Agrawal, A., J. McHale, and A. Oettl (2019) Artificial Intelligence, Scientific Discovery, and Commercial Innovation

Comments

July 30, 2019

Koichi KUME（Visiting senior research fellow, ESRI）

Contributions of this paper

This research paper formulates AI-driven innovation as a multi-stage search process by introducing insightful four ideas:

- A vast potential combinatorial search space (spillover effect, Jones 1995)
- Uncertainty can be reduced by prediction
- The output of the prediction can be expressed in the form of a ranking function (fitness landscape, Kauffman 1993)
- Advances in AI improve the performance of the predictive maps. The innovator seeks to discover all successful combinations and then choose a single success which combinations are to go for testing (sequential search problem Roberts and Weitzman 1981)

Prediction (development of a prediction model choosing a testing threshold)

G is the number of successes and ***Pr*** is a unit step function of the known probability of success of the ***r***-th ranked potential combination

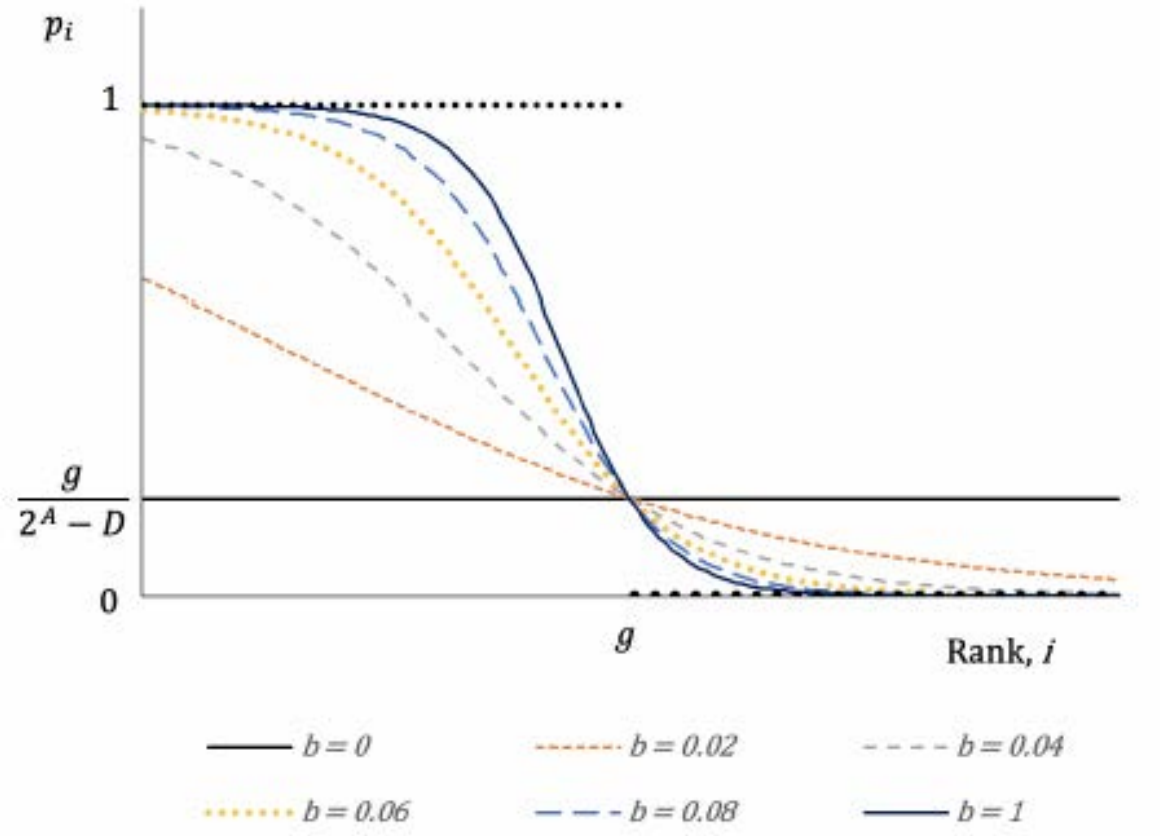
$$p_r = \begin{cases} 1 & for \quad r \leq G \\ 0 & for \quad r > G \end{cases}$$

b is a discrimination power and the logistic decay function holds where b ≥ 0

$$p_r = \frac{1}{1 + Ke^{b(r-G)}}$$

Increases in *b* cause the ranking function curve to rotate in a clockwise direction around the point (G, G/(2^A-D))

