

재료 연구 패러다임 변화: 시행착오적 접근법에서 데이터 기반 연구로

DOI: 10.3938/PhiT.26.034

윤 용·한승우

Paradigm Shift in Material Research: from Edisonian to Mechanistic to Data-driven

Yong YOUN and Seungwu HAN

In materials science, Edisonian trial-and-error approach has been used in identifying a novel material. However, this approach is rather inefficient to meet the high demand of new materials in modern engineering. The development of computer technology makes it possible to produce massive data for materials properties using quantum calculation, which opens a new avenue for material research. This article introduces the data-driven research in materials science and discuss its application in some examples.

들어가는 글

전자기기의 발달과 함께 재료에 대한 관심 또한 함께 증가하고 있다. 하지만 전자 기술의 발달과 달리 재료 분야의 발전은 상대적으로 느리다. 예를 들어 그림 1에 제시되어 있는 노트북 컴퓨터 관련 기술의 성능 발전을 보면 CPU, WIFI, 디스크 용량은 대체적으로 지수적인 성장을 하며 무어의 법칙을 따라가는 것을 볼 수 있다.^[1] 그러나 배터리 분야는 다른 기술의 빠른 발전과 달리 성장이 정체된 모습을 보인다. 반도체 분야가 배터리 분야와 달리 지수적으로 성장할 수 있는 배경은 다음의 두 가지 측면에서 수학적 모델이 잘 구축되어 있다는 점이 있다. 첫 번째는 포토 리소그래피(photo lithography) 기술의 발전에 따른 반도체 공정의 미세화이다. 트랜지스터의 모양은 세대가 증

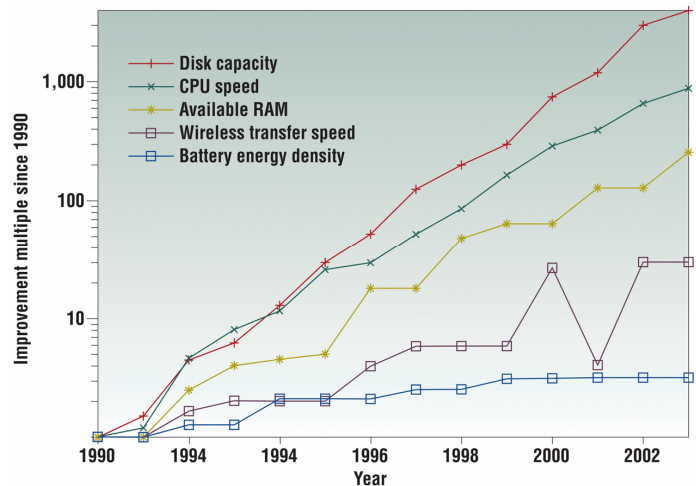


Fig. 1. Relative improvements in laptop computing technology.^[1]

가하더라도 거의 비슷하다. 그래서 빛의 경로를 수학적으로 잘 조절하여 리소그래피에 사용되는 렌즈를 더 정교하게 만들어 트랜지스터의 크기를 지속적으로 줄일 수 있다. 또 다른 이유는 컴퓨터를 이용하여 공정이나 소자의 특성을 정확하게 예측할 수 있다는 점이다. TCAD와 같은 프로그램이 대표적인데, 이러한 프로그램을 사용하면 시행착오적이고 반복적인 실험을 진행하지 않아도 되기 때문에 디자인 최적화에 걸리는 시간을 혁신적으로 단축시킬 수 있다.

반면 재료공학 연구에서는 시행착오적인 접근 방식이 많이 사용된다. 이러한 접근법을 일명 에디슨적인 접근법(Edisonian approach)이라 하는데, 에디슨이 처음 필라멘트의 소재를 결정할 때 다양한 소재를 실험, 실패를 반복하며 좋은 소재를 찾아낸 일화에서 유래되었다. 다시 말해, 에디슨적인 접근법은 연구자의 직관에 따라 수천 가지의 소재에 대해 무차별적으로 시험해 보고 그 중 가장 우수한 재료를 선택하는 방법이다. 이런 시행착오적인 방법을 벗어나 전자공학이나 기계공학과 같이 논리적으로 재료를 탐색하고자 하는 것은 재료공학자의 오랜

