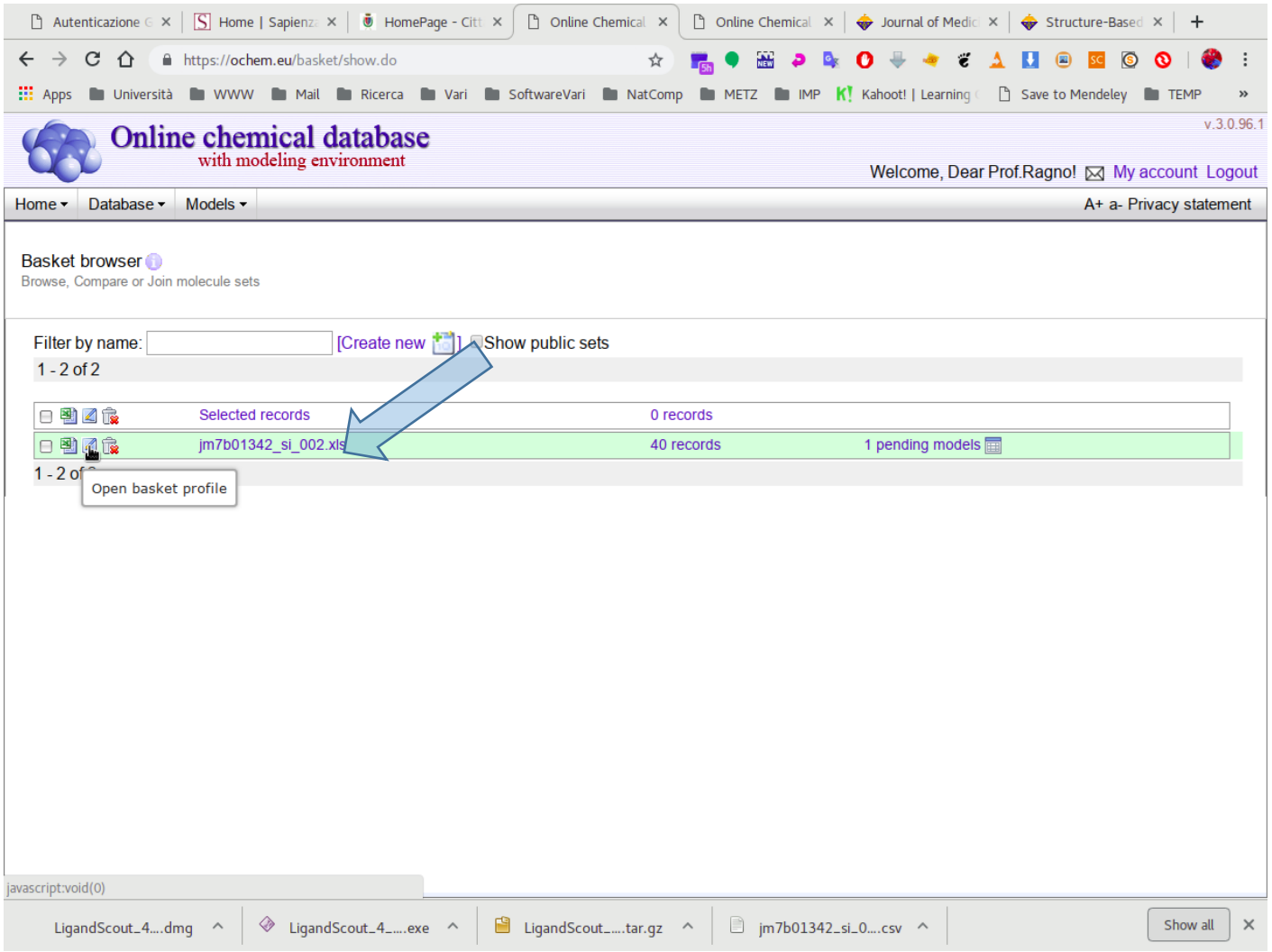
# Preparing the dataset



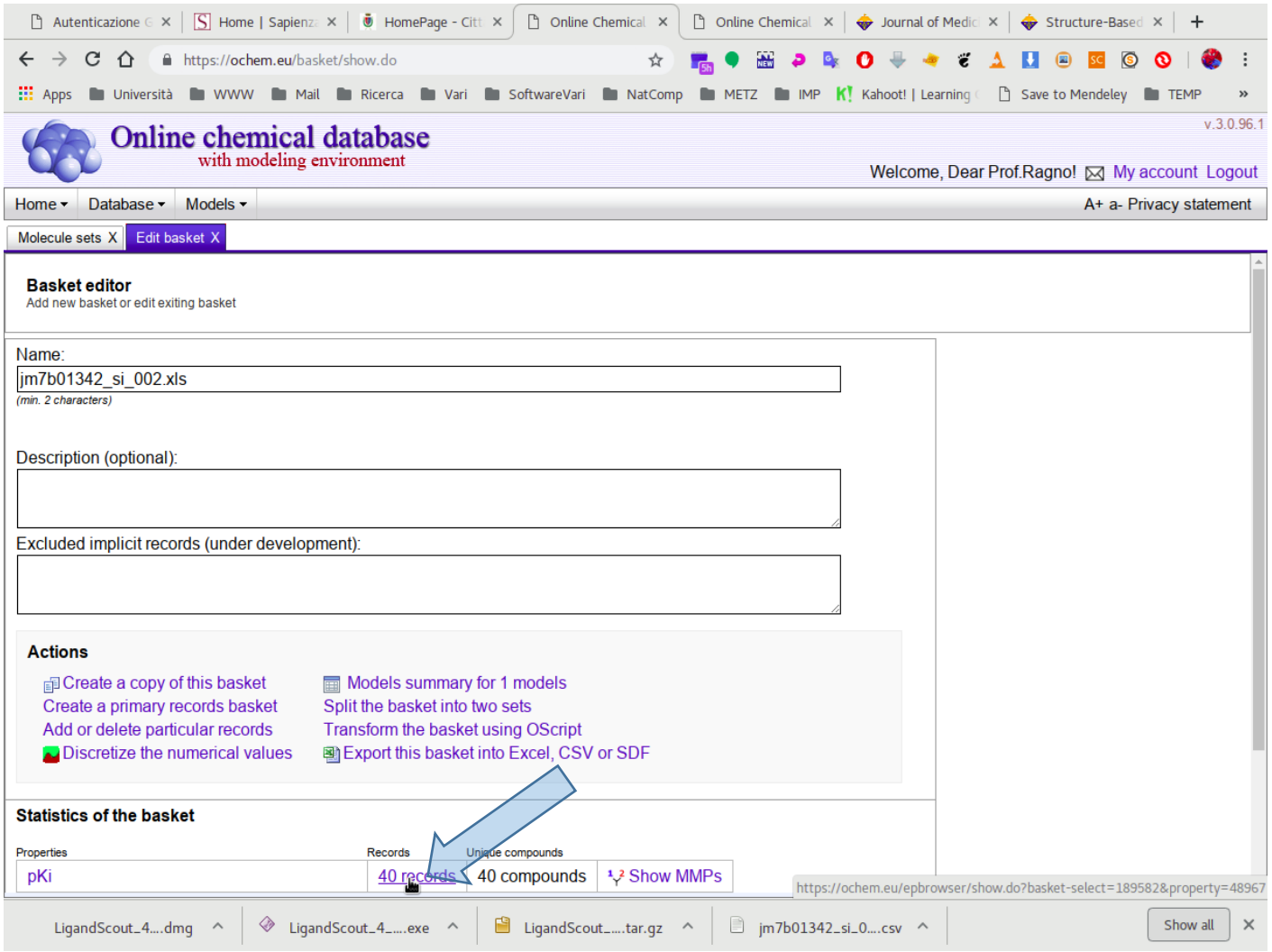
# Preparing the dataset



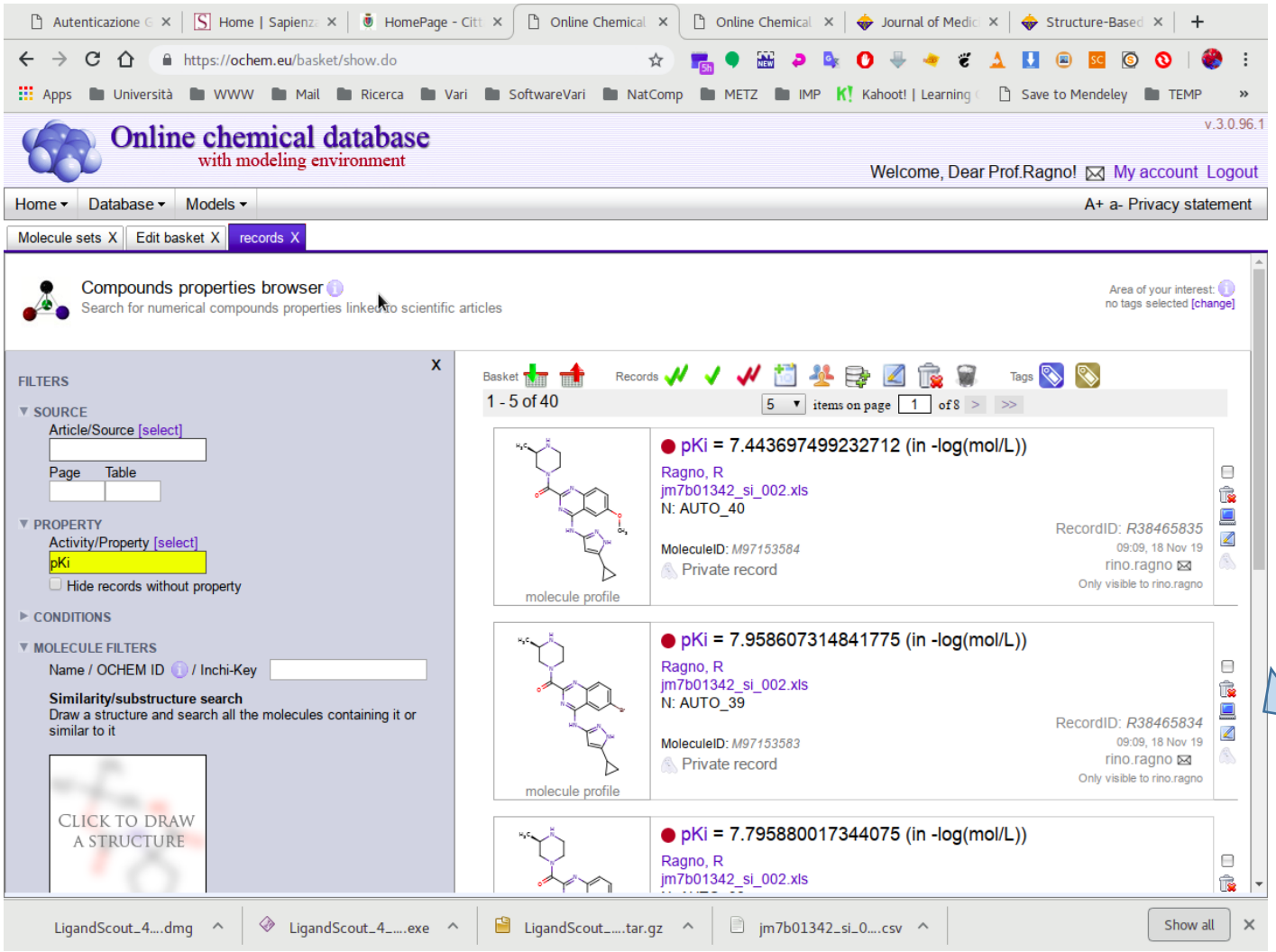
# Preparing the dataset



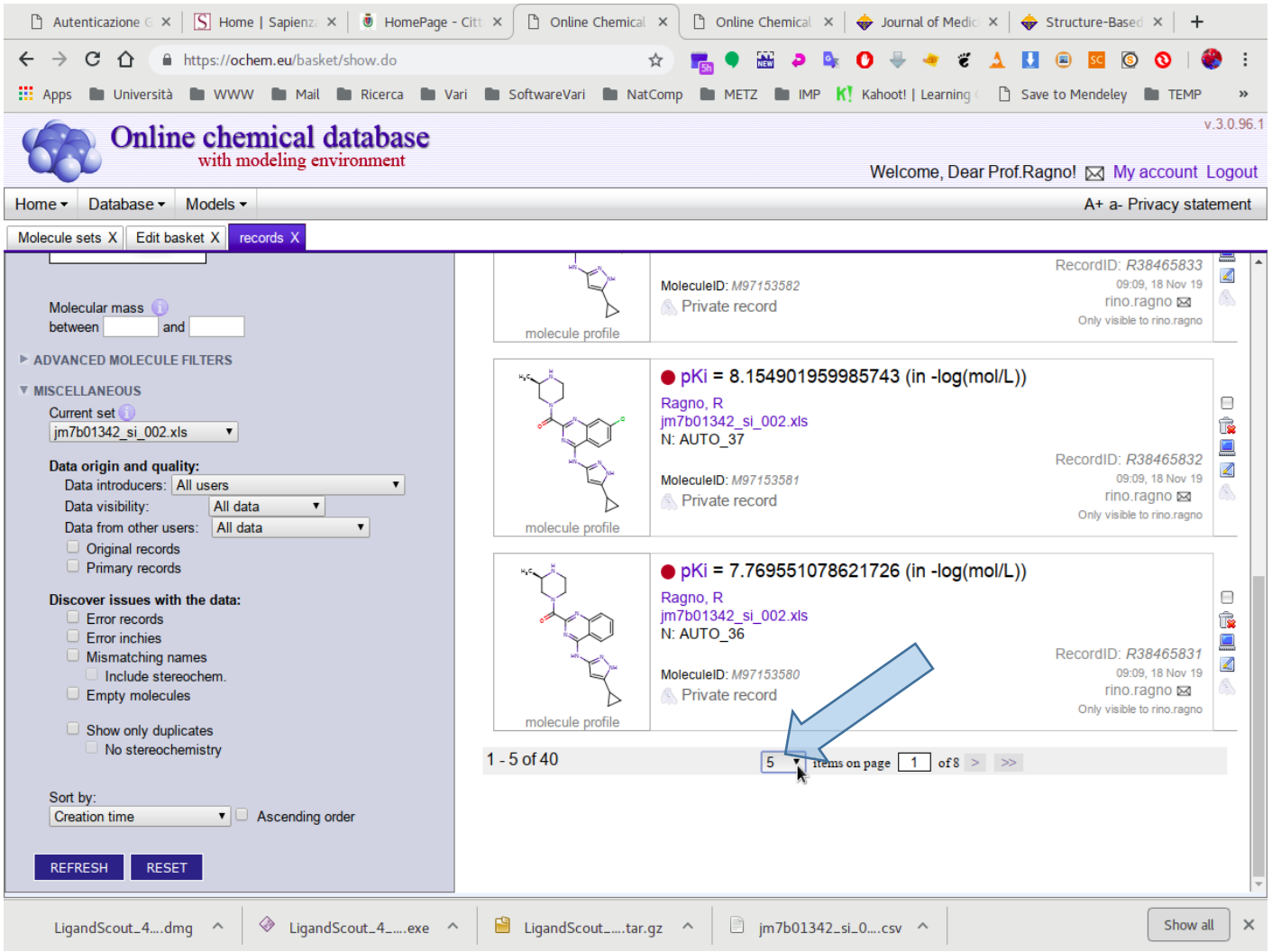
