

## Supplementary Information for Predictive analytics for crystalline materials: Bulk modulus

Al'ona Furmanchuk<sup>a‡</sup>, Ankit Agrawal<sup>a\*</sup>, Alok Choudhary<sup>a</sup>

<sup>a</sup>Department of Electrical Engineering and Computer Science, Northwestern University, USA

Corresponding author: Ankit Agrawal  
E-mail: ankitag@eecs.northwestern.edu

<sup>‡</sup>Current affiliation: Center for Health Information Partnerships, Northwestern University, Feinberg School of Medicine, Chicago, IL 60611, USA

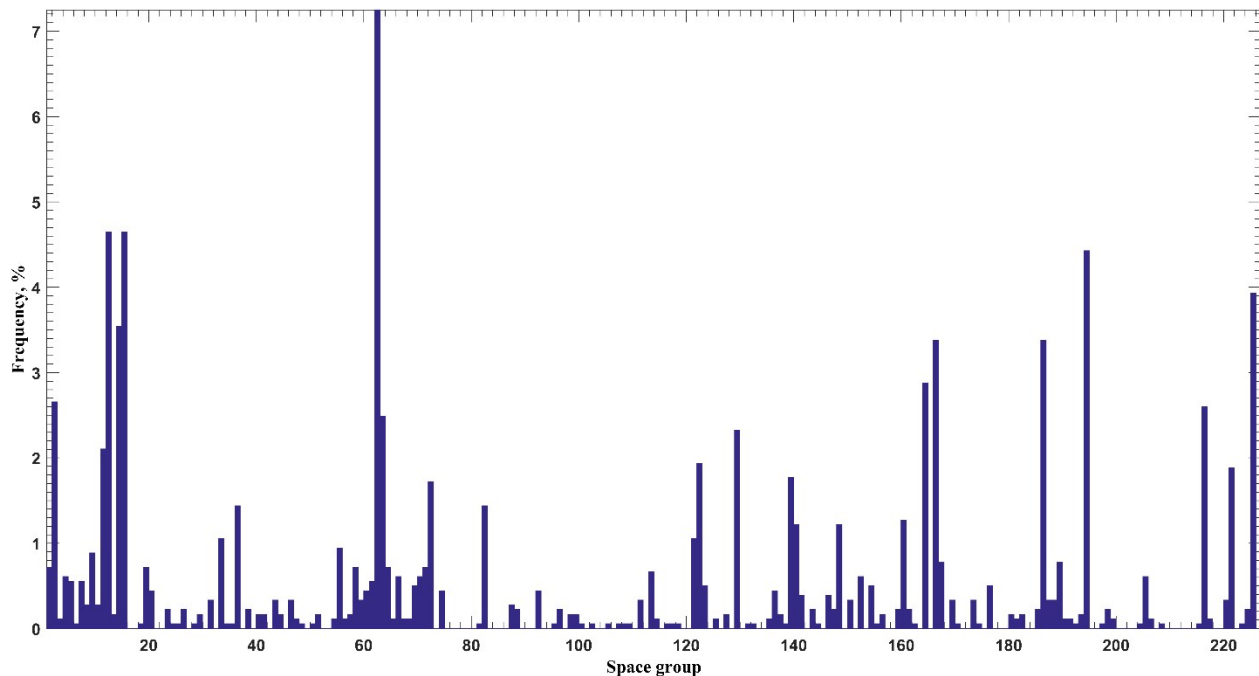
### Contents

<b>Part S1.</b> Data set characterization.	S1
<b>Part S2.</b> Attribute ranking and reduction	S3
<b>Part S3.</b> Experimental vs predicted values of Bulk modulus	S7
<b>Part S4.</b> ThermoEl tool for prediction of Bulk modulus	S16

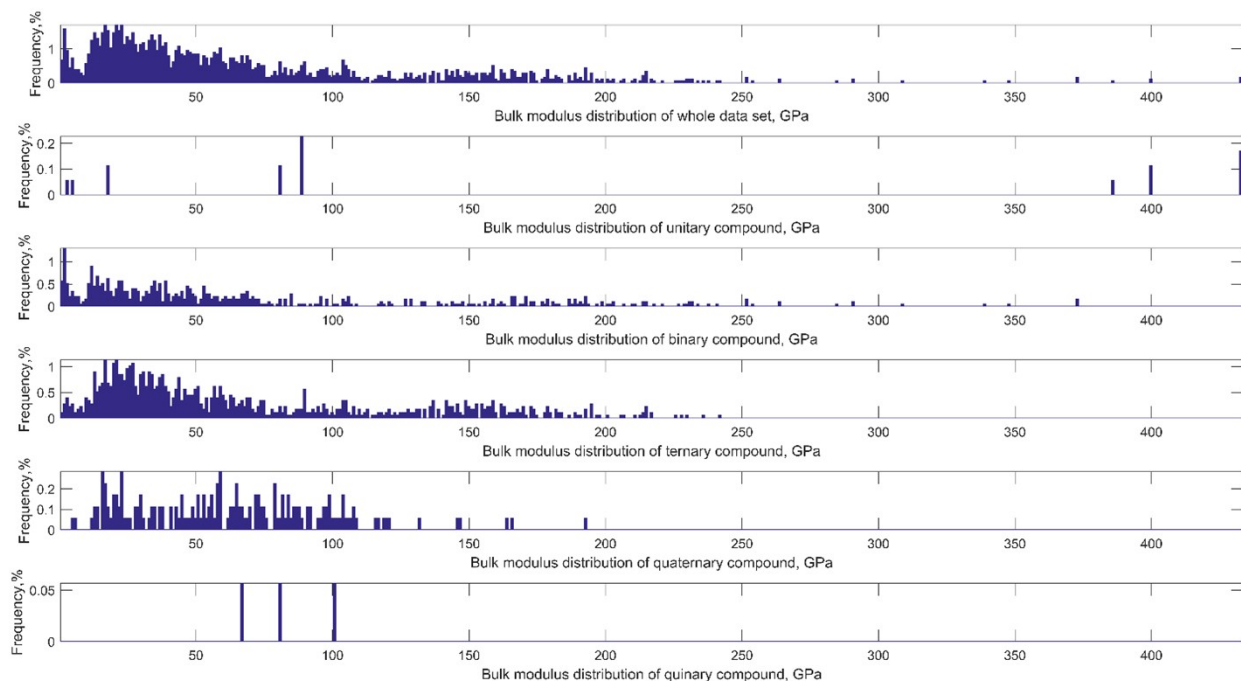
### Part S1. Data set characterization.