Figure 2: Ranking Function Curves for Different Values of the Discrimination Parameter, *b*



Testing (testing all combinations with a probability of success at or above a threshold value that depend on MV^e of success and MC of conducting a test)

Case 1. Search for ***all*** profitable innovation

The value of an innovation is independent of which other valuable innovations are discovered, so the innovator seeks to discover all valuable combination (summation of *Pr* up to r*)with a positive expected value net of cost of testing.

$$V^e = -cr^* + \pi \sum_{r=1}^{r^*} p_r$$

Case 2. Search for a ***single*** target innovation

The innovator has a single target. The value of our innovator of finding additional successful combinations once one success is achieved is zero (V^e is calculated by multiplication of P_r)

$$V^e = -cr^* + \pi \sum_{t=r}^{r^*} \left(\left(\prod_{j=1}^{r-1} (1 - p_{r-j}) \right) p_r \right)$$

In case 1, r^* is the optimal number of combinations to send for testing

$$r^* = G - \frac{ln\left(\frac{2^A - D - G}{G}\right) - ln\left(\frac{\pi}{c} - 1\right)}{b} \qquad where \;\; MV^* = MC$$

Multi-stage discovery process and the options to abandon or enter intermediate stage

Following Roberts and Weitzman (1981) , costs are additive across stages, the value can be received only at the end of the project and the project can be abandoned at the end of each stage. The number of paths is given by the binomial coefficient. The expected probability of success can be calculated as:

$$P_r 0.5^{s-1} \sum_{n_s=1}^{s} \left(\binom{s-1}{n_s-1} (1+u)^{s-1-(n-1)} (1-u)^{n-1} \right) = P_r$$

At the three stages model, the decision to advance combinations along the pipeline is affected by uncertainty (u) and costs at the intermediate (C_2) and final stages (C_3) .

If $Pr < \frac{C_3}{1-u}$, the option to *abandon* will be exercised.

If $Pr \geq \frac{2C_2 + C_3}{1+u}$, the option to *enter intermediate stage* will be exercised.

The optimal number of combinations to advance to intermediate stage:

$$r^{**} = G - \frac{ln\left(\frac{2^A - D - G}{G}\right) - ln\left(\frac{1+u}{2C_2 + C_3}\right) - 1}{b}$$

Implications of the three stage model

An increase in ***D*** the number of observations on prior successes and failures reduces the number of combinations available ***r**** to be discovered, but also provides training data for developing a prediction model for successful new combinations.

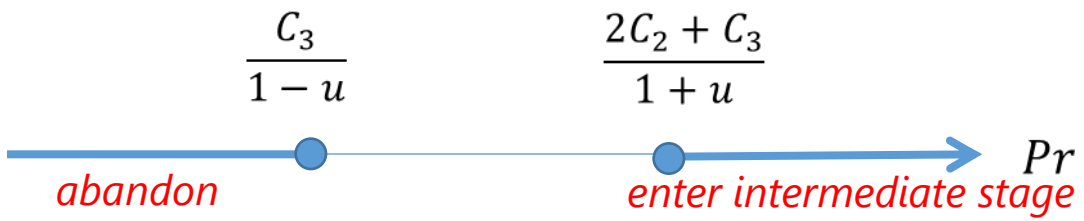
An increase in uncertainty ***u*** will raise the number of combinations ***r**** that advance to the next stage and the expected value MV.