# Preparing the dataset



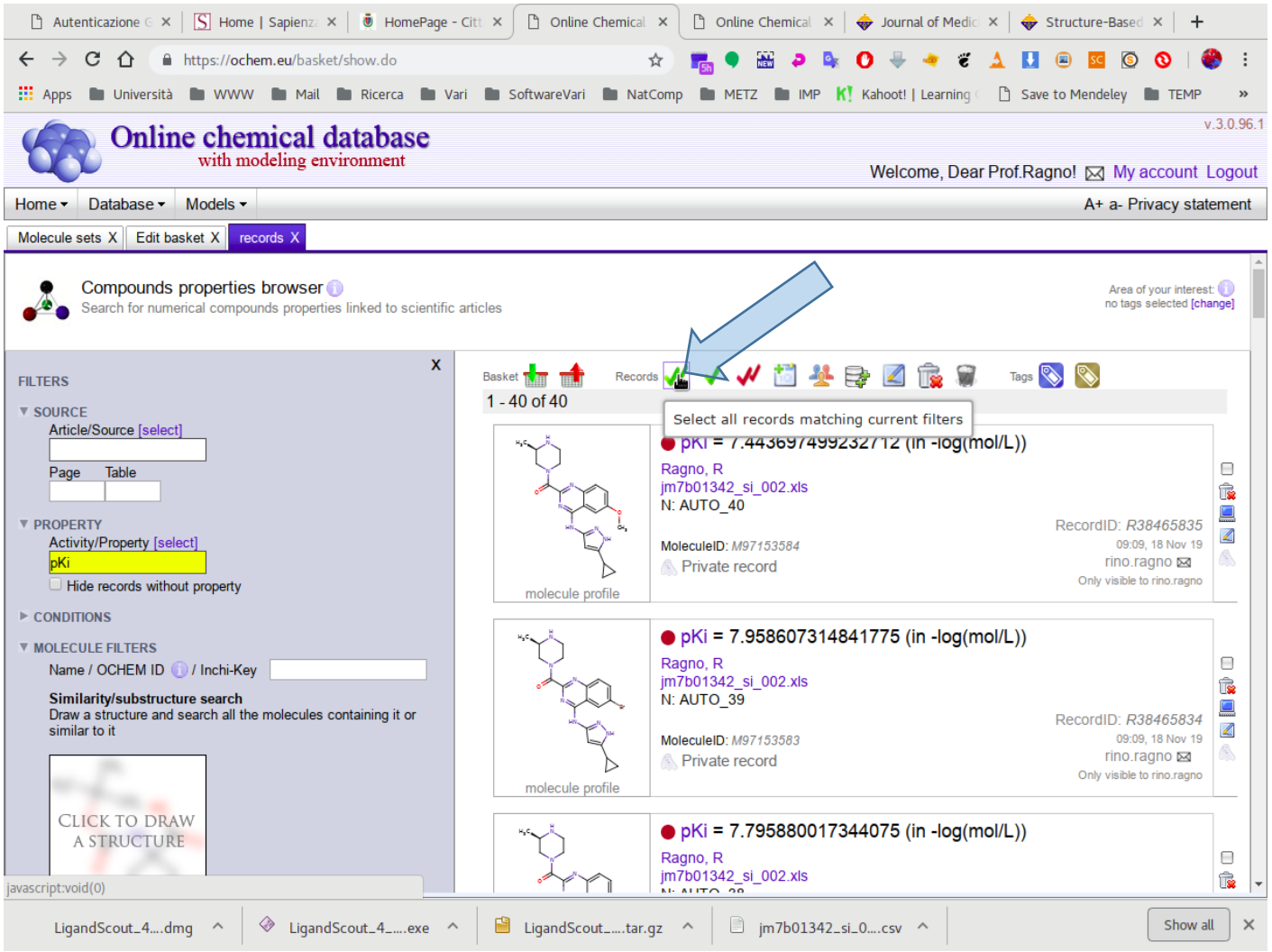
# Preparing the dataset



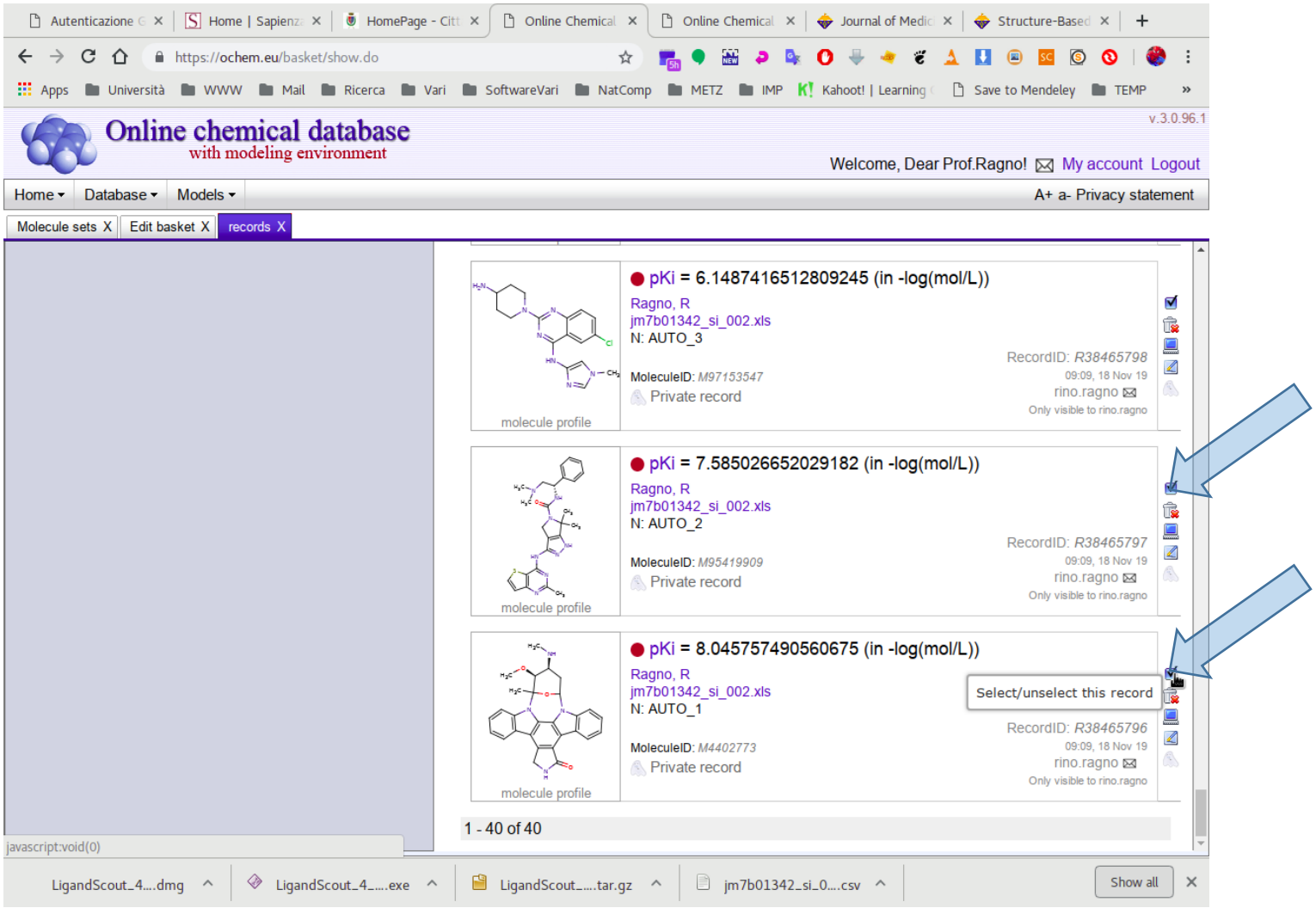
# Preparing the dataset



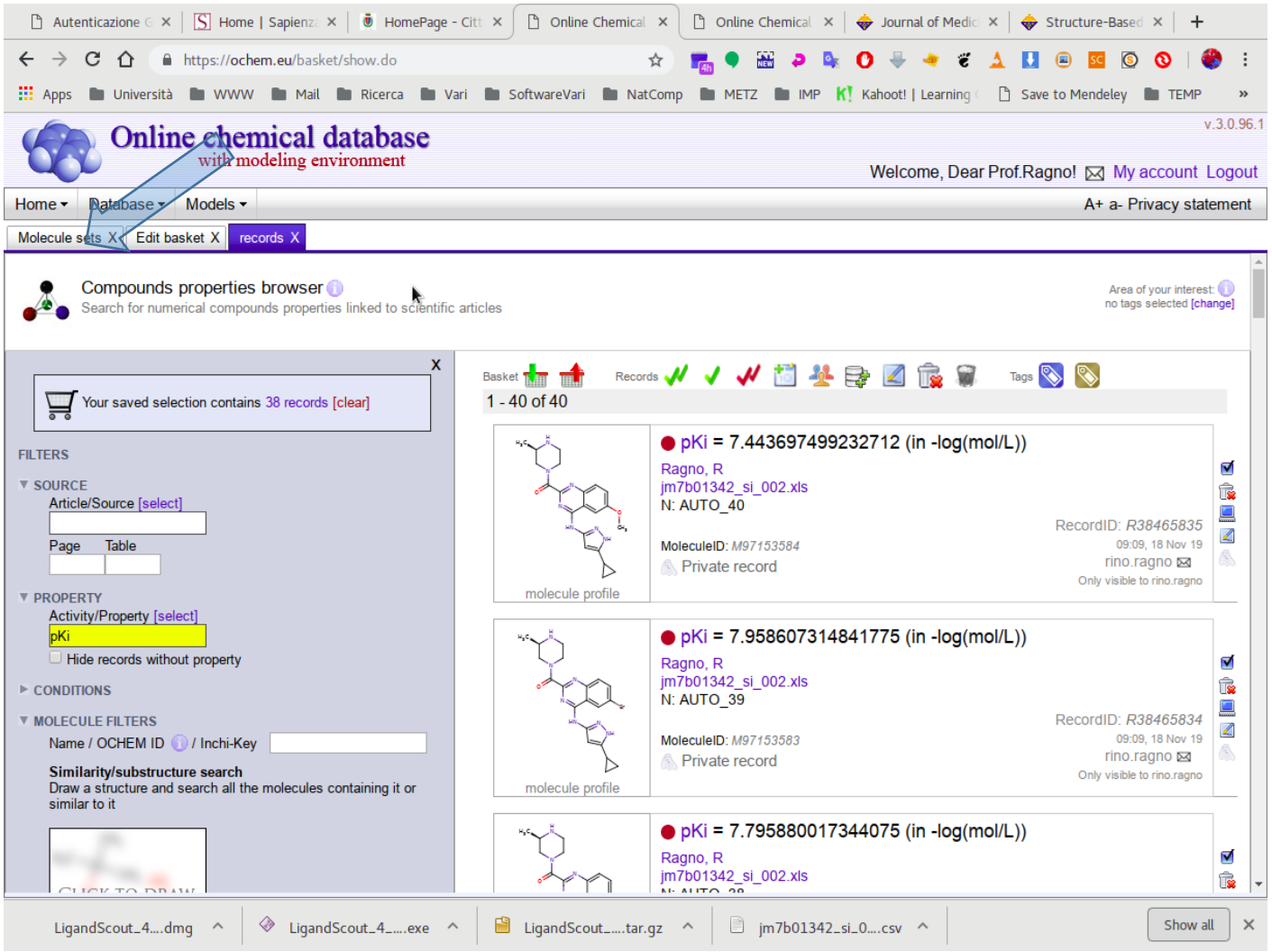
# Preparing the dataset



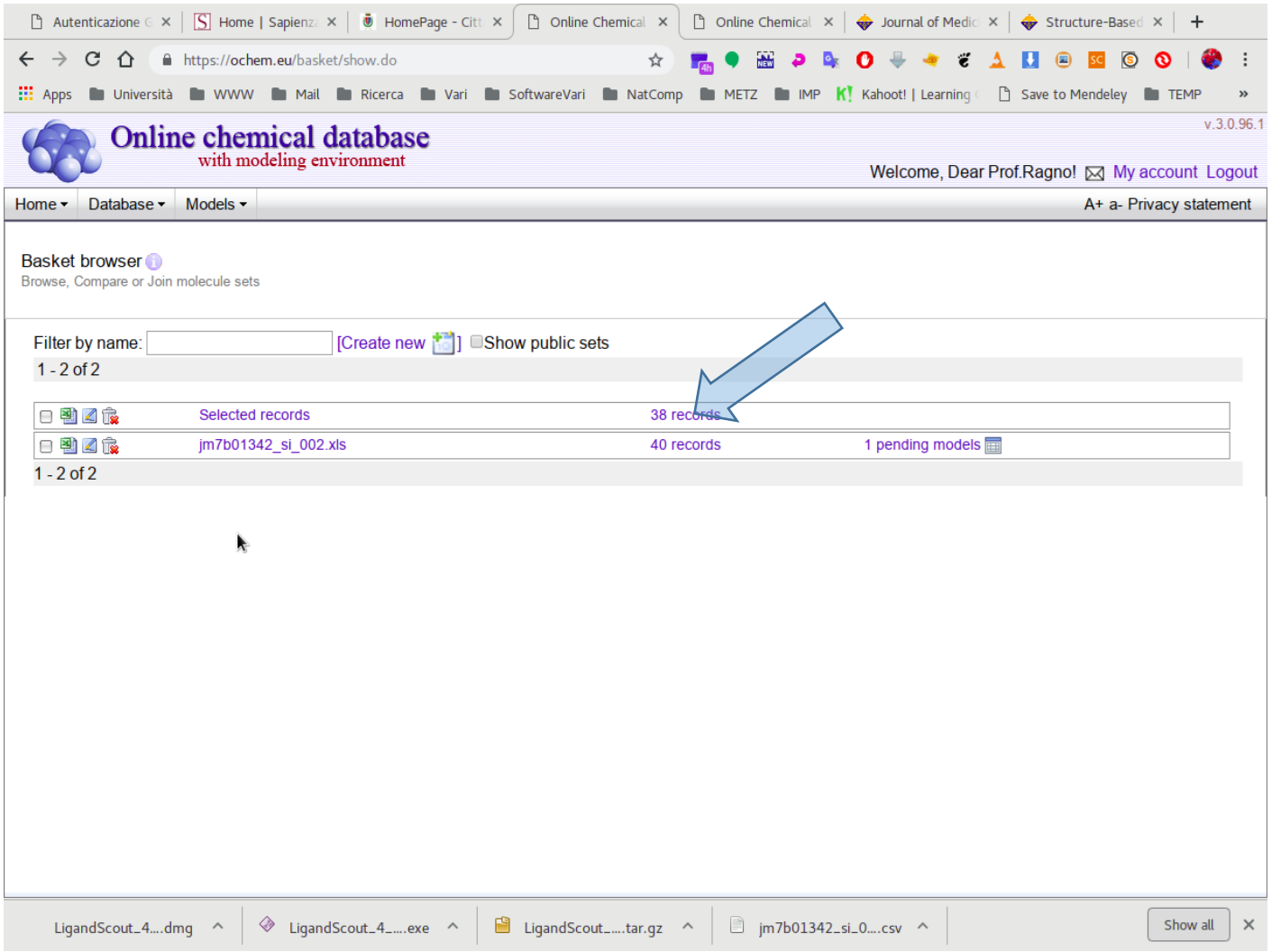