저자약력

윤용은 서울대학교 재료공학부에서 석사(2014) 학위를 받았고 현재 한승우 교수님 연구실에서 박사과정 중이다. (yybyb@snu.ac.kr)

한승우 교수는 서울대학교 물리학과에서 학사(1993), 석사(1995), 박사(2000) 학위를 받았다. 프린스턴 대학교에서 박사후연구원, 이화여자대학교 물리학과 교수를 거쳐, 현재 서울대학교 재료공학부 교수로 재직 중이다. (hansw@snu.ac.kr)

REFERENCES

[1] J. A. Paradiso and T. Starner, IEEE Pervasive Computing 4, 18 (2005).

염원이다. 이를 반영하듯 최근 재료연구에서는 원자적 수준의 메커니즘을 밝혀서 복잡한 현상을 이해하면서 재료를 탐색하려는 시도가 많아지고 있다. 예를 들어 물 분해에 대한 연구에서 촉매가 작동하는 각각의 과정에 작용하는 메커니즘을 이해하고 이를 바탕으로 촉매 재료를 최적화하는 방식이다.^[2,3] 이런 메커니즘적 연구 방식이 가능한 주된 이유는 투과전자현미경(TEM), 주사터널링현미경(STM)과 같은 관측 기술의 발전을 통하여 분자 단위의 현상을 파악할 수 있고,^[2] 또한 양자 계산을 통하여 메커니즘에 대해 파악할 수 있기 때문이다.^[3] 하지만 완전한 메커니즘을 이용하여 실질적인 문제를 해결하는 것이 가능한가에 대한 질문은 여전히 남아 있다. 기본적으로 물질계에서 일어나는 현상은 매우 복잡하기 때문에 논리적으로 생각할 수 있는 단순한 경로를 따라가지 않기 때문이다. 그래서 메커니즘적인 분석은 시행착오적 연구에 비해 추천할 만하지만 여전히 재료 연구에서의 근본적인 문제를 해결하기는 어려워 보인다.

이런 딜레마에서 벗어나기 위해서는 비슷한 문제점을 가지고 있는 근접 학문을 참고할 필요가 있다. 바로 생물학이다. 생체 내에서 단백질이나 호르몬의 움직임은 재료 내 원자의 움직임보다 더 복잡함에도 불구하고 생물학은 최근 유전정보에 대한 연구와 함께 상당한 발전을 이루고 있다. 예를 들어 유전자 검사를 진행하고 유전자를 이용해 다양한 질병의 발병 확률을 예측하는 기술이 있는데 특정 질환을 가진 사람들에게 공통적으로 나타나는 유전정보를 바탕으로 어떤 사람이 그 질환에 걸릴 확률을 예측하는 방법이다. 이와 같이 메커니즘에 기반한 분석을 진행하지 않고 데이터에 기반한 연구가 가능한 것은 십 년 전에는 수조 원이 필요했던 DNA 염기서열 분석 기술이 지금은 일주일에 수백 개의 시퀀스를 천 달러에 분석할 수 있을 정도로 발달했고 이와 함께 막대한 양의 데이터가 쌓이고 있기 때문이다.^[4] 이러한 DNA 정보와 임상 정보를 연결하면 생물학 연구의 혁신을 이뤄내는 것이 가능해진다. 그 결과 빅데이터(big data)를 활용하여 데이터 기반 연구만 진행하는 생물학자들도 늘어나고 있다. 다시 말해 결국에는 메커니즘을 몰라도 데이터 기반 연구를 이용해 문제를 해결할 수 있다는 것이다.