1 <b>H</b>																		2 <b>He</b>
3 <b>Li</b>	4 <b>Be</b>											5 <b>B</b>	6 <b>C</b>	7 <b>N</b>	8 <b>O</b>	9 <b>F</b>	10 <b>Ne</b>	
11 <b>Na</b>	12 <b>Mg</b>											13 <b>Al</b>	14 <b>Si</b>	15 <b>P</b>	16 <b>S</b>	17 <b>Cl</b>	18 <b>Ar</b>	
19 <b>K</b>	20 <b>Ca</b>		21 <b>Sc</b>	22 <b>Ti</b>	23 <b>V</b>	24 <b>Cr</b>	25 <b>Mn</b>	26 <b>Fe</b>	27 <b>Co</b>	28 <b>Ni</b>	29 <b>Cu</b>	30 <b>Zn</b>	31 <b>Ga</b>	32 <b>Ge</b>	33 <b>As</b>	34 <b>Se</b>	35 <b>Br</b>	36 <b>Kr</b>
37 <b>Rb</b>	38 <b>Sr</b>		39 <b>Y</b>	40 <b>Zr</b>	41 <b>Nb</b>	42 <b>Mo</b>	43 <b>Tc</b>	44 <b>Ru</b>	45 <b>Rh</b>	46 <b>Pd</b>	47 <b>Ag</b>	48 <b>Cd</b>	49 <b>In</b>	50 <b>Sn</b>	51 <b>Sb</b>	52 <b>Te</b>	53 <b>I</b>	54 <b>Xe</b>
55 <b>Cs</b>	56 <b>Ba</b>	57-70 *	71 <b>Lu</b>	72 <b>Hf</b>	73 <b>Ta</b>	74 <b>W</b>	75 <b>Re</b>	76 <b>Os</b>	77 <b>Ir</b>	78 <b>Pt</b>	79 <b>Au</b>	80 <b>Hg</b>	81 <b>Tl</b>	82 <b>Pb</b>	83 <b>Bi</b>	84 <b>Po</b>	85 <b>At</b>	86 <b>Rn</b>
87 <b>Fr</b>	88 <b>Ra</b>	89-102 **	103 <b>Lr</b>	104 <b>Rf</b>	105 <b>Db</b>	106 <b>Sg</b>	107 <b>Bh</b>	108 <b>Hs</b>	109 <b>Mt</b>	110 <b>Uun</b>	111 <b>Uuu</b>	112 <b>Uub</b>		114 <b>Uuq</b>				
*Lanthanide series			57 <b>La</b>	58 <b>Ce</b>	59 <b>Pr</b>	60 <b>Nd</b>	61 <b>Pm</b>	62 <b>Sm</b>	63 <b>Eu</b>	64 <b>Gd</b>	65 <b>Tb</b>	66 <b>Dy</b>	67 <b>Ho</b>	68 <b>Er</b>	69 <b>Tm</b>	70 <b>Yb</b>		
**Actinide series			89 <b>Ac</b>	90 <b>Th</b>	91 <b>Pa</b>	92 <b>U</b>	93 <b>Np</b>	94 <b>Pu</b>	95 <b>Am</b>	96 <b>Cm</b>	97 <b>Bk</b>	98 <b>Cf</b>	99 <b>Es</b>	100 <b>Fm</b>	101 <b>Md</b>	102 <b>No</b>		

**Figure S1.** Elemental content of TE design database. Elements present in TE design database are marked in grey.



**Figure S2.** Population of TE design database by space groups.



**Figure S3.** Distribution of Bulk modulus values in data set cleaned from duplicates.

The cutoff 250GPa was applied to curated data set in order to avoid undersampling problem in the model. Therefore, 21 samples (Table S1) whose Bulk modulus values were  $>250$  GPa were excluded from the data set. After this step, one sample for  $B_1N_1$ , 50 samples for  $O_2Si_1$ , six samples for  $Ni_4Si_3$ , and four samples for  $Al_1N_1$  were still present in the data set. No sample for  $C_1$ , and  $Ru_1C_1$  were left.

**Tables S1.** List of ICSD codes for samples excluded from data set due to their high values of Bulk modulus.

<b>ICSD Numbers for samples with Bulk modulus &gt; 250 GPa</b>					
<b>C1</b>	<b>BN1</b>	<b>Ru1C1</b>	<b>O2Si1</b>	<b>N4Si3</b>	<b>AlN1</b>
182761	162876	183170	170531	97567	41358
66464	183257		44271	156337	
182760	614864		40105		
27422	162875		181310		
28859	182731		51702		
88819					
67787					

**Table S2.** Statistics of TE design database, and summary on splitting of cleaned data.

<b>Compound type</b>	<b>Set after cleaning from duplicates</b>	<b>Set after cleaning from outliers and others</b>	<b>Withheld subset for further tests</b>	<b>Working set</b>
Unary	17	10	2	8
Binary	600	588	118	470
Ternary	1006	1002	196	806
Quaternary	179	181	39	142
Quinary	3	3	1	2
<b>Total</b>	<b>1805</b>	<b>1784</b>	<b>356</b>	<b>1428</b>