# Preparing the dataset



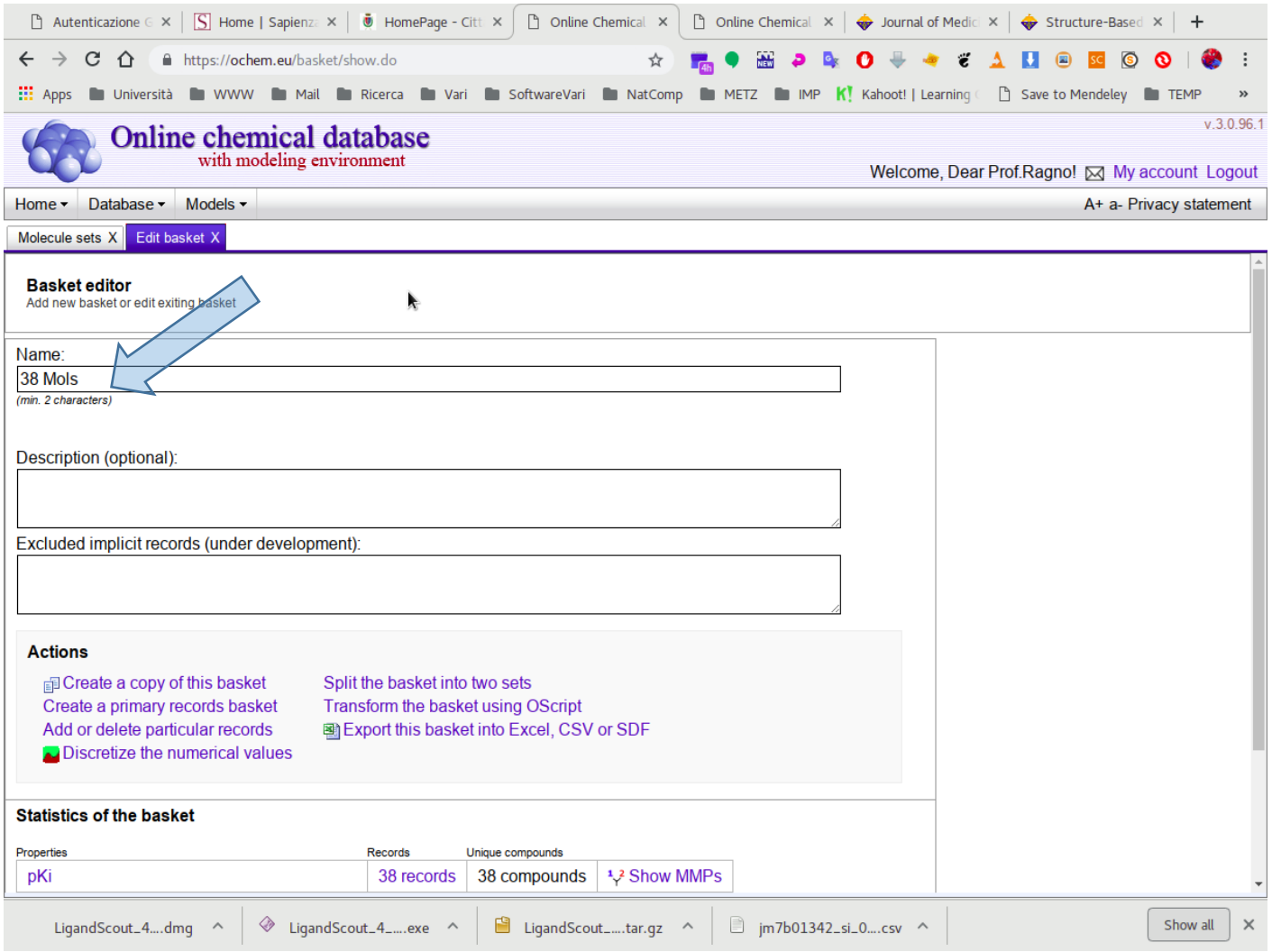
# Preparing the dataset



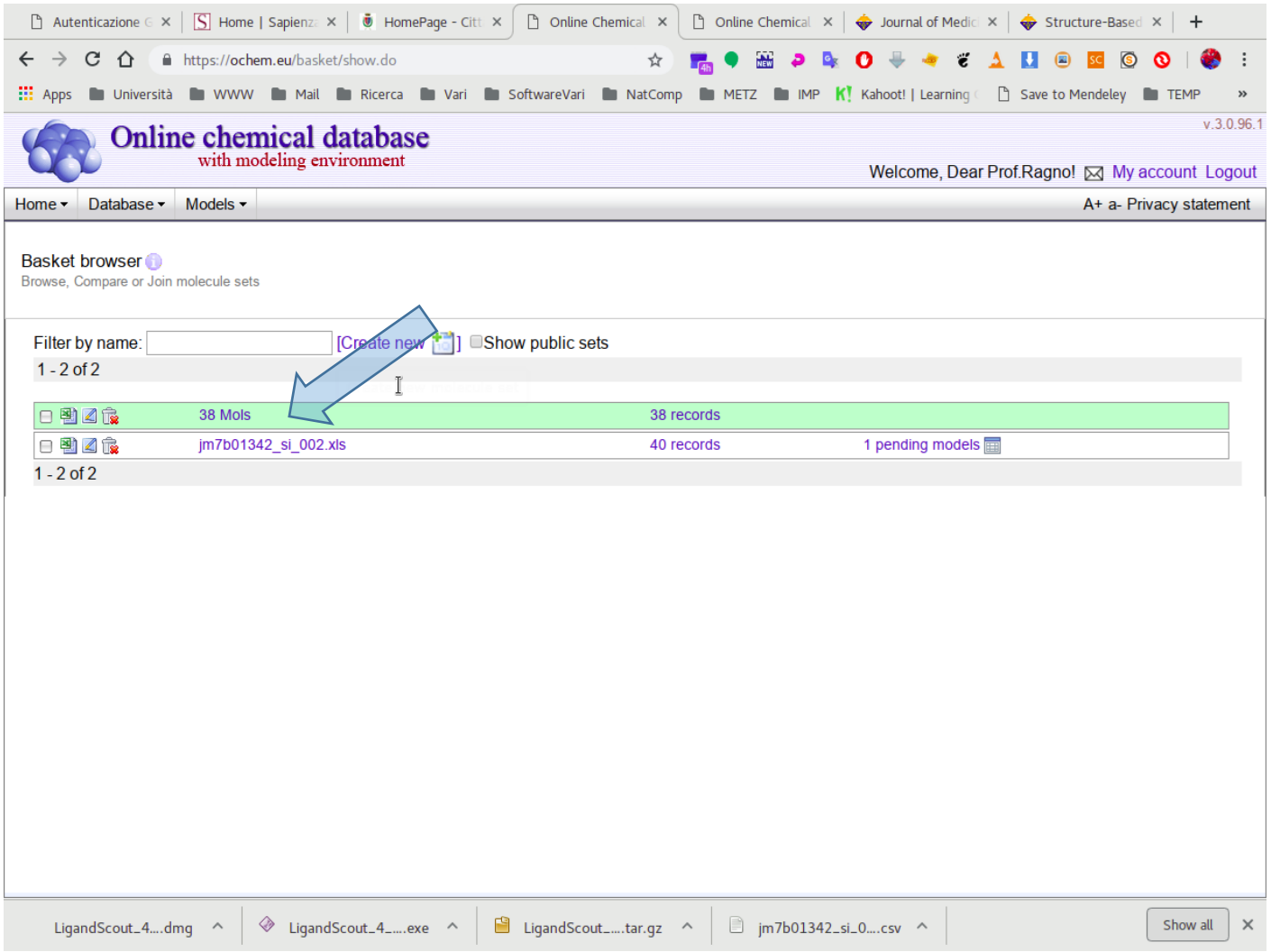
# Preparing the dataset



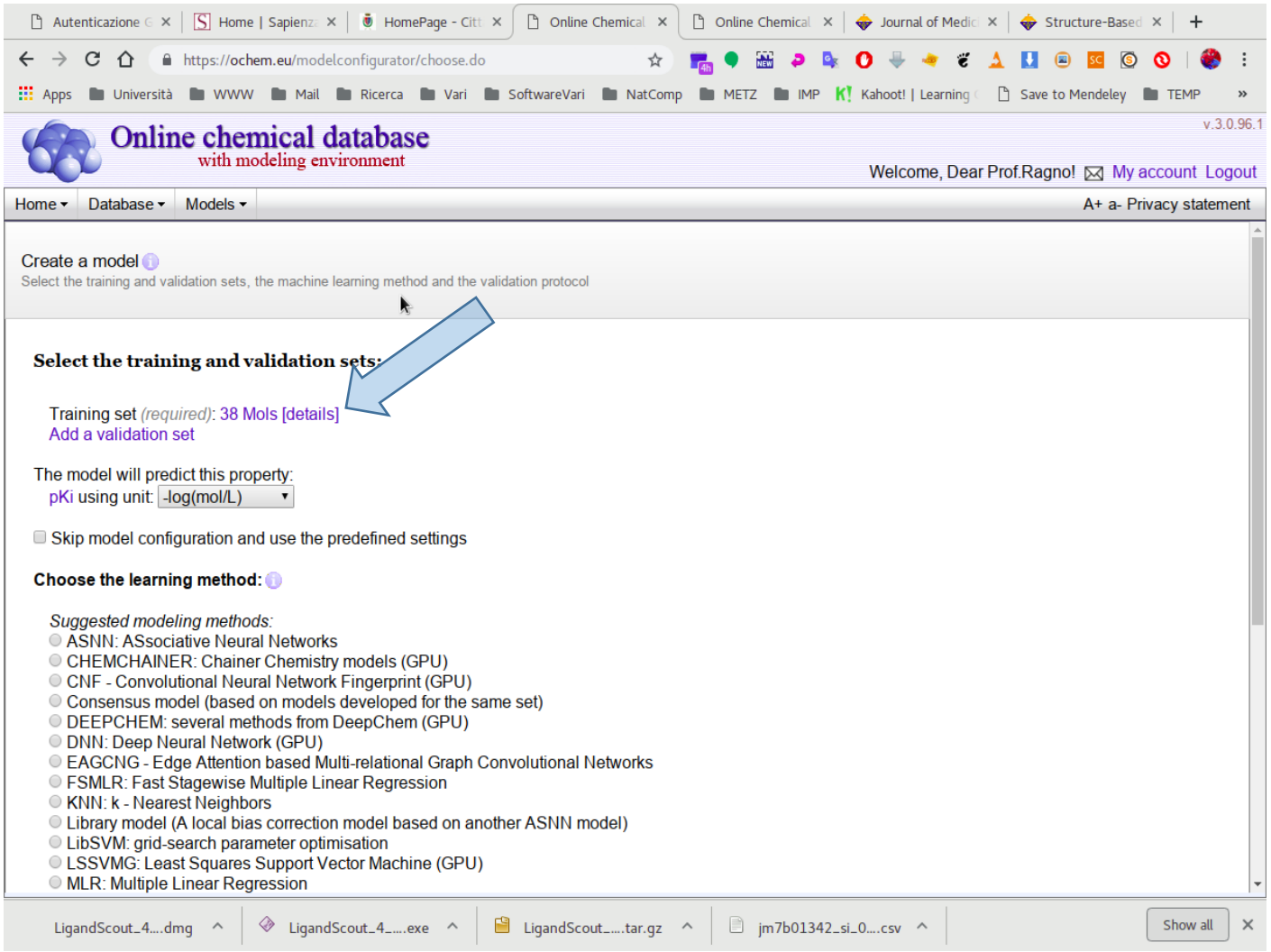
# Preparing the dataset



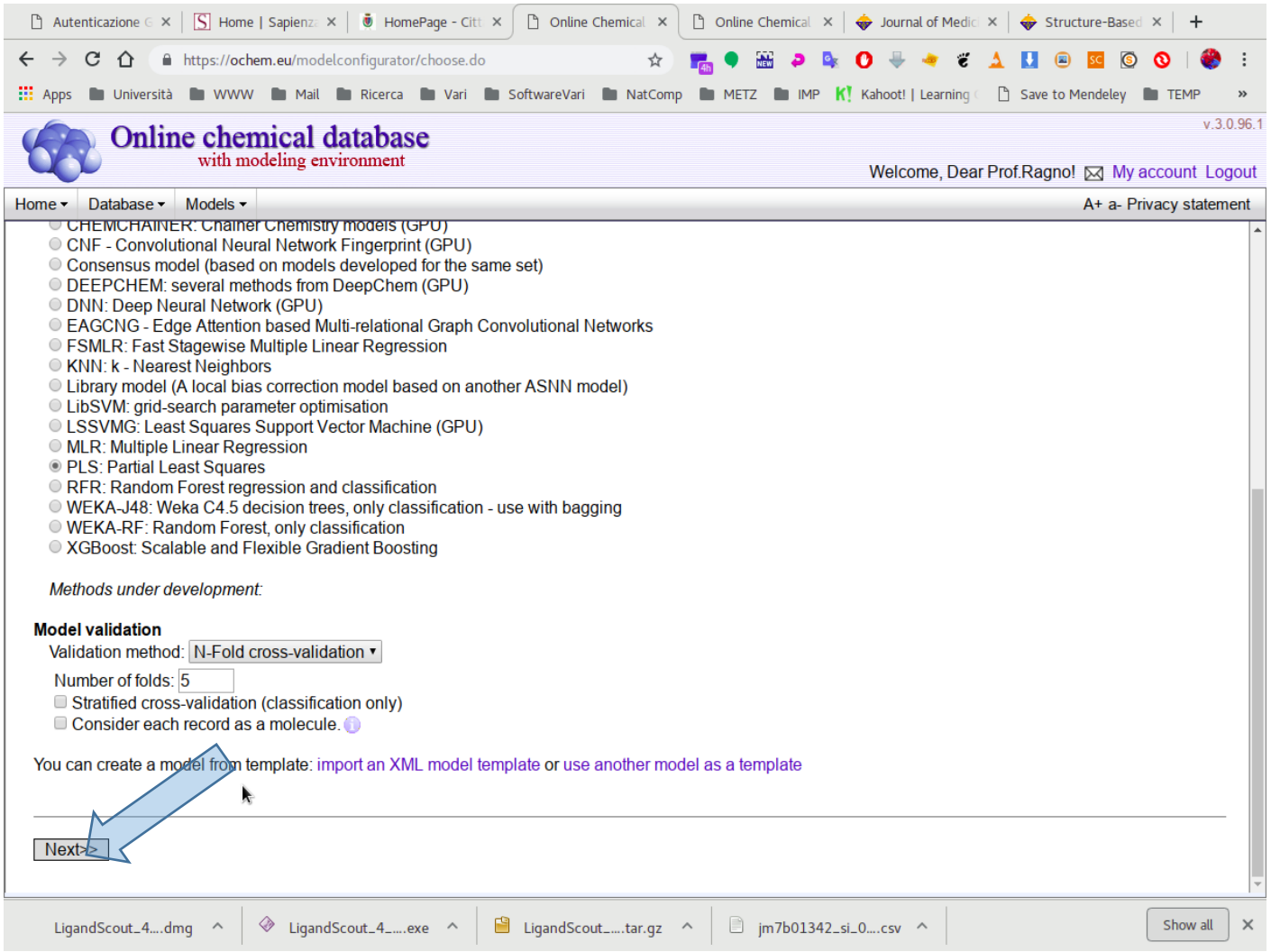