Increase in either intermediate C_2 or final costs C_3 will lower the number of combination ***r**** that advance to the final stage if $C_2 > C_3$

Increases in ***b*** will shift the expected marginal gross value curve upwards, then the number of combinations that advance will increase, and also the expected total net value V^e will increase.

Comment 1. Are there missing assumptions of parameters?

Are the option to abandon and the option to enter intermediate stage compatible or exclusive ?



$$\frac{C_3}{1-u} < \frac{2C_2+C_3}{1+u} \quad \Rightarrow \quad \frac{C_2}{C_3} > \frac{u}{1-u}$$

Need another assumption between uncertainty and costs?

Can we replace K with $\frac{2^A-D-G}{G}$ at r** ?

K *where b=0*

$$r^{**} = G - \frac{ln\left(\frac{2^A-D-G}{G}\right) - ln\left(\frac{1+u}{2C_2+C_3}\right) - 1}{b}$$

We have another b as denominator with no assumption.

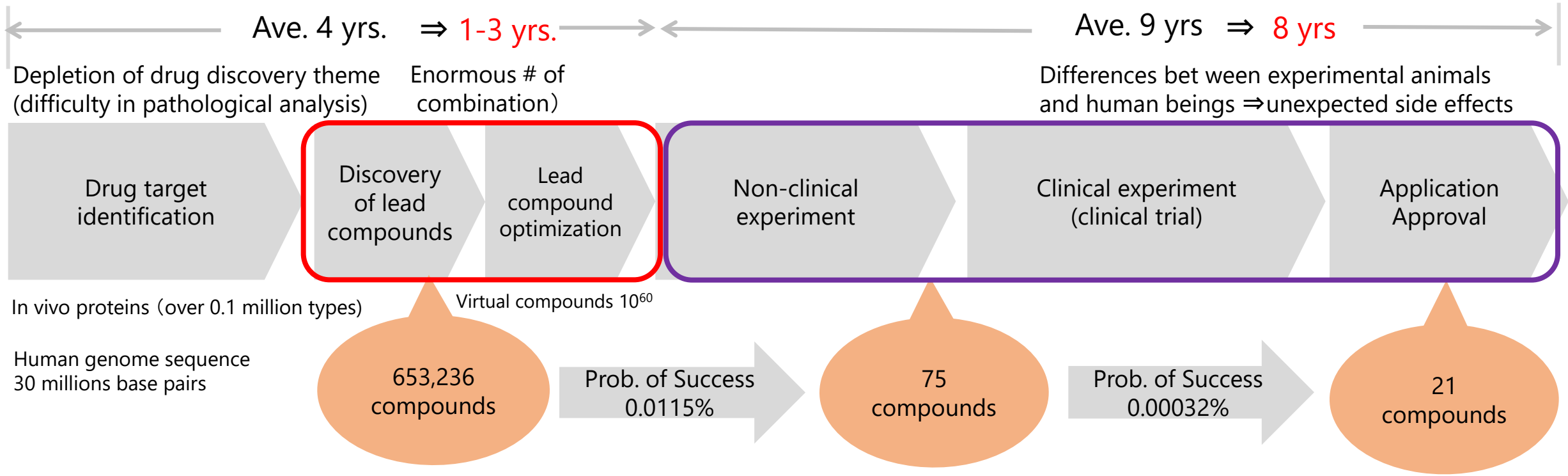
What properties K have ? Is it related with uncertainty?

$$Pr = \frac{1}{1+Ke^{b(r-G)}} \quad \Rightarrow \quad lnK = \ln\left(\frac{1}{P_r}-1\right) - b(r-G) \le ln\, u^s$$

If b=0 and an assumption $u^s \le \sqrt[s]{{p_r}^{-1}-1}$ hold.