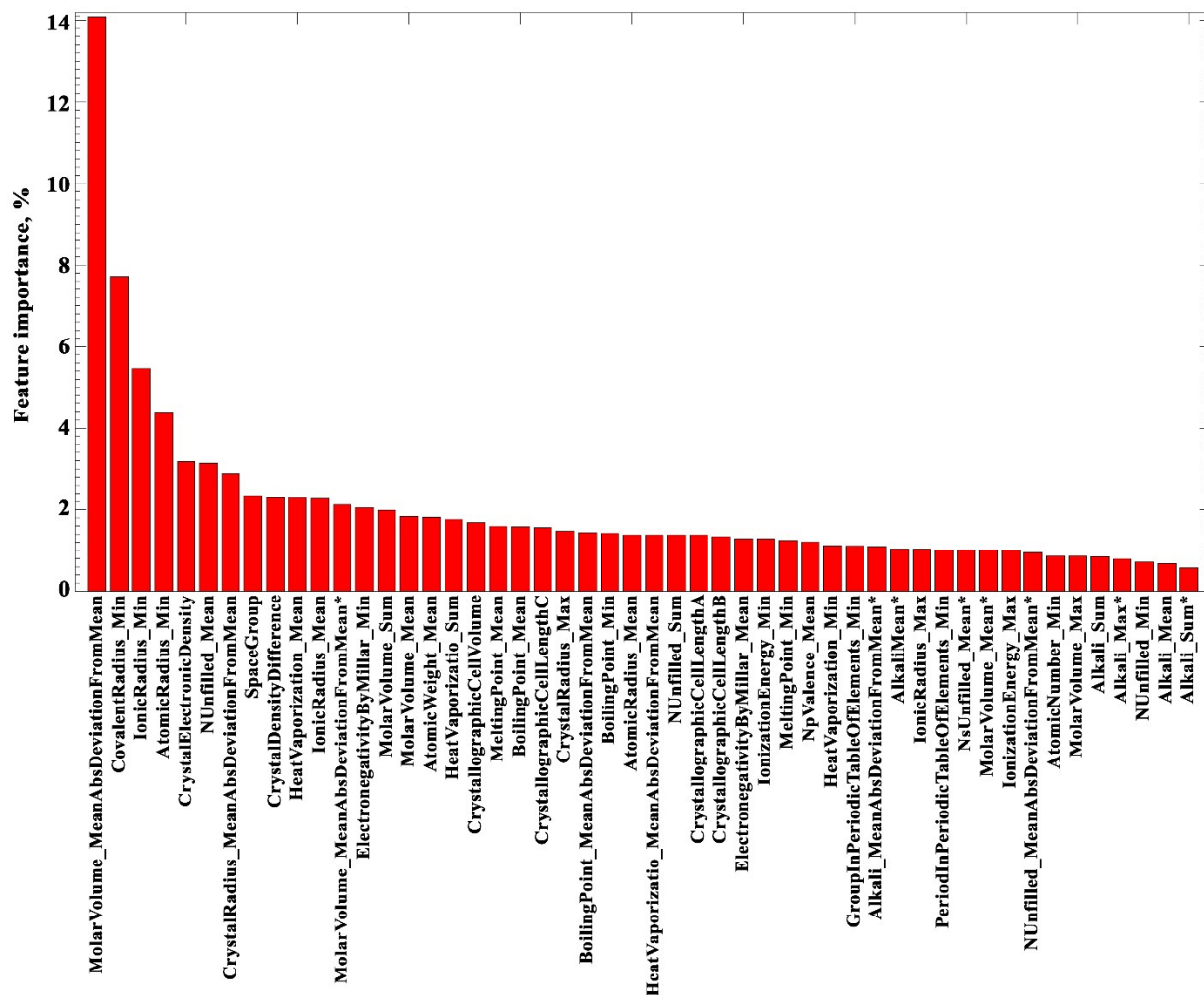
### **Part S2. Attribute ranking and reduction**

The total number of attributes generated for training is 364. The initial Random Forest-based model and grid search for parameters were done with maximum number of attributes. Feature importance ranking is available as part of sklearn package implementation of Random Forest algorithm. Features were sorted according to their importance for the model. Reduced list of ranked attributes from previous model was used for construction of Random Forest model at the next step of optimization. 10-fold cross-validation was used at each step of model optimization. Performance of each model was estimated with different statistical measures (Table S2). Taking into account all statistical estimates one could conclude that the optimum predictive performance is achieved with the regression model built using 50 attributes.

**Table S3.** The 10-fold cross-validated models with varying number of attributes using Random Forest algorithm. Performance of models was estimated with metrics including correlation

coefficient (R), coefficient of determination ( $R^2$ ), mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE), and relative squared error (RSE)

<b># of attributes</b>	<b>R</b>	<b>R<sup>2</sup></b>	<b>MAE</b>	<b>RMSE</b>	<b>RAE,%</b>	<b>RSE,%</b>
364	0.91	0.84	13.58	22.7	30.19	16.41
200	0.92	0.84	13.56	22.65	30.14	16.33
100	0.92	0.84	13.57	22.45	29.98	15.93
50	0.92	0.84	13.58	22.35	29.99	15.79
40	0.92	0.84	13.74	22.47	30.34	15.97
30	0.92	0.84	13.88	22.53	30.66	16.05
20	0.91	0.84	14.23	22.88	31.43	16.55
10	0.89	0.80	15.89	25.42	35.09	20.43