# Preparing the dataset



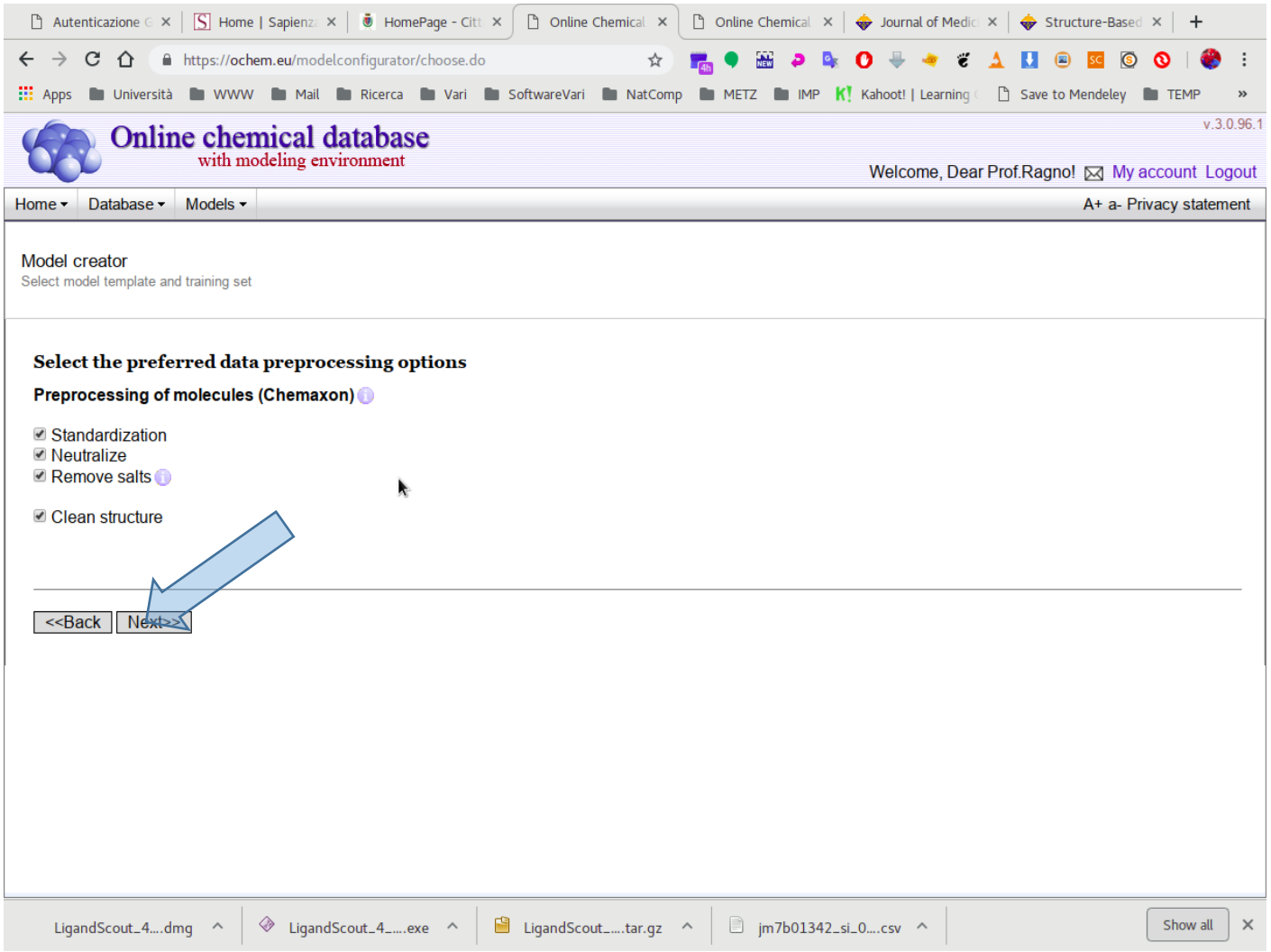
# Preparing the dataset



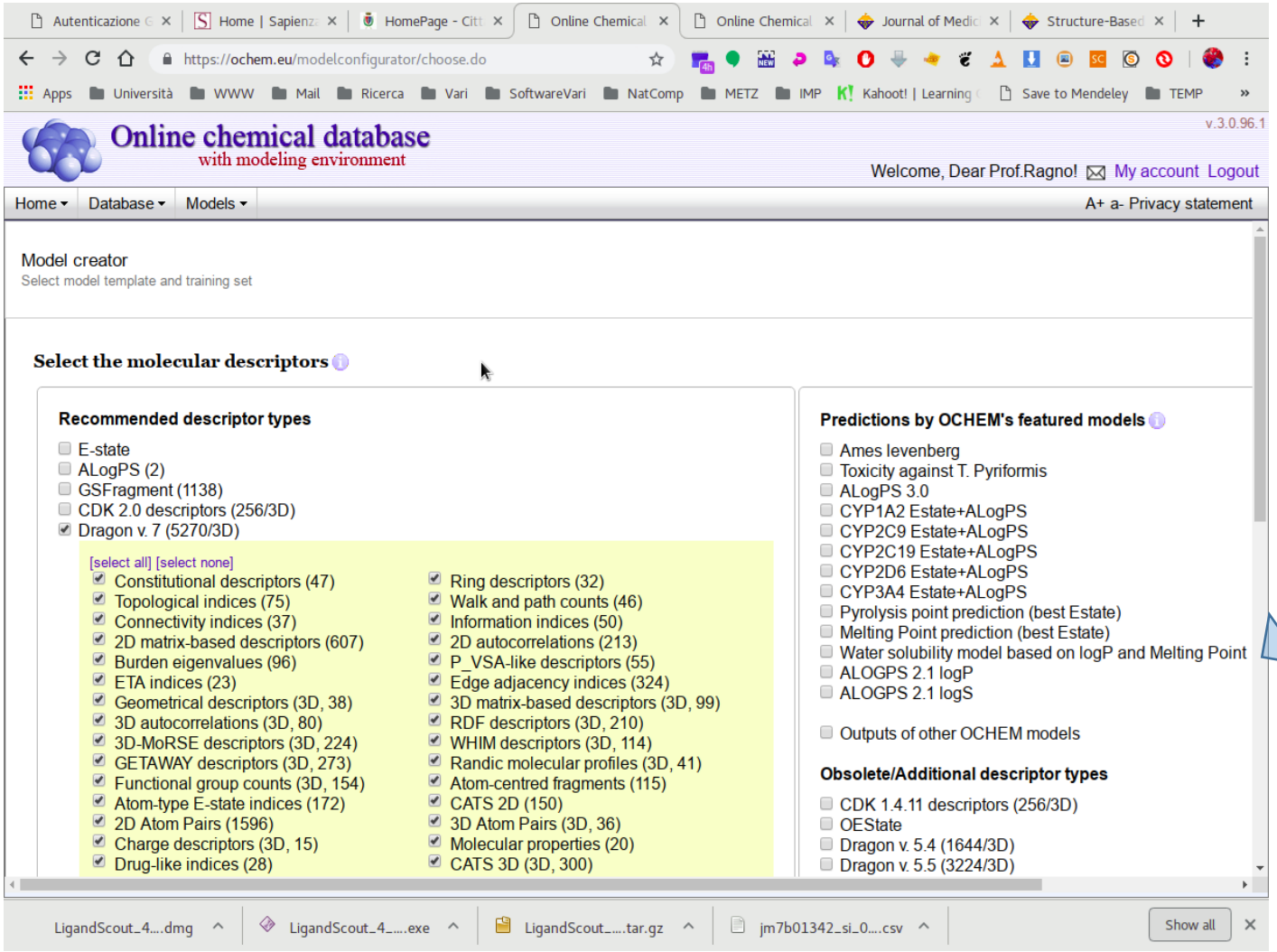
# Preparing the dataset



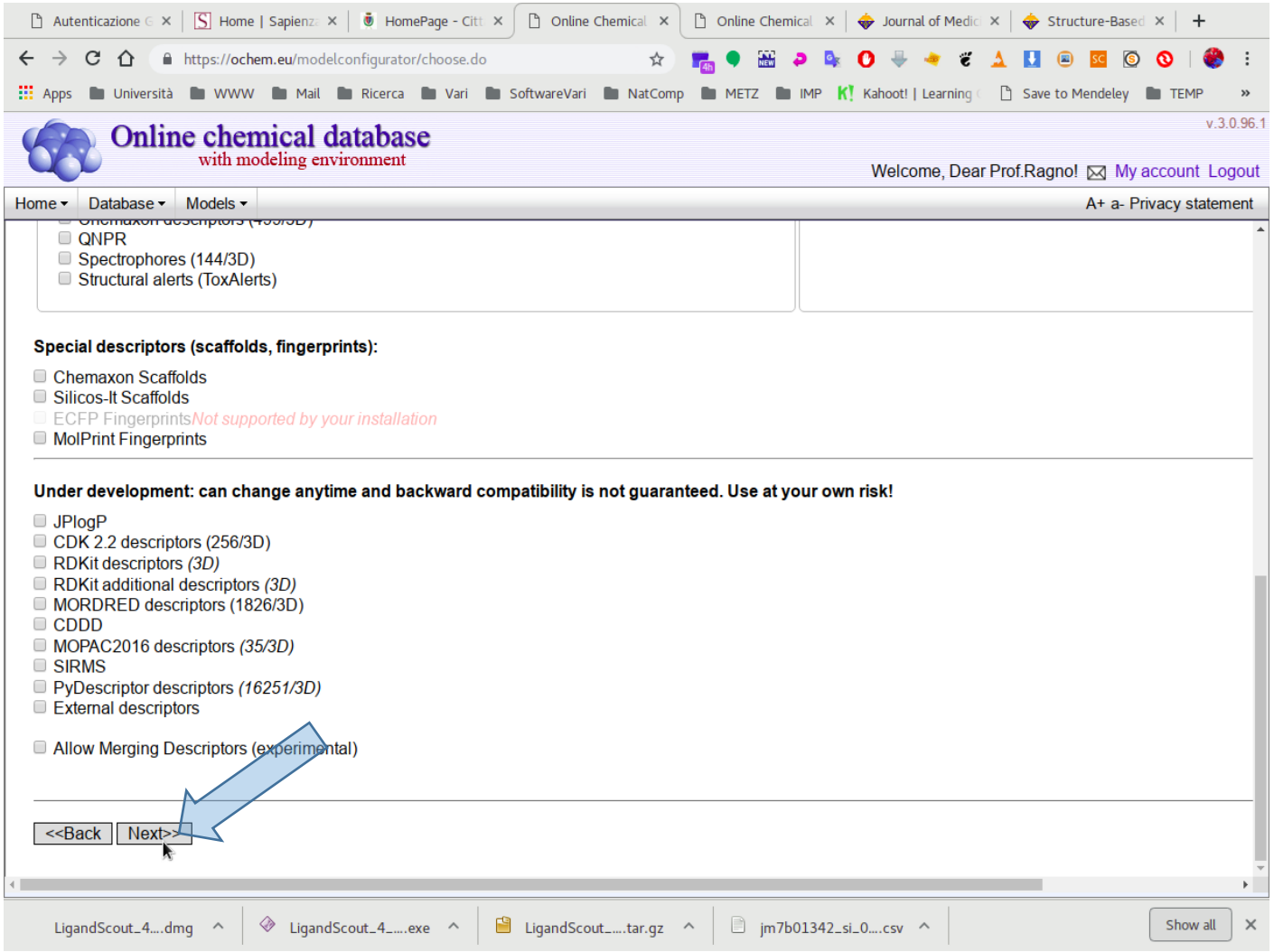