Comment 2: Is the effects of prediction limited ? (Drug industry)

Development period	Development cost	Prob. Of success
13 yrs	$ hundreds of million	0.0032%



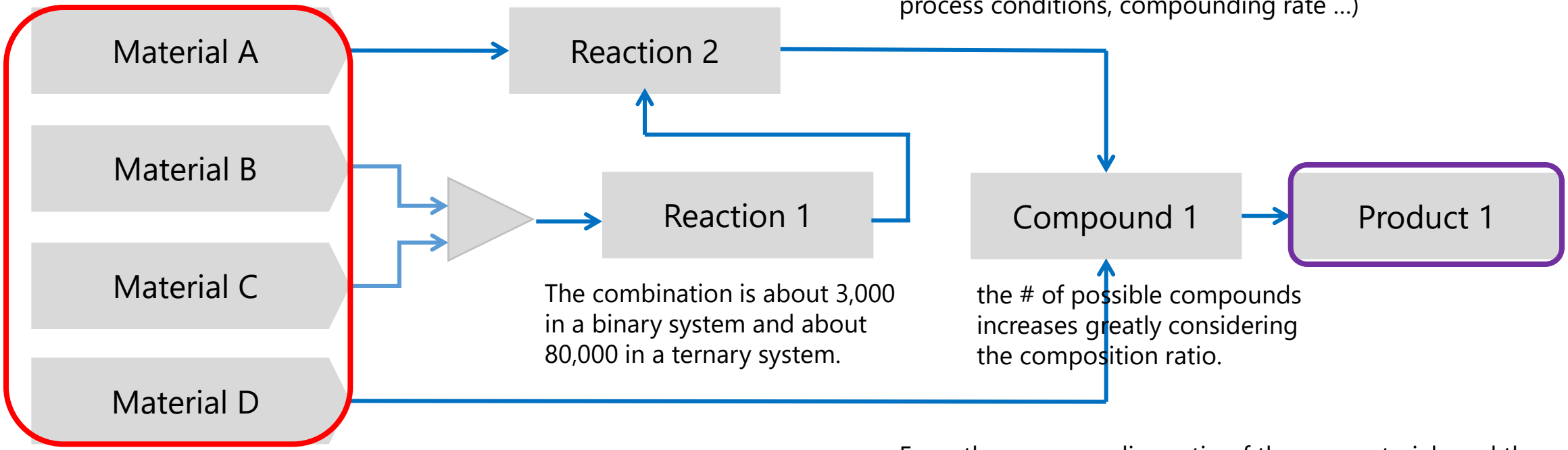
https://cbi-society.org/home/documents/seminar/2017to20/CBI391_Sawada.pdf

https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000154209.pdf

Comment 2: Is the effects of prediction limited ? (Material Informatics)

Conventional approach: Track what reaction happened

The # of elements handled in material science are about 80



The combination is about 3,000 in a binary system and about 80,000 in a ternary system.

the # of possible compounds increases greatly considering the composition ratio.

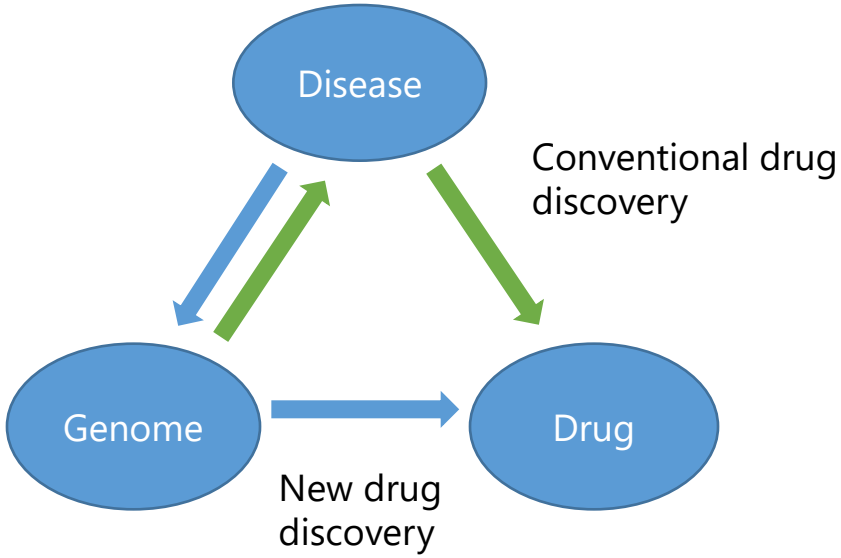
Machine learning approach: Find out what conditions generate a good product

