Supplementary information:

High-throughput Density Functional Perturbation Theory and Machine Learning Predictions of Infrared, Piezoelectric and Dielectric Responses

Kamal Choudhary[1], Kevin F. Garrity[1], Vinit Sharma[2,3], Adam Biacchi[4], Angela R. Hight Walker[4], Francesca Tavazza[1]

1 Materials Science and Engineering Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA.

2 National Institute for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA.

3 Joint Institute for Computational Sciences, University of Tennessee, Knoxville, TN, 37996, USA.

4 Engineering Physics Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA.

Table S1 Comparison of experimental and DFPT IR frequencies (cm^{-1}).

Mats.	JID	DFPT	Experiment
ZnO	1195	379, 410	389,413[1]
AlN	39	600, 653	620,669[2]
GaN	30	532	531[3,4]
SnS	1109	93, 144, 178, 214	99, 145, 178, 220 [5]
SnSe	299	72.6, 98.44, 125.01, 160.3	80,96,123, 150 [5]
KMgF3	20882	160.0, 287.0, 470.8	168, 299, 458[6]
LiNbO3	1240	145.0, 216.6	160, 220[7]
GeS	2169	106.4, 196.2, 236.5, 253.0, 276.9	118, 201,238,258, 280[8]
MAD			8.36

Table S2 Comparison of piezoelectric coefficient max(e_{ij}) data for experiments and DFT. We take average values for the cases where the experimental data are in a range.

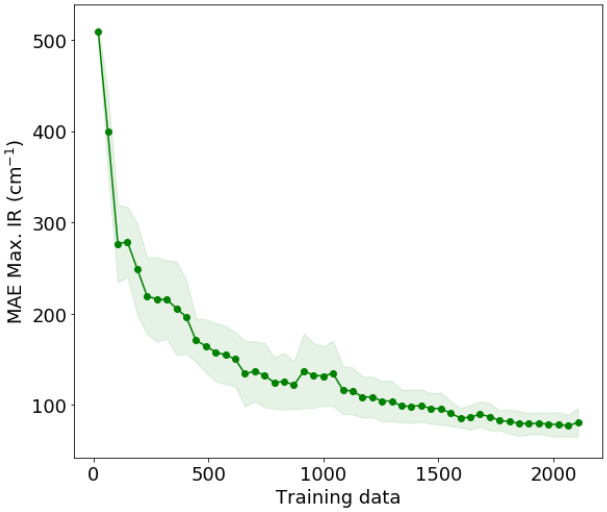
Mats.	**JID**	**Max(e_{ij})**	**DFT**	**Reference**
BN	57695	1.55	1.15	[9]
AlN	39	1.39-1.55	1.39	[10-12]
ZnS	7648	0.38	0.13	[9]
ZnSe	8047	0.35	0.06	[9]
SiO_2	41	0.171	0.16	[13,14]
$BaTiO_3$	110	1.94-5.48	4.13	[15,16 17,18]
$LiNbO_3$	1240	1.33	1.59	[19]
GaSb	35711	-0.07	-0.102	[9]
$PbTiO_3$	3450	3.35-5.0	3.96	[20]
GaN	30	0.73	0.47	[21]
InN	1180	0.97	0.90	[22]
AlP	1327	-0.1	0.004	[9]
AlAs	1372	-0.16	0	[9]
AlSb	1408	-0.13	0.06	[9]
ZnO	1195	1.00-1.55,0.89	1.10	[23]
BeO	20778	0.1	0.22	[23]
MAD			0.21	

Table. S3 Comparison of static dielectric constant for DFPT, MBJ and experiment. Experimental data were obtained from [24-28]. MBJ data were obtained from our optoelectronic property database[29].

Materials	JID	DFPT	MBJ	Experiment
MoS_2	54	ε_{11} =15.56	ε_{11}=15.34	ε_{11}=17.0
$MoSe_2$	57	ε_{11}=16.90	ε_{11}=16.53	ε_{11}=18.0
$MoTe_2$	60	ε_{11}=21.72	ε_{11}=18.74	ε_{11}=20.0
WS_2	72	ε_{11}=13.91	ε_{11}=13.95	ε_{11}=11.5
WSe_2	75	ε_{11}=15.21	ε_{11}=14.32	ε_{11}=11.7
SiC	182	7.10	6.01	6.552
AlP	1327	10.33	6.94	7.54
BN	17	ε_{11}=4.75	ε_{11}=3.72	ε_{11}=5.06
BP	1312	9.03	7.94	11.0
GaP	1393	13.22	8.33	11.11
AlSb	1408	12.27	9.87	12.04
ZnS	1702	9.39	4.8	8.0
CdTe	23	19.59	6.54	10.6
HgTe	8041	ε_{11}=29.44	ε_{11}=11.22	ε_{11}=20
$ZnSiP_2$	2376	ε_{11}=12.44	ε_{11}=8.56	ε_{11}=11.15
$ZnGeP_2$	2355	ε_{11}=14.75	ε_{11}=9.02	ε_{11}=15
MAE	-	2.46	2.78	-

Table S4 K-point dependence of average dielectric constant, maximum piezoelectric coefficient, maximum and minimum optical phonon modes for Si (JVASP-1002), AlN (JVASP-39), MgF_2 (JVASP-20134). Equivalent K-points are represented in length (Å unit), mesh and per atom units.

Materials	**KP-length(Å)**	**KP-mesh**	**KP-per atom**	**ε_{avg}**	**Max(eij)**	**Max mode (cm^{-1})**	**Min mode (cm^{-1})**
Si	35	6x6x6	1728	12.9	0.0	496.5	153.3
Si	50	9X9X9	5832	12.8	0.0	496.4	153.0
Si	65	12x12x12	13824	12.8	0.0	496.3	153.1
AlN	25	9x9x5	1620	8.6	1.39	713.9	240.8
AlN	40	15x15x8	7200	8.6	1.39	714.0	240.8
AlN	55	20x20x11	17600	8.6	1.40	713.9	240.8
MgF_2	20	4x4x7	672	5.26	0.0	507.0	86.1
MgF_2	35	8x8x11	4224	5.26	0.0	507.0	86.2
MgF_2	50	11x11x16	11616	5.26	0.0	506.9	86.2



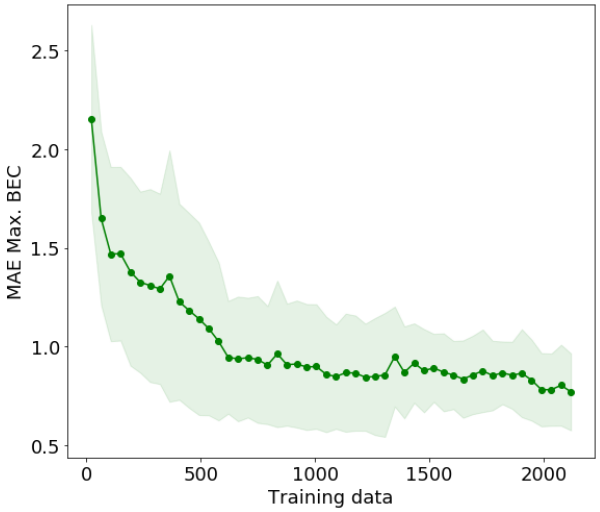


Fig S1 Regression learning curves for maximum IR mode and Born effective charge. The filled regions show standard deviation as we carry out five-fold cross-validation for each training data points in the figures.