# Preparing the dataset



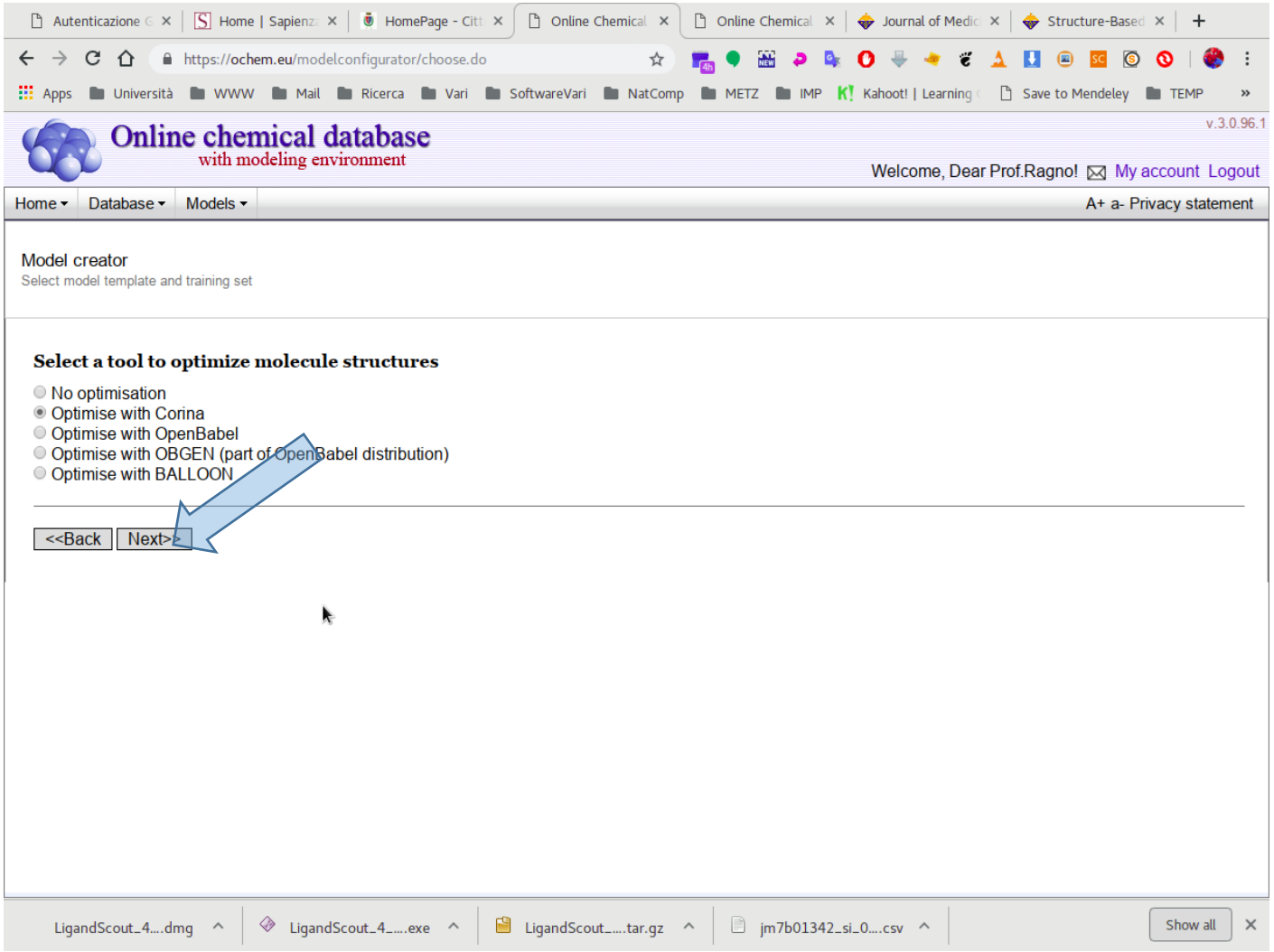
# Preparing the dataset



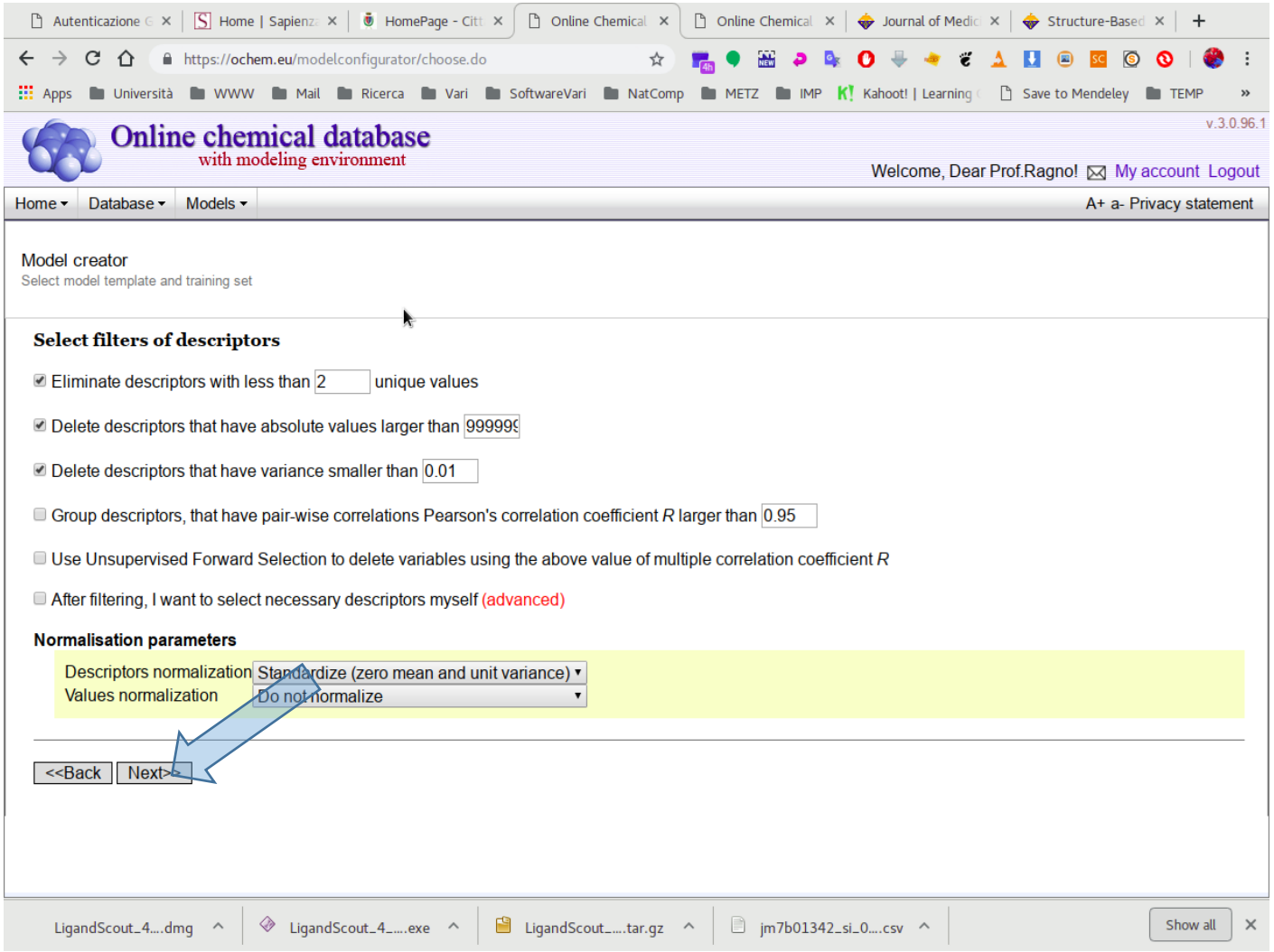
# Preparing the dataset



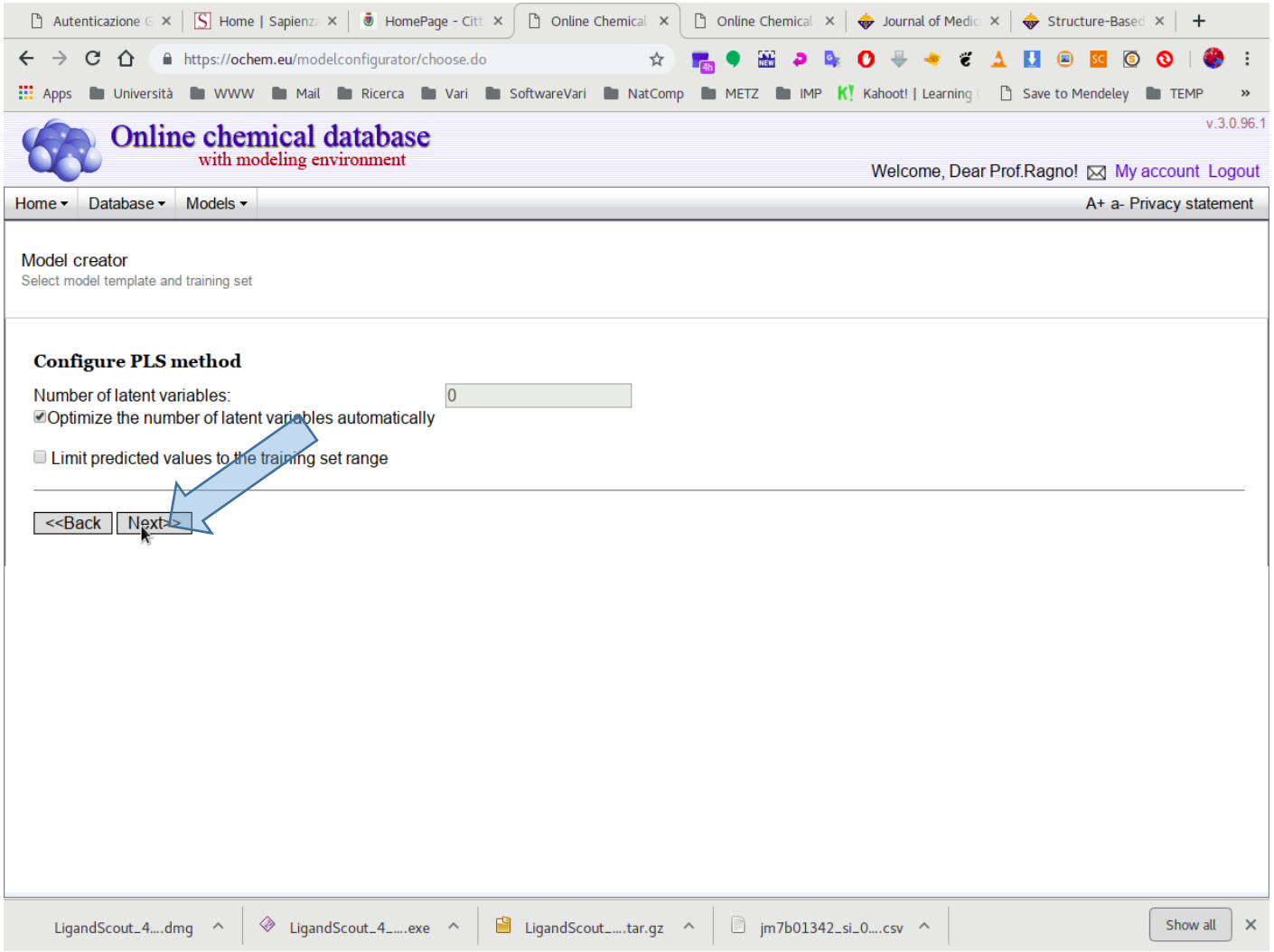
# Preparing the dataset



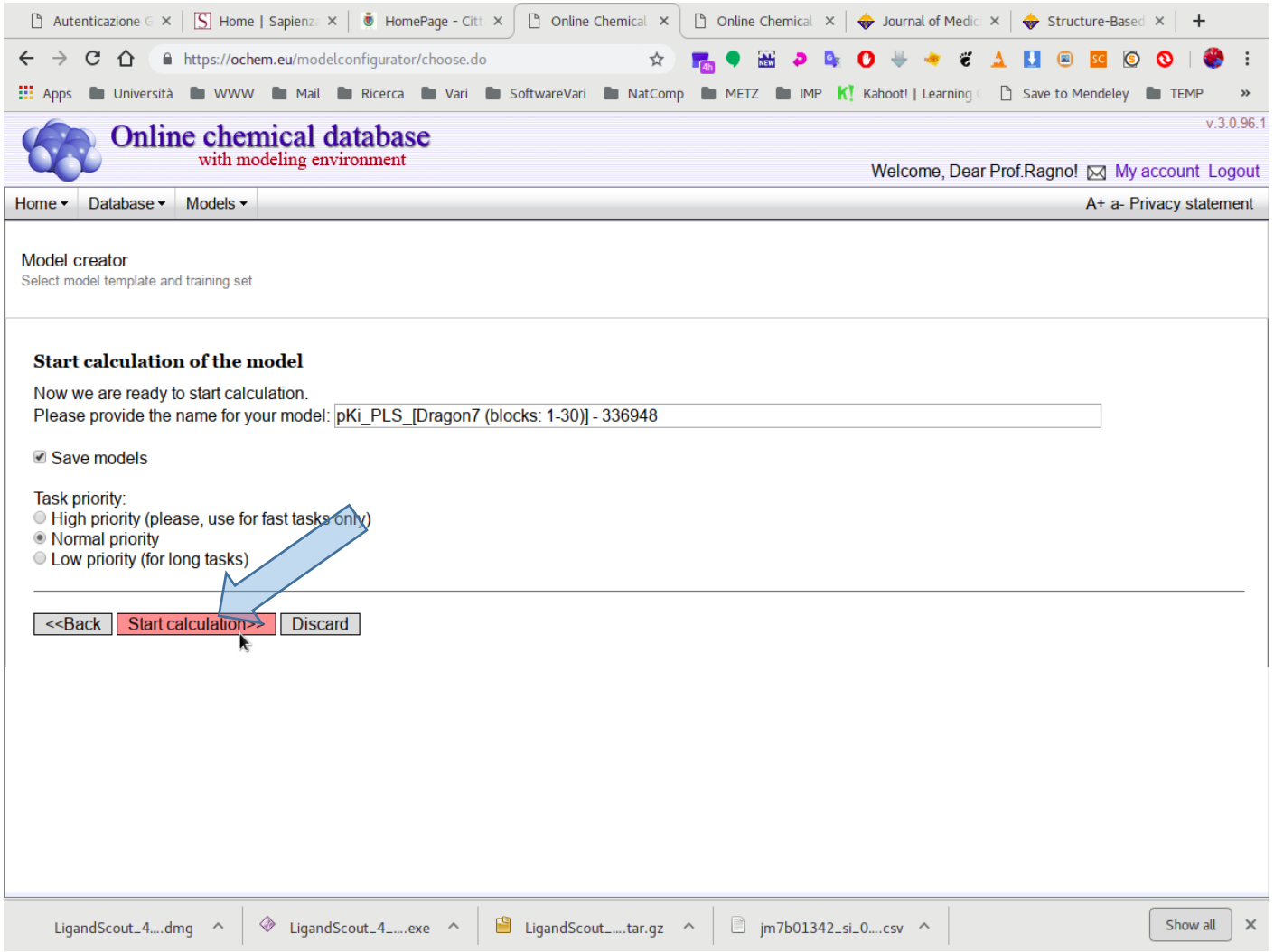