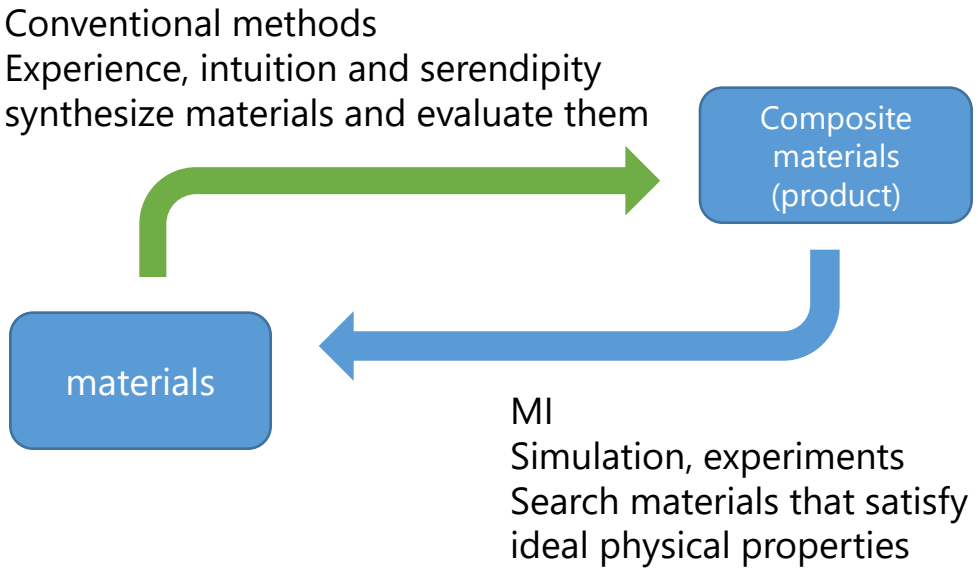
Estimate the physical properties and characteristics of a product by constructing equation model using reaction principles, physical properties of materials, and process control factors
Physical property of product=f (physical property of each material, process conditions, compounding rate …)

From the compounding ratio of the raw materials and the actual value of the physical property, an estimation equation that gives the highest explanatory power is calculated.
Physical property of product=f (compounding rate)

http://www.kri-inc.jp/tech/1269556_11451.html
http://cms.mtl.kyoto-u.ac.jp/informatics.html

Comment 3: Sequence of development has been changed by AI ?





Rather than aiming to *understand* the complete mechanism of disease, we will simply find drugs for the gene mutation from disease-specific genomic information

http://www.jpma.or.jp/opir/journal/journal_001.pdf

Shifting from the existing material development direction (*synthesizing* materials and *investigating* their physical properties) to the opposite direction (determining the material to be synthesized from the *required* physical properties)

https://www.mizuho-ir.co.jp/publication/giho/pdf/009_11.pdf

Comment 4: Closed loop and credit assignment problem*

In a *closed loop* of development, prediction-testing model can automate development process by using landscape fitting. In reality, state space *S* is open, and defining state space S properly is another challenge for innovators.

Even if the outcome is successful, there is still a serious problem: *credit assignment problem*** (Minsky, Pearl). Can we identify causes of success (or failure) without understanding mechanism of disease or structure of materials? How can *we* (not intelligence machine) conceive drug discovery theme or ideal materials without rich experiences or intuition? Can *we* step forward only with prediction machine? These concerns may have effects on social acceptance of AI driven innovation.