**Figure S4.** Attribute importance ranking in the Random Forest model. The attribute labels could be seen in Table S3.

**Table S4.** The list of attributes and their descriptions for the final Random Forest model discussed in this work.

Attribute importance, %	Attribute label	Attribute type based on
14.09	MolarVolume_MeanAbsDeviationFromMean	properties of atoms
7.72	CovalentRadius_Min	properties of atoms
5.46	IonicRadius_Min	properties of atoms
4.38	AtomicRadius_Min	properties of atoms
3.18	CrystalElectronicDensity	property of crystallographic cell
3.14	NUnfilled_Mean	properties of atoms
2.89	CrystalRadius_MeanAbsDeviationFromMean	properties of atoms
2.35	SpaceGroup	property of crystallographic cell
2.30	CrystalDensityDifference	property of crystallographic cell

2.29	HeatVaporization_Mean	properties of atoms
2.27	IonicRadius_Mean	properties of atoms
2.12	MolarVolume_MeanAbsDeviationFromMean*	coefficient weighted properties of atoms
2.05	ElectronegativityByMillar_Min	properties of atoms
1.99	MolarVolume_Sum	properties of atoms
1.83	MolarVolume_Mean	properties of atoms
1.81	Atomic_weightMean	properties of atoms
1.76	HeatVaporization_Sum	properties of atoms
1.68	CrystallographicCellVolume	property of crystallographic cell
1.59	MeltingPoint_Mean	properties of atoms
1.58	BoilingPoint_Mean	properties of atoms
1.56	CrystallographicCellLengthC	properties of atoms
1.47	CrystalRadius_Max	properties of atoms
1.44	BoilingPoint_MeanAbsDeviationFromMean	properties of atoms
1.42	BoilingPoint_Min	properties of atoms
1.38	AtomicRadius_Mean	properties of atoms
1.38	HeatVaporization_MeanAbsDeviationFromMean	properties of atoms
1.38	Nunfilled_Sum	properties of atoms
1.38	CrystallographicCellLengthA	properties of atoms
1.34	CrystallographicCellLengthB	properties of atoms
1.29	ElectronegativityByMillar_Mean	properties of atoms
1.29	IonizationEnergy_Min	properties of atoms
1.25	MeltingPoint_Min	properties of atoms
1.21	NpValence_Mean	properties of atoms
1.12	HeatVaporizatio_Min	properties of atoms
1.11	GroupInPeriodicTableOfElements_Min	properties of atoms
1.10	AlkaliMeanAbsDeviationFrom_Mean*	coefficient weighted properties of atoms
1.04	Alkali_Mean*	coefficient weighted properties of atoms
1.03	IonicRadius_Max	properties of atoms
1.02	PeriodInPeriodicTableOfElements_Min	properties of atoms
1.02	NsUnfilled_Mean*	coefficient weighted properties of atoms
1.02	MolarVolume_Mean*	coefficient weighted properties of atoms
1.01	IonizationEnergy_Max	properties of atoms
0.95	Nunfilled_MeanAbsDeviationFromMean*	coefficient weighted properties of atoms
0.86	AtomicNumber_Min	properties of atoms
0.86	MolarVolume_Max	properties of atoms
0.85	Alkali_Sum	properties of atoms
0.79	Alkali_Max*	coefficient weighted properties of atoms
0.71	Nunfilled_Min	properties of atoms
0.68	Alkali_Mean	properties of atoms
0.57	Alkali_Sum*	coefficient weighted properties of atoms

### Part S3. Experimental and predicted values of Bulk modulus

**Table S5.** Predicted vs experimentally measured values of Bulk modulus, structural and composition details of materials.

ICSD#	Formula	SG	TN	Bulk <sub>Predicted</sub>	Formula (if different from ICSD)	Bulk <sub>Exp</sub>	SG (if reported)
58289	Ag <sub>0.5</sub> In <sub>0.5</sub> La <sub>1</sub>	Pm-3m	221	67.2316	Ag <sub>0.89</sub> In <sub>0.02</sub> La <sub>1</sub>	43.9 <sup>1</sup>	Pm-3m
58289, 173984, 174087, 174088, 174089, 174090, 174091, 174092, 246904, 247821, 281041; 31058, 35540, 605623	Ag <sub>2</sub> O <sub>1</sub>	Pn-3mZ Pn-3mS	224	126.86±0.26		84 <sup>2</sup>	Pn-3m
70055	Ag <sub>6</sub> Ge <sub>10</sub> P <sub>12</sub>	I-43m	217	56.22		71.4 <sup>1</sup>	I-43m
35331	AlIB <sub>12</sub>	P41212	92	122.68		139 <sup>3</sup>	monophase with impurities
159334, 43851, 52282, 606058, 654653	AlIB <sub>2</sub>	P6/mmm	191	151.55±0.05		170 <sup>4</sup>	
187973, 57596, 606532, 606533, 606535, 657493	AlICo <sub>1</sub>	Pm-3m	221	115.22±1.62		162 <sup>5</sup>	PM-3m
151203, 159135, 185281, 57695, 607008, 607009, 607010, 607011, 607012, 607014, 607016, 607017	AlICu <sub>2</sub> Mn <sub>1</sub>	Fm-3m	225	73.97±0.38		109 <sup>1</sup>	
106247, 58155, 609224, 609238	AlIRu <sub>1</sub>	Pm-3m	221	136.89±2.15		207 <sup>6</sup>	multiphase
24804, 151218, 190410, 44325, 44978, 52651, 58160, 609287, 609288, 609290	AlISb <sub>1</sub>	F-43m	216	48.46±0.13		59 <sup>7</sup>	F-43m
43359, 43424, 610534, 43950, 43951	AsIGa <sub>1</sub>	F-43m	216	59.30		75 <sup>8</sup>	F-43m
150571, 58437, 611778	Au <sub>1</sub> Cu <sub>1</sub> Zn <sub>2</sub>	Fm-3m	225	72.15		132 <sup>1</sup>	
108027, 20384	B <sub>1</sub> Be <sub>2</sub>	Fm-3m	225	137.92		124 <sup>3</sup>	monophase with impurities
25535	B <sub>1</sub> Be <sub>4</sub>	P4/nmmS	129	136.96		123 <sup>3</sup>	monophase with impurities