# Preparing the dataset



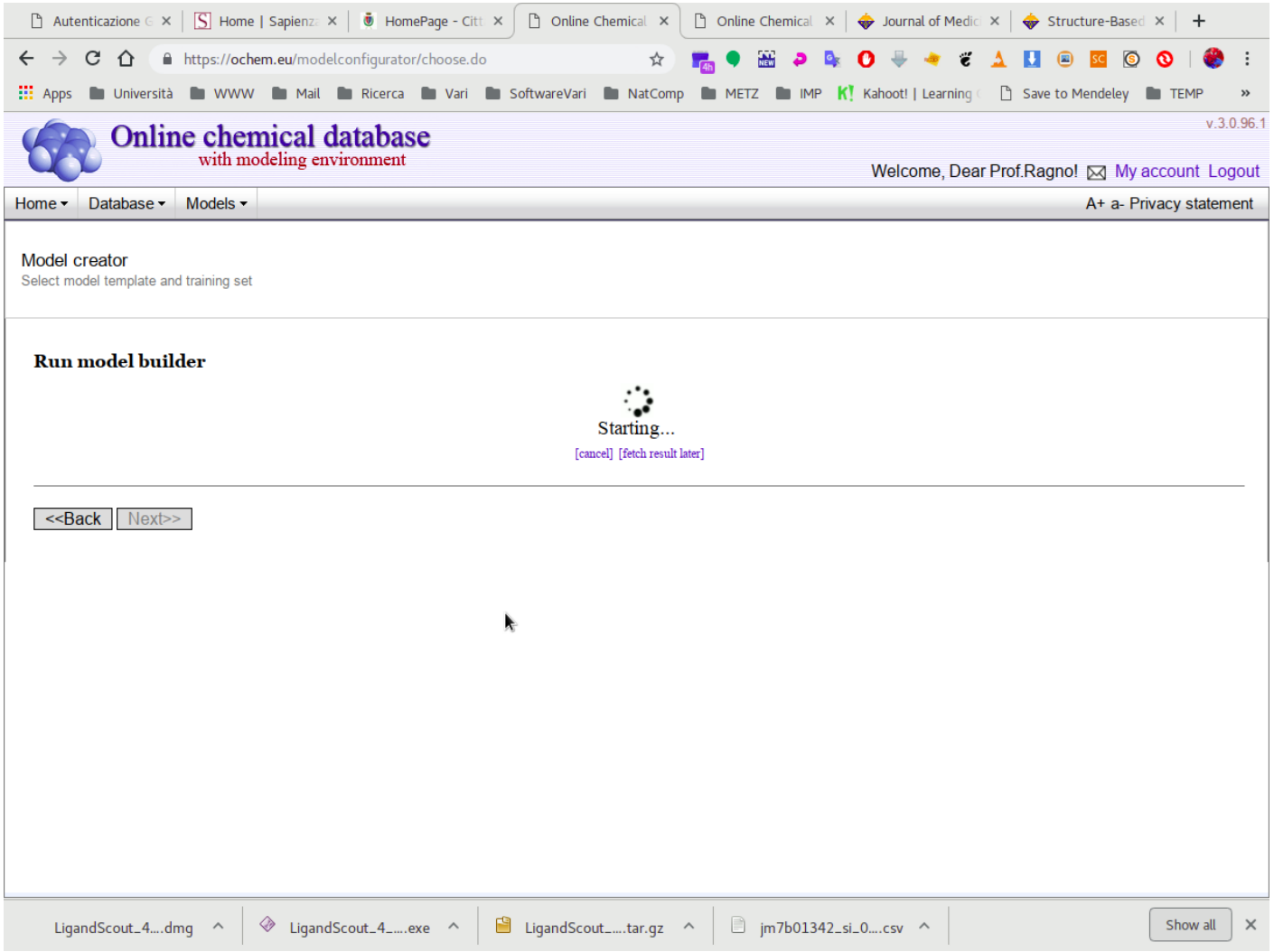
# Preparing the dataset



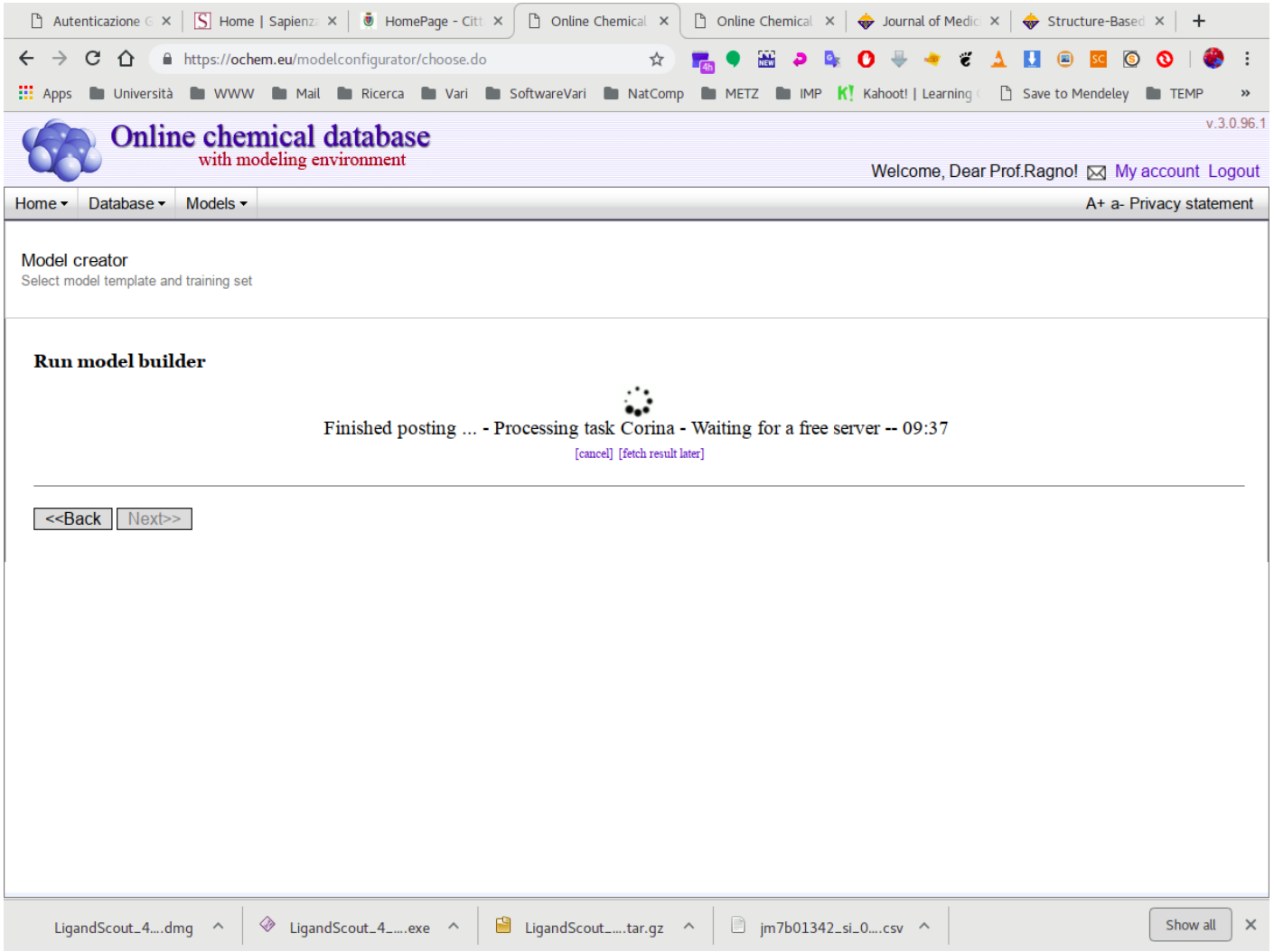
# Preparing the dataset



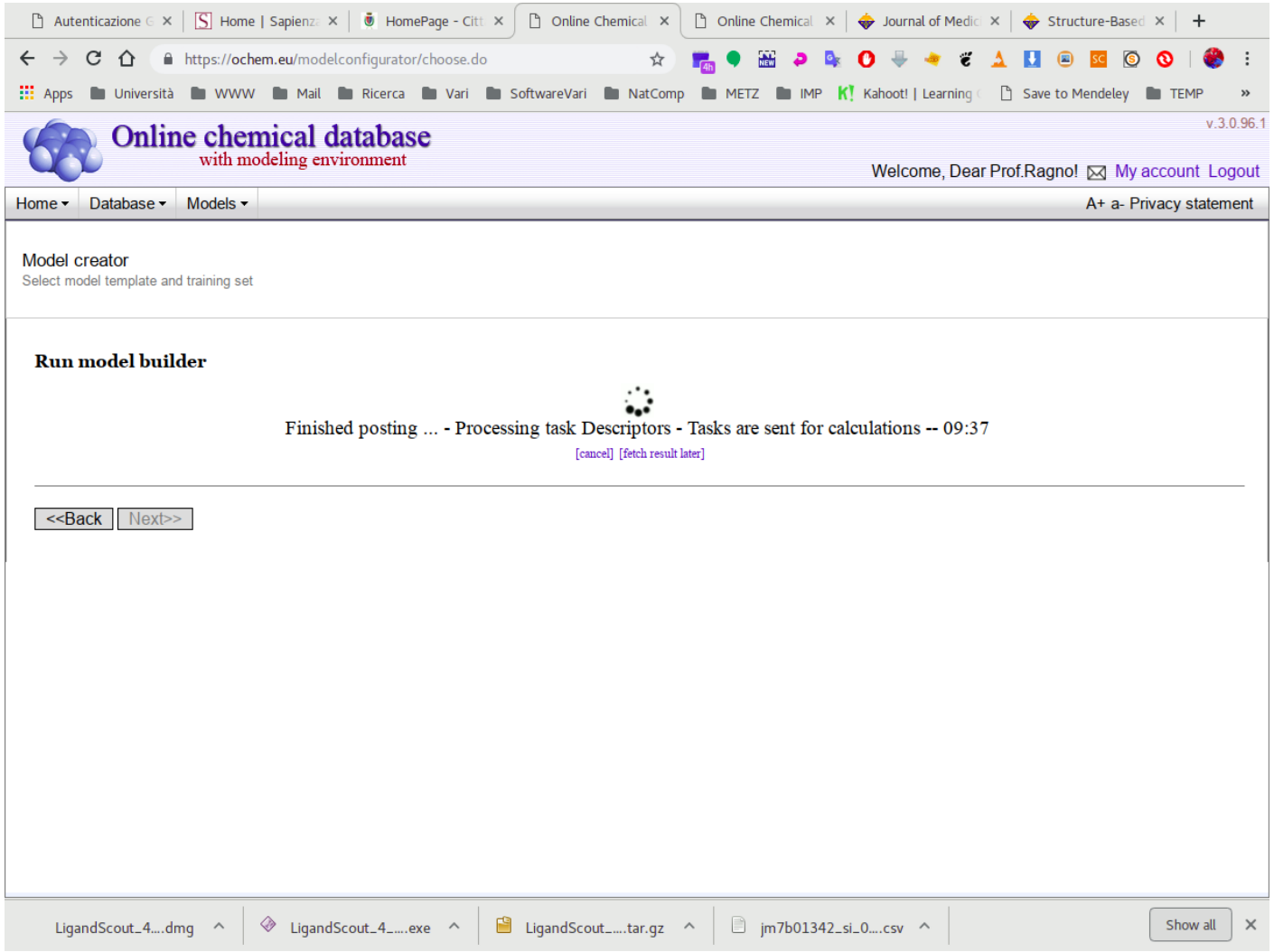
# Preparing the dataset



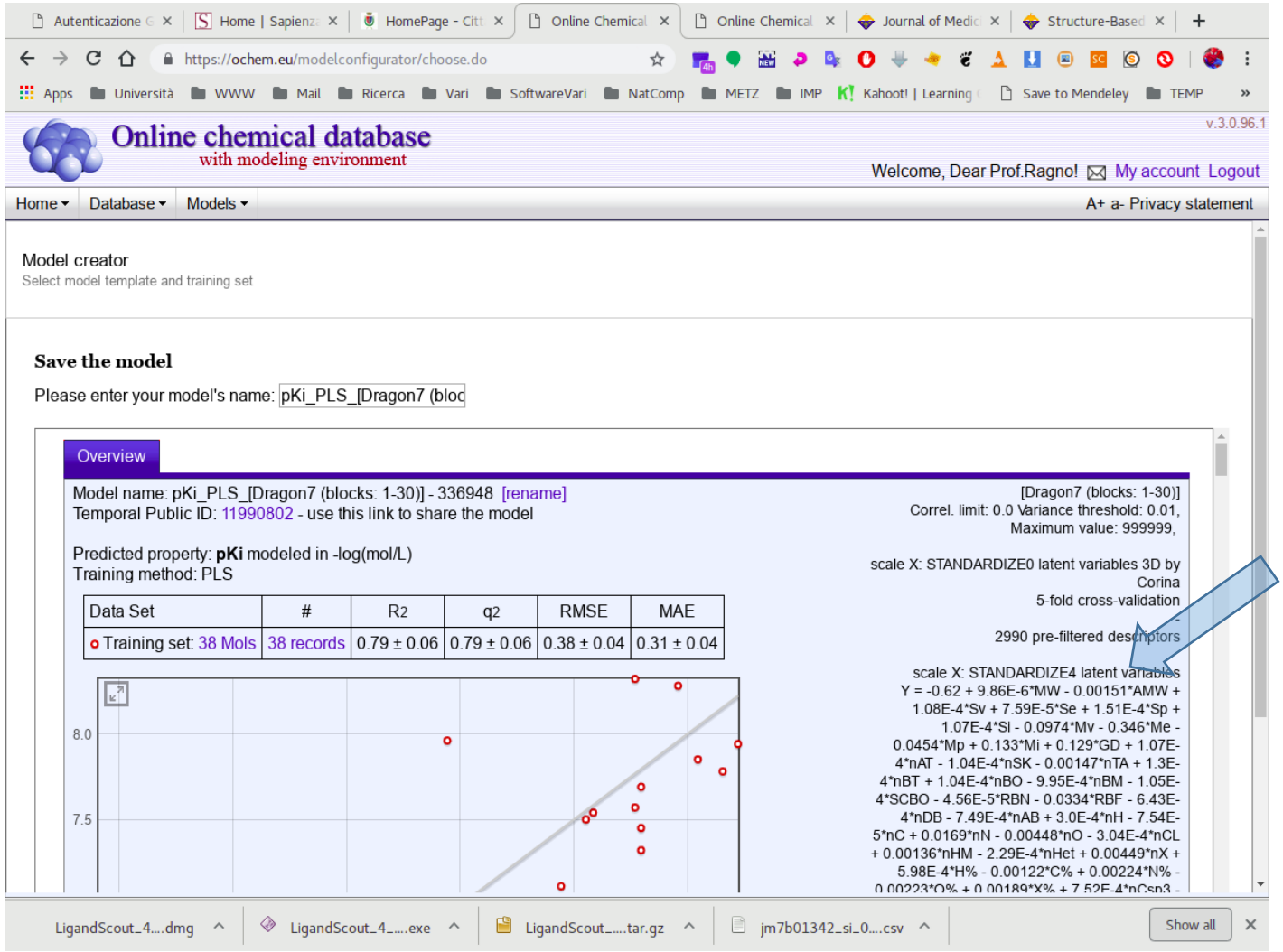