재료 연구에서도 다양한 데이터가 수집되고 있다. 특히 어떤 물질의 밴드 갭이나 구조, 탄성계수와 같은 데이터들은 잘 알려져 있다. 하지만 데이터 기반의 연구를 위해서는 촉매의 특성, 박막의 누설전류와 같이 공학적 관점에서 보다 중요한 물성에 대한 데이터베이스 구축이 이루어져야 한다. 그러나 이러한 특성들은 재료에 의해서만이 아니라 디바이스를 만드는 과정에도 의존하는 것으로 알려져 있다. 이렇게 어떤 데이터가 나올 때 그 데이터가 생성되는 조건을 메타 데이터(meta-data)라고 하

는데, 디바이스 특성을 데이터화하기 위한 메타 데이터는 복잡하고 많은 경우 알려져 있지 않다. 예를 들어 어떤 후처리 공정을 진행하였는지, 어떤 용매에서 반응했는지를 모두 기록한다면 데이터화할 수 있을 것 같지만, 실제 실험에서는 같은 조건하에서도 물성이 다르게 측정되기 때문에 실험상의 결과를 데이터베이스로 구축하는 것은 상당히 어렵다.

재료 데이터베이스

그럼에도 불구하고 최근 재료분야에서 데이터에 기반한 연구가 관심을 받고 있다. 바로 범밀도함수(DFT)나 양자 계산을 이용하여 데이터를 생산하고 데이터베이스를 구축하는 것이다.^[5,6] DFT 계산을 통한 재료물성값들은 방법에 따라 정확도가 달라질 수 있다. 예를 들어 그림 2에서와 같이 밴드 갭을 다양한 방법을 통하여 계산할 수 있는데 이때, 정밀도가 높고 시간이 많이 소요되는 계산($GGA < HSE06 < scGW < scGW \text{ e-h}$)일수록 실험에 더 가까운 값을 얻을 수 있다.^[7] 여기에서 강조할 점은 고차원적 계산 방법을 사용하는 것은 마치 테일러 급수에서 1차항에서 값을 구한 다음 만족할 만한 정확도가 아니면 2차, 3차항으로 확장하는 것과 같이 논리적이라는 것이다. 결과적으로 비용과 시간을 충분히 들이면 실험에 준하는 정확한 물성값들의 데이터베이스를 구축할 수 있다. 또한 빨라진 컴퓨터의 속도도 계산을 이용한 대용량 데이터베이스 구축을 가능하게 하였다. 과거에는 컴퓨터 자체의 성능이 뛰어나지 않았을 뿐만 아니라 그 가격 또한 매우 비싸서 많은 양의 결과를 얻기 어려웠다. 하지만 무어의 법칙에 따라 컴퓨터가 발전하면서 10년 전과 비교해 계산 능력이 수천 배 향상됐기 때문에 막대한 양의 데이터 생산이 가능해졌다.

이러한 기술의 발달을 바탕으로 많은 나라에서 계산을 통해

REFERENCES

- [2] A. Fujishima, X. Zhang and D. A. Tryk, *Surf. Sci. Rep.* **63**, 515 (2008).
- [3] E. M. Sproviero, J. A. Gascón, J. P. McEvoy, G. W. Brudvig and V. S. Batista, *J. Am. Chem. Soc.* **130**, 3428 (2008).
- [4] National Human Genome Research Institute, <https://www.genome.gov/27565109/the-cost-of-sequencing-a-human-genome> (accessed 14/8/2017).
- [5] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder and K. A. Persson, *APL Mater.* **1**, 011002 (2013).
- [6] S. Curtarolo, W. Setyawan, G. L. W. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. Mehl, H. T. Stokes, D. O. Demchenko and D. Morgan, *Comp. Mat. Sci.* **58**, 218 (2012).
- [7] S. Park, B. Lee, S. H. Jeon and S. Han, *Curr. Appl. Phys.* **11**, s337 (2011).

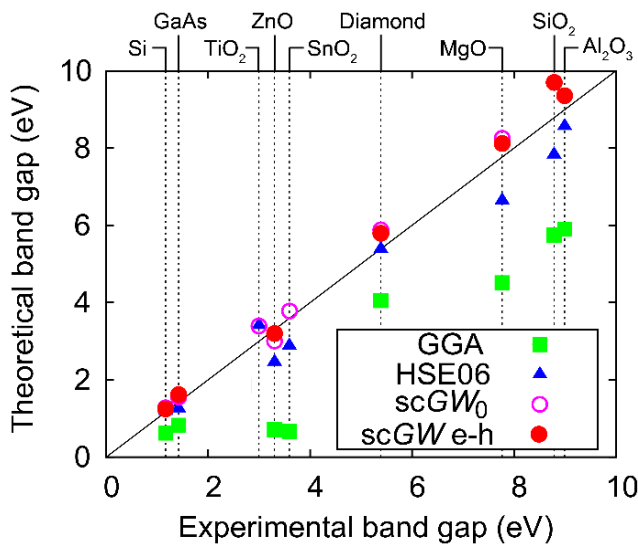


Fig. 2. The comparison of band gaps of various materials between experiment and theory.^[7]