150598, 236951, 92832, 93925	B2MgI	P6/mmm	191	129.143		145 <sup>9</sup> , 120 <sup>10</sup> , 151 <sup>11</sup>	polycrystalline aggregate, P6/mmm; P6/mmm
24701; 237013, 615593, 615594, 659126	BTiI	Pnma	62, 191	154.11±2.26		255 <sup>12</sup> , 240 <sup>13</sup> , 251 <sup>3</sup> , 276 <sup>14</sup> , 237 <sup>15</sup> , 346 <sup>15</sup> , 323 <sup>15</sup>	NA, (polycrystalline aggregate), monophase with impurities, two-phase ceramics, TiB2 98.89, the rest is impurities, NA, NA, NA
30327, 44492, 615766, 615771	B2ZrI	P6/mmm	191	144.24		218 <sup>16</sup>	P6/mmm
58745	Be12TiI	P6/mmm	191	113.09		134 <sup>3</sup>	monophase with impurities
58758, 616516	Be13ZrI	Fm-3c	226	128.33		122 <sup>3</sup>	Bi-phase
183373, 186889, 44724, 616412, 616413	Be1SI	F-43m	216	89.19±0.56		105 <sup>17</sup>	F-43m
183374, 192056, 53944, 616419	Be1SeI	F-43m	216	84.43±0.07		92 <sup>17</sup>	F-43m
183375, 290008, 53945, 616439	Be1TeI	F-43m	216	56.88±0.12		67 <sup>17</sup>	F-43m
22174, 236387, 44290, 53834, 53848	Br1CsI	Pm-3m	221	27.01±0.06		15.4 <sup>18</sup>	F-43m



27982, 44274, 52236, 53819	BrLi1	Fm-3m	225	30±0.02		26 <sup>18</sup>	F-43m
41549, 75816, 79574, 83231, 89142	C1B2Ni2Y1	I4/mmm	139	124.51		200 <sup>19</sup>	polycrystalline thin slabs
15325, 164970, 164971, 164972, 24169, 27051, 618777, 159380, 161172, 18136, 23887, 24168, 27475, 28187, 28310, 28341, 28895, 164973, 164974, 164975, 618779, 38150, 27635, 28360	C1Si1	P63mc R3mH F-43m P3m1	186 160 216 156	169.46±9.52		184 <sup>3</sup> , 211 <sup>7</sup>	Polycrystalline (?) with impurities
181020, 181244, 28730, 29008, 40938, 41413, 44618, 44937, 52754, 53978, 60368, 60369, 60370, 60371, 60559, 655133, 656448, 76045, 82707	Ca1F2	Fm-3m	225	80.58±6.01		84 <sup>20</sup>	Fm-3m
183368	Cd0.5Mn0.5Te1	F-43m	216	42.139±0.90	Cd0.55Mn0.45Te	40.3 <sup>1</sup>	
40413	Cd0.4Mn0.6Te1	Fm-3m	225	42.74		42.3 <sup>1</sup>	
102073, 620528, 620546; 108237; 246692, 246693; 620524, 620526, 620544; 620527, 620531, 620533, 93942, 93943, 93944, 93945; 620534; 67862, 67863, 67864	Cd1Se1	I41/amdS Pmm2 Cmcm Fm-3m F-43m P63mc P3121	186	44.01±0.24		53 <sup>8</sup>	
67862, 67863, 67864; 93942, 93943, 93944, 93945	Cd1Te1	P3121 F-43m	141	37.15		42 <sup>8</sup>	
22173, 44289, 53847, 622366, 622367, 655032	Cl1Cs1	Pm-3m	221	35.87±0.10		17.6 <sup>18</sup>	
26909, 27981, 44273, 52235,	Cl1Li1	Fm-3m	225	42.38±0.24		32.9 <sup>18</sup>	

53818							
61563	Cs1F1	Pm-3m	221	41.64		19.3 <sup>18</sup>	
44291, 44938, 53835, 53849, 56522, 56523, 56524, 60498	Cs1I1	Pm-3m	221	23.79±2.53		12.5 <sup>18</sup>	
103056, 628351	Cu1Mg1Zn1	Fd-3mS	227	57.11±0.01		58 <sup>1</sup>	
108422	Dy0.73Fe2Tb0.27	Fd-3mS	227	76.68		113 <sup>1</sup>	
160213	Dy2O3	P-3m1	164	111.68		231 <sup>21</sup>	
180837, 184538, 185242, 185606, 248578, 248579, 41269, 61269, 630042, 630047, 66736, 82421, 96208	Dy2O3	Ia-3	206	112.02±0.06		150 <sup>22</sup> , 191 <sup>21</sup>	
180843, 184540, 186592, 27774, 39186, 39187, 39188, 39189, 39190, 39521, 53406, 55832, 630890, 630897, 94888, 96210	Er2O3	Ia-3	206	117.03±0.36		155 <sup>23</sup> , 140 <sup>22</sup> , 200 <sup>24</sup> , 136 <sup>25</sup>	
194513	Eu2O3	P-3m1	164	92.21		134 <sup>25</sup>	P-3m1
40472, 631460, 631468, 659185, 96953	Eu2O3	Ia-3	206	90.36±4.13		115 <sup>25</sup>	Ia-3
53828	F1Rb1	Fm-3m	225	27.97		24.1 <sup>18</sup>	Fm-3m
53828, 168800, 181245, 248431, 248432, 260881, 262348, 40414, 41402	F2Sr1	Fm-3m	225	64.79±0.28		71.3 <sup>20</sup>	Fm-3m
158504, 158505, 158506, 158740, 158741, 158742, 158743, 158744, 158745, 158746, 165105, 174321, 183969, 183970, 26410, 30860, 49549, 50567, 64829, 65339, 65340, 65341, 75627, 77588, 77589, 77590, 77591, 77592, 84098, 84611, 85177;	Fe3O4	Fd-3mZ Fd-3mS	227	169.52±0.86		187 <sup>26</sup>	Fd-3mZ