# Preparing the dataset



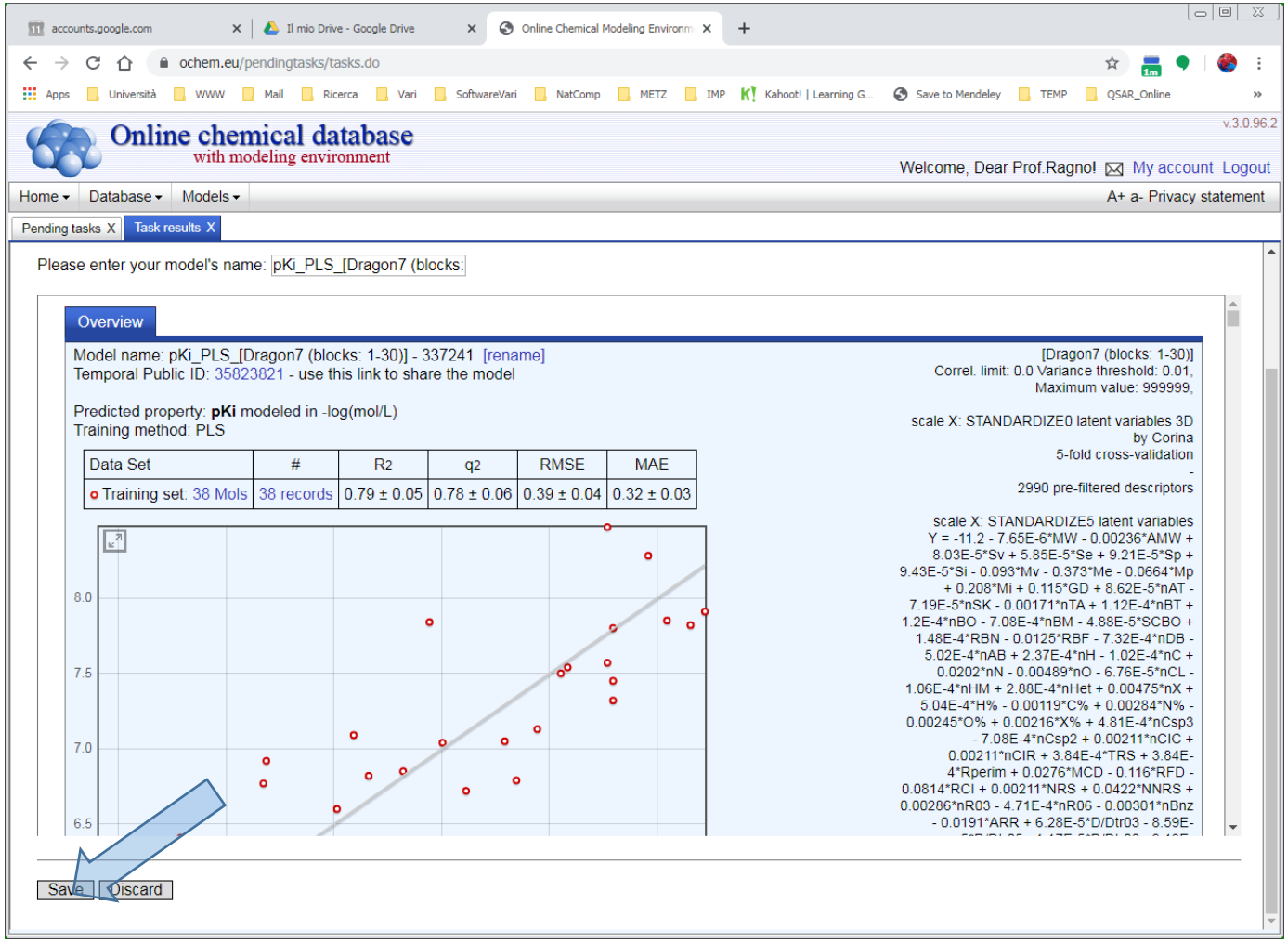
# Preparing the dataset



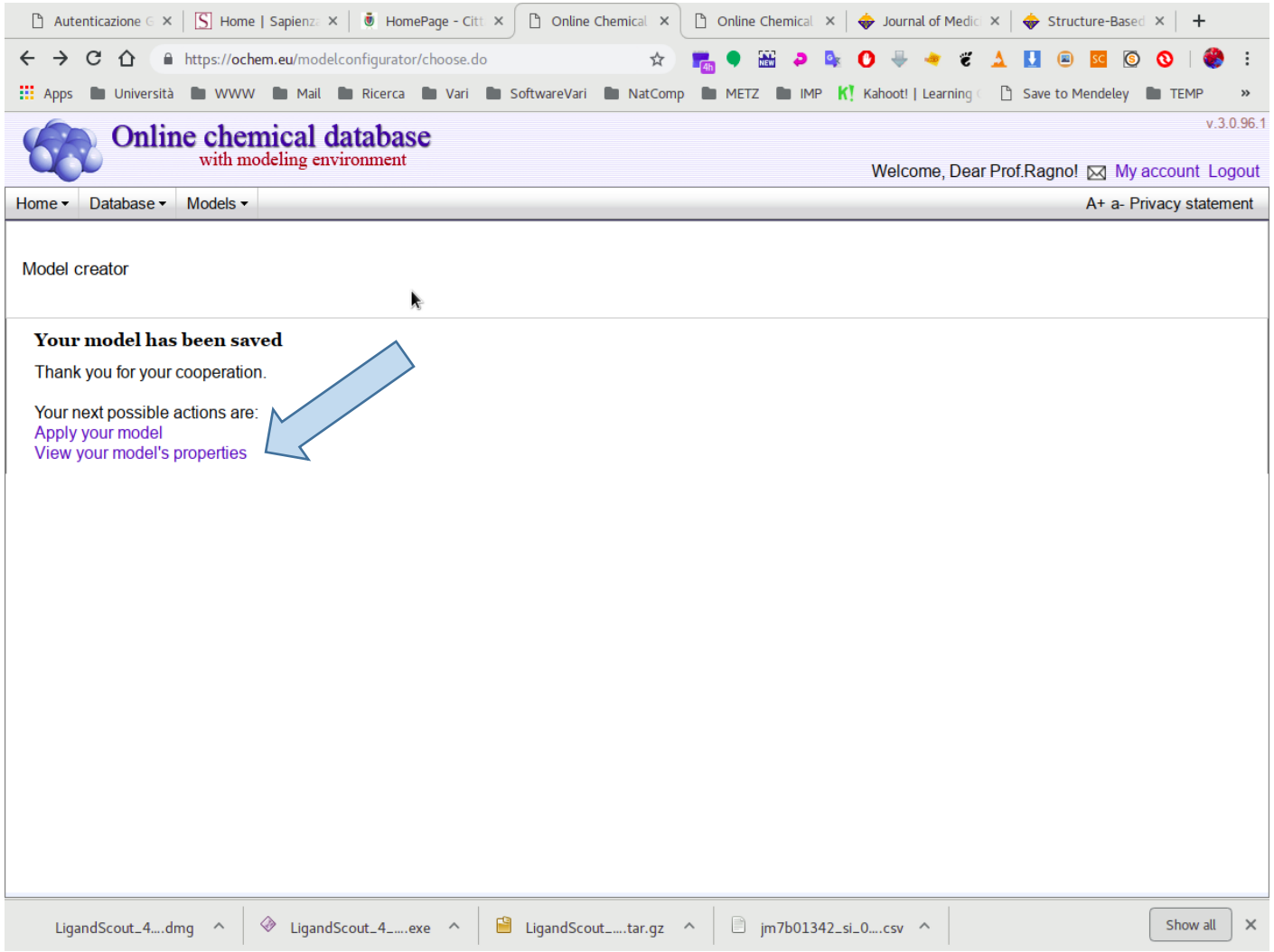
# Preparing the dataset



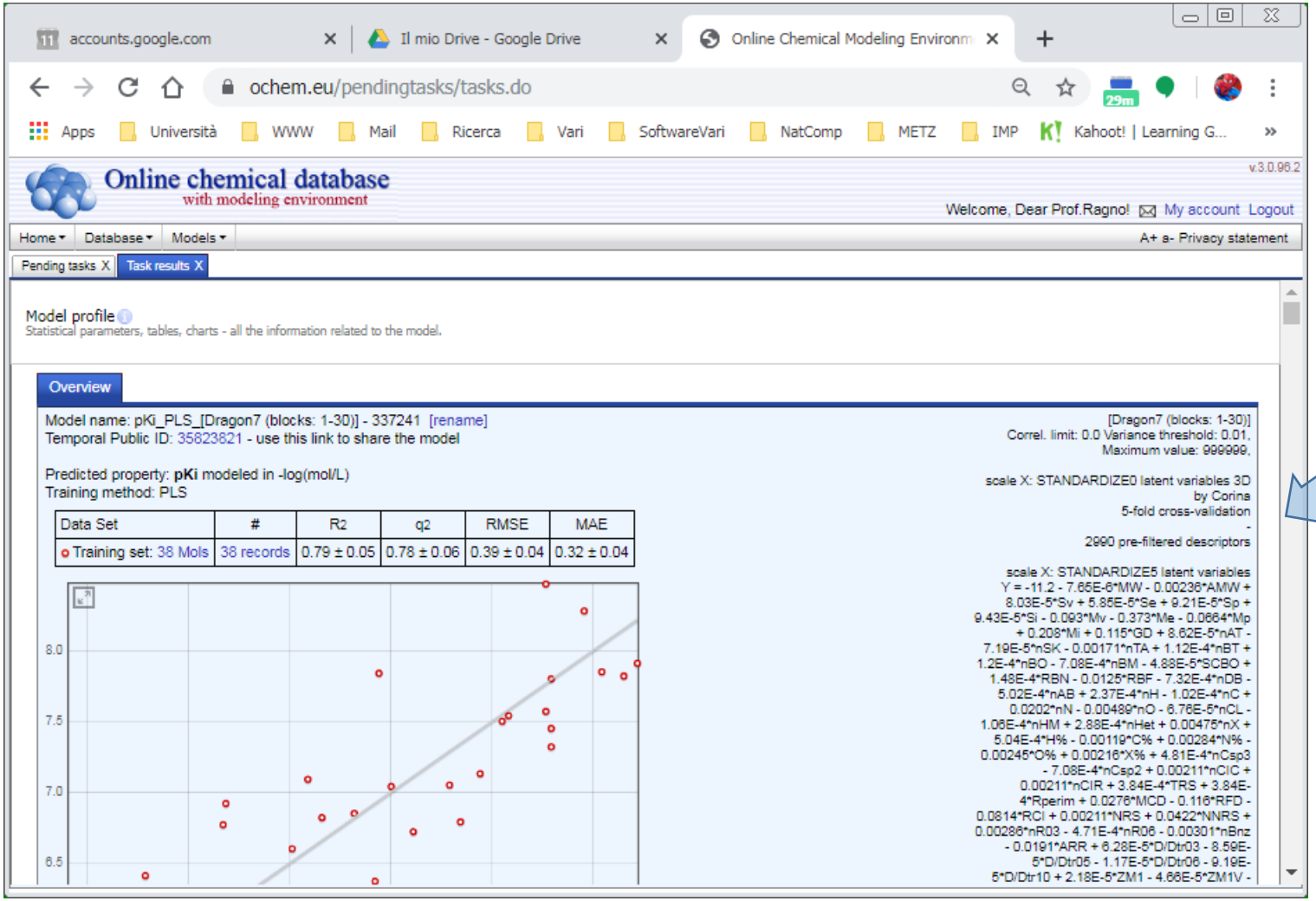
# Preparing the dataset



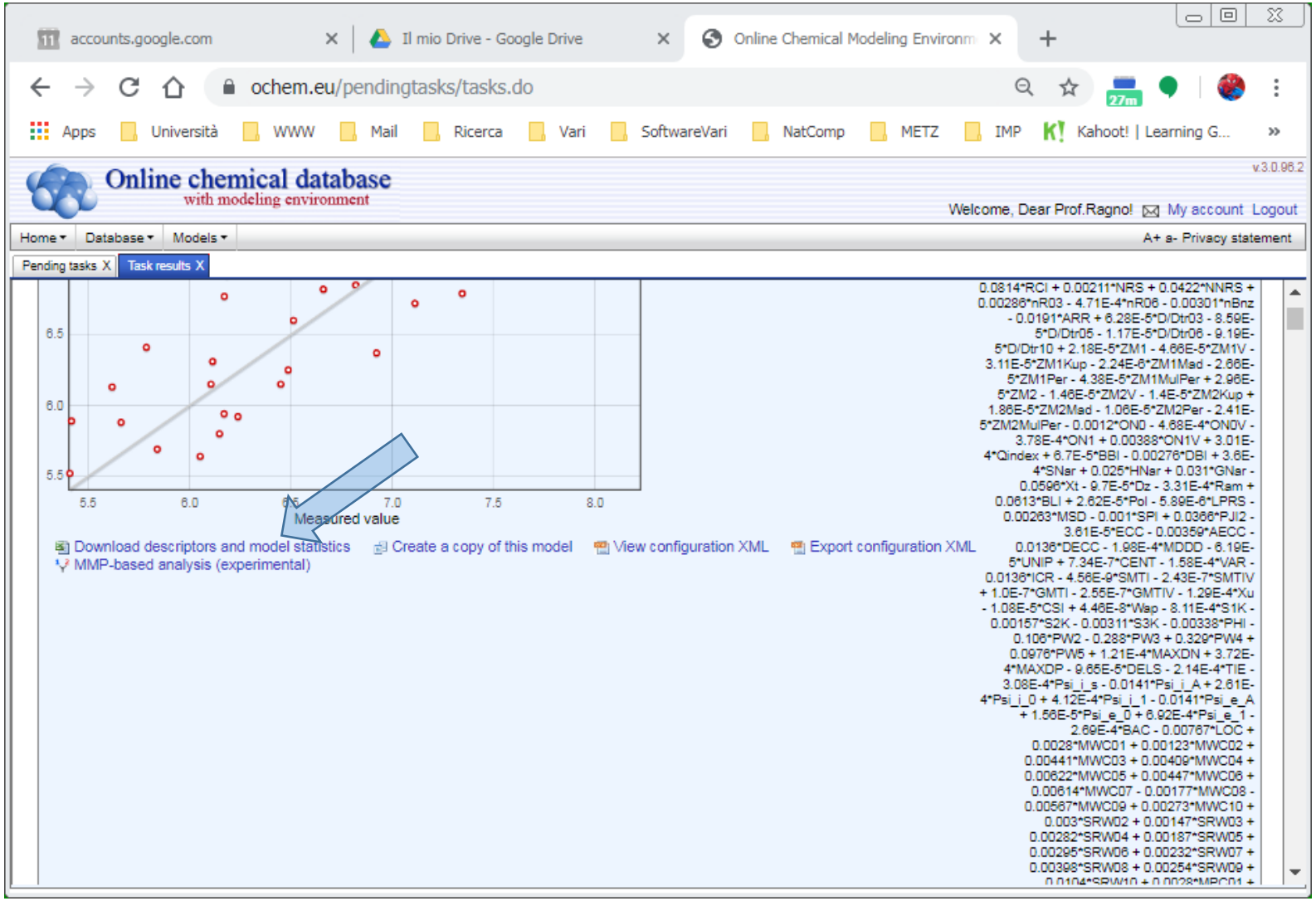