생성된 데이터를 활용하여 데이터베이스를 구축하는 사업을 진행하고 있다. 특히 미국에서 진행하고 있는 Materials Genome Initiative가 가장 대표적이고^[8] 이를 뒤따라 일본의 AtomWork,^[9] 한국의 미래소재디스커버리 사업에서 데이터베이스를 구축하고 있다. 대표적으로 Material Project나 AFlow와 같은 웹사이트에서는 수십만에서 수백만 개에 이르는 물질의 엔탈피(enthalpy)나 상평형도(phase diagram)를 찾아볼 수 있다. 이렇게 많은 재료의 물성을 파악하게 된 것은 인류 역사상 최초의 사건이고, 이것이 재료연구에 어떤 영향을 미칠지는 앞으로 기대된다.

고유전체 물질 탐색

이와 관련하여 본 연구실에서도 자체적으로 AMP²라는 계산 패키지를 만들고 있는데, 사용자들이 데이터를 자동화하여 생산할 수 있는 패키지를 만드는 것이 목표이다. 이 패키지를 활용한 예로 전자소자의 고유전체(high-k) 물질을 찾는 연구를 진행한 바 있다.^[10] CPU나 DRAM에서 고유전체 물질이 많이 사용되는데, 유전율과 밴드 갭이 모두 큰 물질을 사용하여 높은 특성의 디바이스를 얻을 수 있다. 하지만 유전율과 밴드 갭은 서로 반비례하는 관계이기 때문에 두 가지 물성이 모두 좋은 물질을 찾기는 쉽지 않다. 그래서 본 그룹에서는 양자 계산을 통하여 서로 다른 수천 가지의 산화물들의 유전율과 밴드 갭을 계산하였다. 그림 3은 약 1600개의 산화물에 대해 밴드 갭과 유전율을 구한 결과이다. 약 1600개의 산화물에 대해서 계산하였음에도 불구하고 눈에 띄는 물질은 입방정계(cubic) 상의 산화베릴륨(BeO) 하나였다. 하지만 이 물질은 고압에서 형성되는 구조로 실제로 사용하기 위해서는 도핑 등을 통해서 상을 안정

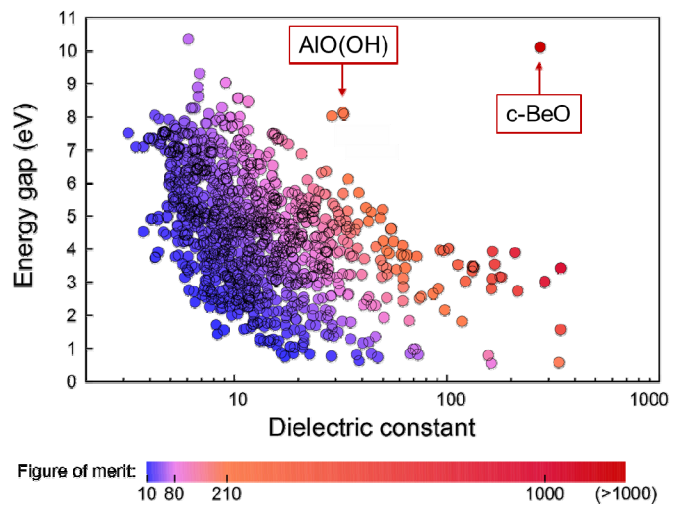


Fig. 3. Band gap (E_g) versus dielectric constant (κ) for computed 1600 oxides. Color code in each point indicates the figure of merit ($E_g \times \kappa$).^[10]

화하는 것이 필요하다. 그럼에도 불구하고 이 물질은 기존에 알고 있던 밴드 갭과 유전율 사이의 관계를 벗어나는 상당히 흥미로운 결과였다. 이 물질을 제외한다면 AlO(OH) 물질만이 기존에 사용하고 있는 산화하프늄(HfO₂)을 능가할 수 있다고 보인다. 결과적으로 자동화 계산을 통하여 많은 수의 신물질의 특성을 계산해 보아도 기존에 사용하고 있는 물질을 능가하는 것을 찾기가 상당히 어렵다는 것을 알 수 있다.