85806, 85807, 96012, 98085, 98087							
103795, 635033, 635040	GalP1	F-43m	216	66.22		89 <sup>8</sup>	
44843, 635317	GalSb1	F-43m	215	40.65		57 <sup>8</sup>	
160211, 162247	Gd2O3	P-3m1	164	125.64±0.03		142 <sup>27</sup> , 160 <sup>28</sup> , 174 <sup>29</sup> , 145 <sup>25</sup>	P-3m1
150677, 152449, 160886, 165505, 181377, 183132, 184536, 184590, 184591, 184592, 184593, 184594, 184595, 193365, 40473, 41270, 636104, 636111, 636112, 94892, 96207	Gd2O3	Ia-3	206	131.56±0.12		118 <sup>27</sup> , 188 <sup>28</sup> , 134 <sup>29</sup> , 125 <sup>25</sup>	Ia-3
167204, 189805; 636533; 16954, 173528; 173890; 189803, 189804, 248478, 248479, 94305, 94306; 245956, 245958, 245959, 245960, 245961; 246628; 53643; 636526; 636527, 636528, 43422, 53788, 53642	Ge1	P63/mmc P63mc P43212 Ia-3 Cmca R-3H Imma I41/amdS Fm-3m Fd-3mS	194 186 96 206 64 148 74 141 225 227	64.68±5.26		4.06 <sup>8</sup> , 78 <sup>7</sup>	
56204	Hg0.7Mn0.3Te1	F-43m	216	37.19	Hg0.8 Mn0.2 Te1	40.8 <sup>1</sup>	F-43m
168445	Hg5In2Te8	F-43m	216	36.09		33.6 <sup>1</sup>	F-43m
160214	Ho2O3	P-3m1	164	116.84		204 <sup>30</sup>	P-3m1
152458, 184539, 185248, 27773, 41268, 44516, 639519, 639523, 639525, 80038, 82422, 96209	Ho2O3	Ia-3	206	116.57±0.94		206 <sup>30</sup> , 127 <sup>23</sup> , 134 <sup>22</sup> , 178 <sup>27</sup> , 178 <sup>31</sup>	Ia-3
27983, 414244, 44275, 53820	IlLi1	Fm-3m	225	27.98±0.09		19.3 <sup>18</sup>	Fm-3m
10021, 108558,	In1Sb1	Pmm2	25	37.80±3.17		47 <sup>8</sup>	

57396, 655298, 659843; 10022; 44641; 108559, 659364; 157947; 44331, 640427, 640431; 640432		PmnnS Pm-3m I41/amdS Cmcm F-43m P63mc					
100205, 151763, 154586, 160205, 184524, 192270, 24693, 56771, 641599, 641603, 96196	La2O3	P-3m1	164	99.13±4.16		113 <sup>25</sup> , 135.6 <sup>32</sup>	P-3m1
173180, 173193, 173206, 180569, 182024, 182025, 182026, 182027, 182028, 22402, 54368, 57411, 60431, 642216, 642219	Li2O1	Fm-3m	225	67.15		88 <sup>33</sup>	Fm-3m
151762, 168341, 193361, 194467, 40471, 428543, 642475, 642477, 96211	Lu2O3	Ia-3	206	144.91±0.29		144 <sup>25</sup> , 214 <sup>34</sup> , 113.5	Ia-3
187171, 28903, 30241, 41234, 44930, 53939, 603166, 642784, 642786, 642787, 659124	Mg1S1	Fm-3m	225	67.51±0.25		80 <sup>35</sup>	Fm-3m
160208, 169442, 184526, 191536, 28180, 32514, 645656, 645658, 645659, 645663, 96199	Nd2O3	P-3m1	164	107.84±0.20		135.6 <sup>32</sup>	P-3m1
60210, 647455, 77685	O3Sm2	P-3m1	164	88.24±0.49		224 <sup>36</sup> , 130 <sup>25</sup> , 155 <sup>37</sup>	P-3m1
165779, 184535, 40475, 647450, 96206	O3Sm2	Ia-3	206	95.99±0.20		142 <sup>36</sup> , 116 <sup>25</sup> , 149 <sup>37</sup>	Ia-3
184541, 61270, 647578, 647581, 78582, 90669	O3Tm2	Ia-3	206	94.63±0.29		143 <sup>38</sup> , 130 <sup>22</sup> , 154.5 <sup>31</sup>	Ia-3
152465, 163730, 165770, 191322,	O3Yb2	Ia-3	206	59.66±0.10		181 <sup>39</sup>	Ia-3