* I thank Professor Hiroyuki Chuma (Seijo University) for his valuable comment and discussion.

** The process of identifying among the set of actions chosen in an episode the ones which are responsible for the final outcome

Comment 5: Skill sets

To satisfy the demand for new tasks generated by AI, organizing interdisciplinary team or combining computational expertise with knowledge of specific scientific domains is important.

The next question is *how* we can choose an expertise or a scientific domain. What types of *competency* are needed for us to utilize prediction machines?

For example, to choose expertise at work in the near future, *meta recognition* must be needed. It is an unnecessary addition but my research team proposed a concept of “self-evolvability” as a driving force of individuals in digitalization era.

One example: Self-evolvability

Self-evolvability is defined as a mixed concept of meta-cognition, informating everything, and pursuing a shared benefit and desire of self-change. The questionnaires are as follows;

(*meta-recognition*) I'm trying to engage in more advanced and high-level work.
(*meta-recognition*) work while judging the possibility of the value of my skills and my job being lost.
(*benefit from knowledge sharing*)I share my knowledge and experience with others as much as possible.
(*desire to change)* I expect that my ideas will change.
(*informating everything*) I try to improve myself, having knowledge of my co-workers' work processes and outcomes

Effects of Self-evolvability on attitudes toward new technology

	More and more new technology should be incorporated in work				I want to use technology to enhance the efficiency of my work			
	Japan		U.S.		Japan		U.S.	
Self-evolvability	0.294	***	0.366	***	0.255	***	0.454	***
	(0.04)		(0.05)		(0.04)		(0.05)	

note: coefficients in upper row, standard error in lower row, *** p<0.001

# of obs. engineer 400s, sales manager 400s, teacher 400s in each country.

Dependent variables: disagree=1～agree=5

Control variables: age, education, occupations, personality traits and others.

KUME Koichi, Hiroyuki Chuma, Susumu Hayashi and Akihito Toda (2017) "Self-evolvability and Attitude toward Technological Changes: An empirical analysis using a survey", RIETI Discussion Paper Series 17-J-053.

Thank you for your kind attention

AI and drug discovery in Japan

Pharmaceutical companies	Overview
Astellas	Biovista：drug repositioning technology compound analysis by Clinical Outcome Search Space(COSS) Numedii：big data Intelligence technology to discover applicable disease of new compounds
Takeda	Numerate: in silico drug design technology to discover low molecular candidates for a new drug
Sumitomo Dainippon	Exscientia: automated bispecific-small-molecule design technology to design new compounds for a new drug.
Mitsubishi Tanabe	Introduce AI systems to drug discovery, Non-clinical experiment, clinical experiment in 2019 to increase probability of success of drug discovery and shorten the development period.
Shionogi	Semi-automation of analysis of clinical experiment by SAS machine learning

Impact estimation of introduction of AI : 4 years reduction of development period and 9 billion dollars cost reduction, # of approved new drugs increases from 5 to 20 products per year.

https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000154209.pdf

https://cbi-society.org/home/documents/seminar/2017to20/CBI391_Sawada.pdf

Material informatics in Japan

Matrials	Inventors	Algorithm
Electrolyte of Li-ion secondary battery	Fujitsu, RIKEN	Bayesian optimization
Dielectric material	TDK, Kyoto Univ.	data driven full searching
Thermoelectric material	NEC	Heterogeneous Mixture Learning, game tree search , combinatorial optimization
Magnetic material	Fujitsu Laboratories, Fujitsu	Genetic algorithm
Heat insulating material	NIMS	Ensemble regression tree
Power semiconductor	Nagoya Univ.	neural network
Organic EL light emitting material	Kyulux	neural network

Hitachi: developed electrodes for power semiconductors and reduced one-sixth of development period.
Asahi kasei: reduced development period of additives for polymer from 1-2 years to 8 months

https://www.nikkan.co.jp/releases/view/37806

https://tech.nikkeibp.co.jp/atcl/nxt/mag/ne/18/00030/00001/