차세대 물성 데이터베이스

위와 같은 계산과학의 발전에서 자칫 계산을 통해 모든 것을 예측할 수 있다는 장밋빛 미래를 그리기 쉽지만, 실제로는 그렇지 못하다. 우선 계산할 수 있는 물성을 보면, 물질의 구조, 탄성계수, 밴드 구조와 같이 물질이 독립적인 순수한 상으로 존재할 때 가지는 물성이 대부분이다. 우리는 이런 특성들을 하나로 묶어서 컴포넌트 물성(component property)이라고 지칭한다. 반면 우리가 응용적인 측면에서 관심 있는 것은 더 복잡한 특성, 예를 들어 소자 안에서 물질의 수명, 효율, 이동도, 전류-전압 특성 같은 시스템 물성(system property)이다. 그러나 이러한 시스템 물성을 계산하는 것은 메커니즘이 완벽하게 규명되어 있지 않는 이상 불가능하다. 따라서 컴포넌트 물성을 바탕으로 시스템 특성을 예측 혹은 계산할 수 없다. 현재까지

REFERENCES

- [8] Materials Genome Initiative, <https://www.mgi.gov> (accessed 14/8/2017).
- [9] AtomWork, <http://crystdb.nims.go.jp> (accessed 14/8/2017).
- [10] K. Yim, Y. Youn, J. Lee, K. Lee, H.-H. Nahm, J. Yoo, C. Lee, C. S. Hwang and S. Han, NPG Asia Materials 7, e190 (2015).

재료 연구에서 데이터베이스가 구축되고 있는 것을 보면 컴포넌트 물성은 컴퓨터를 이용하여 굉장히 많은 데이터가 구축되어 있다. 반면 시스템 물성은 실험을 통하여 굉장히 느린 속도로 데이터가 쌓이고 있다. 이런 극심한 차이를 어떻게 극복하여 데이터 기반 연구를 진행할 것인가는 큰 문제이다. 이에 대한 한 가지 방향성을 제시하자면 데이터만을 이용해 재료연구를 하는 것은 분명 한계가 있을 것이기 때문에 메커니즘에 대한 적절한 가설을 세우고 가설의 핵심적인 서술자(descriptor)에 대해서 데이터베이스화된 컴포넌트 물성과 연결시키는 형태의 연구가 성공적인 결과를 줄 수 있을 것이다. 또한 많은 시스템 물성이 순수한 결정 상태의 물성과 더불어 결합 물성이 중요하다는 것이 널리 알려져 있다. 이런 점에서 현재 벌크 물성에 집중되어 있는 데이터베이스를 발전시켜 결합 등의 물성으로 확장한 차세대 물성 데이터 베이스 구축이 필요할 것으로 예상된다.

맺음말

기존의 재료 연구는 시행착오적인 과정을 반복하면서 최적의 물질을 찾아가는 방향으로 진행되어 왔다. 실제로 이러한 방향의 연구들이 성공적인 결과를 낳기도 하였다. 하지만 에디슨적인 연구를 통해서만 다른 분야에서처럼 빠른 발전을 이루기가 어려웠고 또한 컴퓨터의 발전과 함께 계산을 통하여 막대한 양의 재료 물성 데이터를 얻는 것이 가능하게 되면서 데이터 기반의 연구가 새로운 연구 방향으로 대두되었다. 실제로 고유전체 물질을 찾는 과정에서 데이터 기반의 연구를 진행하여 의미 있는 결과들을 얻을 수 있었다. 하지만 재료 연구에 있어서는 순수한 데이터 기반 연구만으로는 성공을 거두기 어렵다. 따라서 메커니즘에 대한 분석과 이를 바탕으로 서술자를 찾고 관련된 디지털 데이터 베이스를 활용하는 데이터 기반의 재료 연구가 새로운 길을 제시해 줄 수 있을 것이라 기대한다.