202904, 27775, 39179, 39180, 39181, 39182, 39183, 39184, 41266, 420428, 62871, 62872, 62873, 62874, 647660, 647662, 647665, 78583, 84138, 96959							
156404	Pb <sub>0.71</sub> Sn <sub>0.29</sub> Te <sub>1</sub>	Pnma	62	33.53		36.3 <sup>1</sup>	
250762, 53932, 648434, 648437, 648443, 648450	Pb <sub>1</sub> Si	Fm-3m	225	49.74±0.01		52.9 <sup>40</sup>	
15736, 15737, 15738, 15739, 15740, 15741, 15742, 15743, 15744, 291065, 291066, 37393, 37394, 37395, 37396, 37397, 37398, 37399, 37400, 42178, 42179, 42180, 42181, 42182, 42183, 42184, 42185, 42186, 42187, 42188, 42189, 42190, 42191, 42192, 42193, 42194, 42195, 42196, 42779, 42780, 42781, 42782, 42783, 42784, 42785, 42786, 42787, 42788, 42789, 42790, 42791, 42792, 42793, 42794, 42795, 42796, 42797, 42798, 42799, 42800, 42801, 42802, 42803, 42804, 42805, 42806, 42807, 42808, 42809, 42810, 42811, 42812, 42813, 42814,	SiZn <sub>1</sub>	P63mc R3mH P3m1 F-43m P63mc	186 160 156 216 186	54.34±3.37		77 <sup>8</sup>	

42815, 42816, 42817, 42818, 42819, 42820, 42821, 42822, 42823, 42824, 42825, 42829, 42830, 42831, 42832, 42833, 42834, 42835, 42836, 42837, 42838, 42839, 42840, 42841, 42842, 42843, 42844, 42845, 42846, 42847, 42848, 42849, 42850, 42851, 42852, 42853, 42854, 42989, 42990, 42991, 42992, 42993, 42994, 42995, 42996, 42997, 43068, 43069, 43070, 43071, 43072, 43073, 52223, 67790, 76954						
181761  53952	Se1Zn1	F-43m	216	49.50±0.05		62 <sup>8</sup>
109025, 16955, 246372, 30101, 41991, 43403, 43610, 51688, 52266, 52458, 52459, 57187, 652255, 652258, 659044, 89414, 89415, 89416	Si1	I41/amdS Ia-3 P63mc Fd-3mS Fm-3m P63/mmc Cmca	141 206 186 206 227 225 194 64	90.59±4.73		100 <sup>7</sup>
184481, 184482, 184483, 184484, 184485, 184486, 184487, 184488, 184489, 184490, 184491, 184492, 184493, 184494, 184495, 184496, 184497, 78310, 80076	Te1Zn1	F-43m P3121 Cmcm Fm-3m P31	216 152 63 225 144	43.42±1.00		51 <sup>8</sup>



## Part S4. ThermoEl tool for prediction of Bulk modulus

The Bulk modulus prediction component is an extension of our ThermoEl toolkit (Fig. S4). In order to obtain prediction one shall upload the cif-file of material. References are given on the web for appropriate literature sources. An example of cif-file content on experimentally studied  $\text{Cd}_{0.4}\text{Mn}_{0.6}\text{Te}$  could be also viewed by exploring window on the right side of the page.

### ThermoEl toolkit

HomeSeebeck coefficientBulk modulus

**Welcome to Bulk modulus component of our toolkit**

It is developed based on 2500 data points as they reported in "TE design lab" database <http://www.tedesignlab.org/>. In order to obtain predictions, various properties of atoms as well as some details of crystal structures were fed into Random Forest regression model.

**Obtaining prediction is easy**

Prediction for the compound under investigation is done based on information extracted from its crystallographic information (cif) file. Tool is working with cif-files as they reported on ICSD database [Hall, S. R., Allen, F. H. and Brown, I. D. (1991). "The Crystallographic Information File (CIF): A New Standard Archive File for Crystallography", *Acta Cryst.*, A47, 655-685 or [http://www.iucr.org/\\_data/assets/pdf\\_file/0019/22618/cifguide.pdf](http://www.iucr.org/_data/assets/pdf_file/0019/22618/cifguide.pdf)]

In order to get prediction of Bulk modulus (B), user should simply upload corresponding cif-file.

File:

I am currently working toward final version of the tool. Please, report bugs to [alona.furmanchuk@northwestern.edu](mailto:alona.furmanchuk@northwestern.edu).

Developed by Alona Furmanchuk  
Center for Ultra-scale Computing and Information Security (CUCIS), EECS Department, Northwestern University, Evanston, IL 60208, USA

The cif-file for compound  $\text{Cd}_{0.4}\text{Mn}_{0.6}\text{Te}$  reported to have  $B = 42.3$  GPa

```
#(C) 2015 by Fachinformationszentrum Karlsruhe. All rights reserved.
data_40413-ICSD
_database_code_ICSD 40413
_audit_creation_date 1999-06-02
_audit_update_record 2004-10-01
_chemical_name_systematic 'Cadmium manganese telluride (0.4/0.6/1)'
_chemical_formula_structural '(Cd0.4 Mn0.6) Te'
_chemical_formula_sum 'Cd0.4 Mn0.6 Te1'
_chemical_name_structure_type NaCl
_exptl_crystal_density_diffrn 5.17
_publ_section_title 'Atomic substitution in Cd(1-x)Mn(x)Te for 0.1<= x= 0.4'
loop
_citation_id
_citation_journal_full
_citation_year
_citation_journal_volume
_citation_page_first
_citation_page_last
_citation_journal_id_ASTM
primary 'Acta Crystallographica, Section C: Crystal Structure Communications'
1989 45 545 548 ACSCEE
loop
```

**Figure S5.** Screenshot of Bulk modulus predictive tool as it is implemented in ThermoEl tool kit. It is publicly available online at <http://info.eecs.northwestern.edu/ThermoEl>