# Preparing the dataset



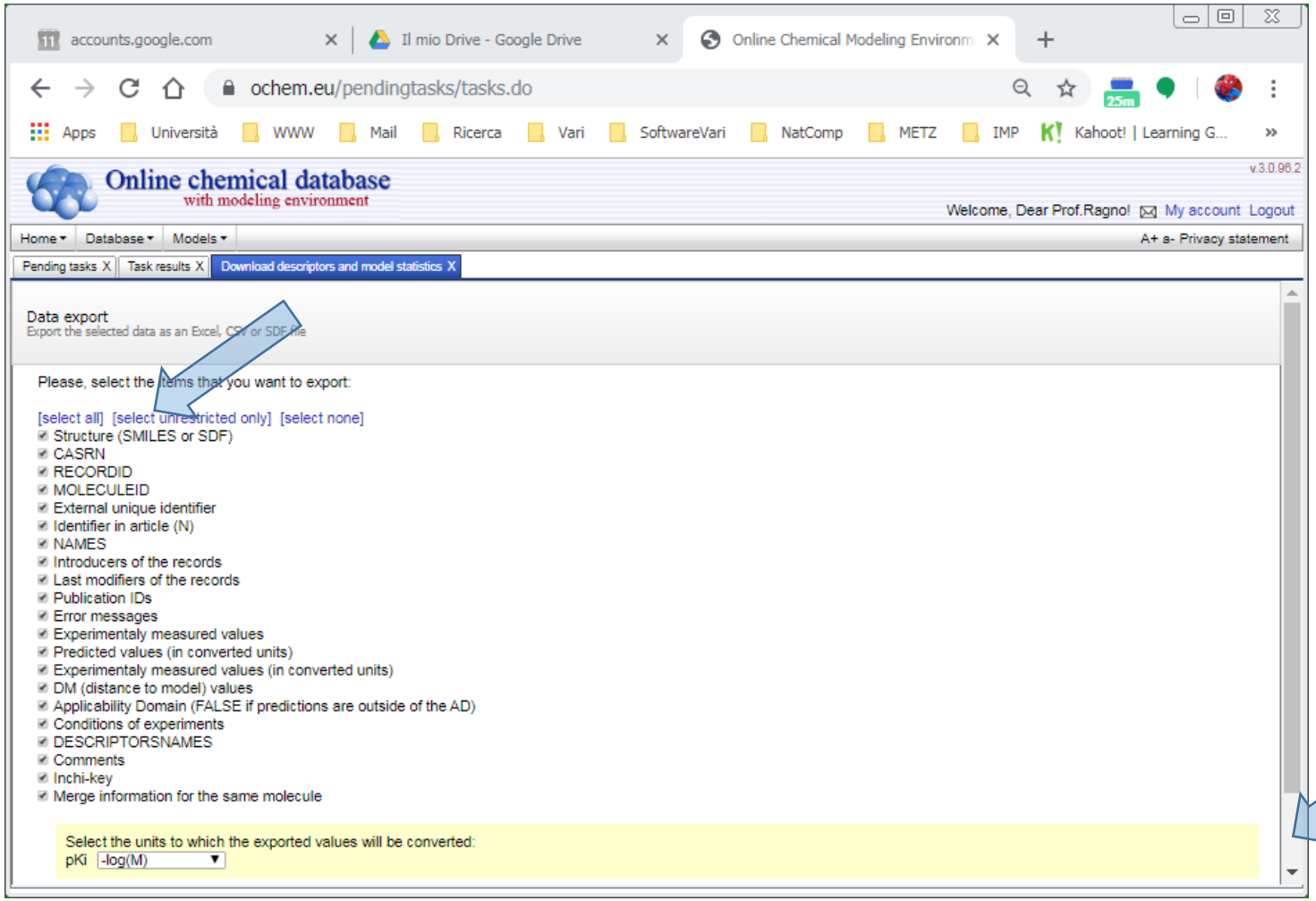
# Preparing the dataset



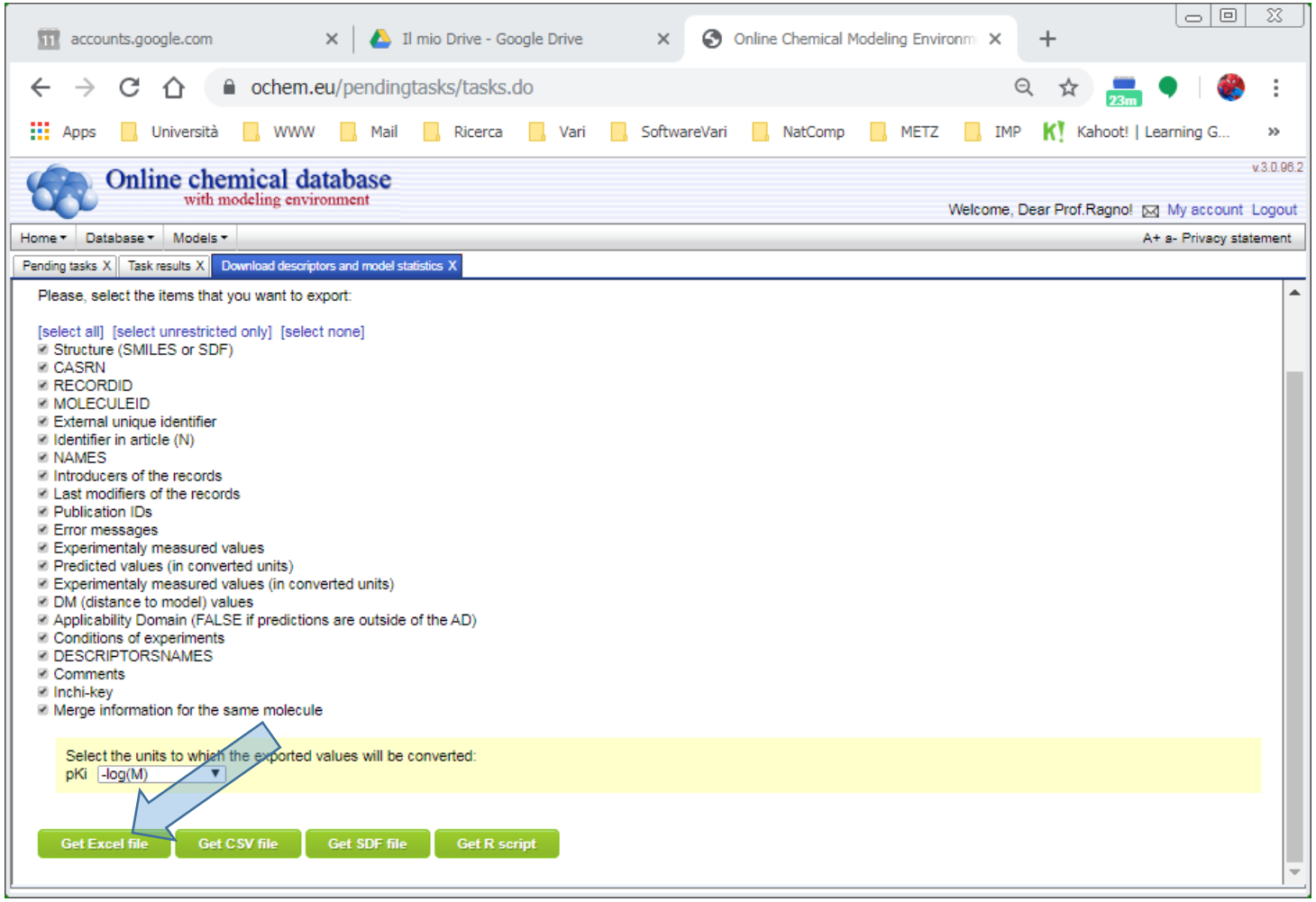
# Preparing the dataset



# Preparing the dataset



# Preparing the dataset



# Preparing the dataset