### References

1. C. Li, Y. Li and P. Wu, *Intermetallics*, 2004, **12**, 103-109.
2. R. G. Greene, H. Luo, A. L. Ruoff, S. S. Trail and F. J. DiSalvo, Jr., *Phys. Rev. Lett.*, 1994, **73**, 2476–2479.
3. W. H. Gust, A. C. Holt and E. B. Royce, *J. Appl. Phys.*, 1973, **44**, 550-560.
4. I. I. Mazin and V. P. Antropov, *Physica C*, 2003, **385**, 49-65.
5. R. L. Fleischer, *J. Mater. Res.*, 1993, **8**, 49-58.
6. R. Fleischer, L. , R. D. Field and C. L. Briant, 1991, **22A**, 403-414.
7. S. Kamran, K. Chen and L. Chen, *Phys. Rev. B* 2008, **77**, 094109
8. M. L. Cohen, *Phys. Rev. B*, 1985, **32**, 7988-7991.
9. U. U. Harms, A. A. Serquis, R. B. Schwarz and V. F. Nesterenko, *J. Supercond.: Incorporating Novel Magn.*, 2003, **16**, 941-944.
10. K. Prassides, Y. Iwasa, T. Ido, D. H. Chi, K. Uchara, E. Nishibori, M. Tanaka, M. Sakata, Y. Ohishi, O. Shimomura, T. Muranaka and J. Akimitsu, *Phys. Rev. B*, 2001, **64**, 12509.
11. T. Vogt, G. Schneider, J. A. Hriljac, G. Yang and J. S. Abell, *Phys. Rev. B*, 2001, **62**, 220505.
12. C. A, *Engineered Materials Handbook*, ASM International, Metals Park, OH, 1991.
13. R. G. Munro, *J. Res. Natl Inst. Stand. Technol.*, 2000, **105**, 709-720.
14. S. P. Dodd, M. Cankurtaran, G. A. Saunders and B. James, *J. Mater. Sci.*, 2001, **36**, 3989-3996.



15. D. P. Dandekar and D. C. Benfanti, *J. Appl. Phys.*, 1993, **73**, 673-679.
16. T. I. Serebriakova, V. A. Neronov and P. D. Pashev, *High-Temperature Borides*, Metallurgy, Moscow, 1991.
17. K. Li, Z. Ding and D. Xue, *Phys. Status Solidi B* 2011, **248**, 1227–1236.
18. S. Narain, *Phys. Stat. Sol.*, 1994, **182**, 273-278.
19. S. Meenakshi, V. Vijayakumar, R. S. Rao, B. K. Godwal, S. K. Sikka, Z. Z. Hossain, R. R. Nagarajan, L. C. L.C. Gupta and R. R. Vijayaraghavan, *Physica B*, 1996, **223&224**, 93-95.
20. G. K. White, *J. Phys. C: Solid State Phys.* , 1980, **13**, 4905–4913
21. S. Jiang, J. Liu, C. Lin, L. Bai, Y. Zhang, X. Li, Y. Li, L. Tang and H. Wang, *Solid State Comm.*, 2013, **169**, 37-41.
22. Y. Makino and S. Miyake, *Alloys Compd.*, 2000, **313**, 235-241.
23. W. R. Manning and O. Hunter, , Jr., *J. Am. Ceram. Soc.*, 1969, **52**, 492-496.
24. Q. Guo, Y. Zhao, C. Jiang, W. L. Mao, Z. Wang, J. Zhang and Y. Wang, *Inorg. Chem.*, 2007, **46**, 6164-6169.
25. J. P. McClure, PhD thesis, University of Nevada, 2009.
26. D. L. Anderson and O. L. Anderson, *J. Geophys. Res.* , 1970, **75**, 3494–3500.
27. D. Lonappan, C. Shekar, N. V. Sahu, P. C. Kumarasamy, B. V. Bandyopadhyay and K. Rajagopalan, *Philos. Mag. Lett.*, 2008, **88**, 473-479
28. F. X. Zhang, M. Lang, J. W. Wang, U. Becker and R. C. Ewing, *Phys. Rev. B*, 2008, **78**, 064114.
29. L. Bai, J. Liu, X. Li, S. Jiang, W. Xiao, Y. Li, L. Tang, Y. Zhang and D. Zhang, *J. Appl. Phys.*, 2009, **106**, 073507.
30. S. Jiang, J. Liu, X. Li, L. Bai, W. Xiao, Y. Zhang, C. Lin, Y. Li and L. Tang, *J. Appl. Phys.*, 2011, **110**, 013526.
31. P. C. Sahu, D. Dayana Lonappan and S. N.V.C., *J. Phys.: Conf. Ser.*, 2012, **377**, 012015.
32. K. K. Pandey, N. Garg, A. K. Mishra and S. M. Sharma, *J. Phys. : Conf. Ser.* , 2012, **377**, 012006.
33. A. Otero-de-la-Roza and V. Luan~a, *J. Phys. Chem. A*, 2011, **115**, 12953–12961.
34. S. Jiang, J. Liu, C. Lin, L. Bai, W. Xiao, Y. Zhang, D. Zhang, X. Li, Y. Li and L. Tang, *J. Appl. Phys.*, 2010, **108**, 083541.
35. P. E. Van Camp, V. E. Van Doren and J. L. Martins, *Phys. Status Solidi B*, 1995, **190**, 193–197.
36. Q. Guo, Y. Zhao, C. Jiang, W. L. Mao and Z. Wang, *Solid State Comm.*, 2008, **145**, 250-254.
37. S. Jiang, J. Liu, C. Lin, X. Li and Y. Li, *J. Appl. Phys.*, 2013, **113**, 113502.
38. S. L. Dole, O. Hunter and F. W. Calderwood, *J. Am. Ceram. Soc.*, 1980, **63**, 136-139
39. C. Meyer, J. P. Sanchez, J. Thomasson and J. P. Itié, *Phys. Rev. B*, 1995, **51**, 12187.
40. A. Chroneos and R. V. Vovk, *J. Mater. Sci. – Mater. Electron.*, 2015, **26**, 2113-2116.

