

인공지능을 활용한 규제샌드박스 개선 : 규제신속확인을 중심으로

황하 부연구위원

한국행정연구원, 사회통합재난안전연구실

1

혁신적 기술의 등장

- 많은 국가들이 삶의 질 향상을 위해 혁신 기술을 도입하는 추세
- 혁신 기술의 상용화를 위해서는 신뢰가 전제되어야 하고, 이를 위해 안전성 확보가 필수

기술규제에서 기술을 통한 규제(RegTech)로

금융 기관 :

- 2015년 까지 금융회사가 납부한 Post-GFC 과징금 2,000억 달러 초과 (CNBC)
- 금융규제가 복잡해지면서 규제준수 비용이 크게 증가

감독 기관:

- 신기술을 활용한 금융상품 출시가 급격히 늘어나면서, 이에 대한 규제 당국의 이해도가 저하

-> 정보의 비대칭성으로 인해 금융기관과 감독기관 모두 신기술 활용을 통해 규제준수와 모니터링을 용이하게 할 인센티브가 발생

2

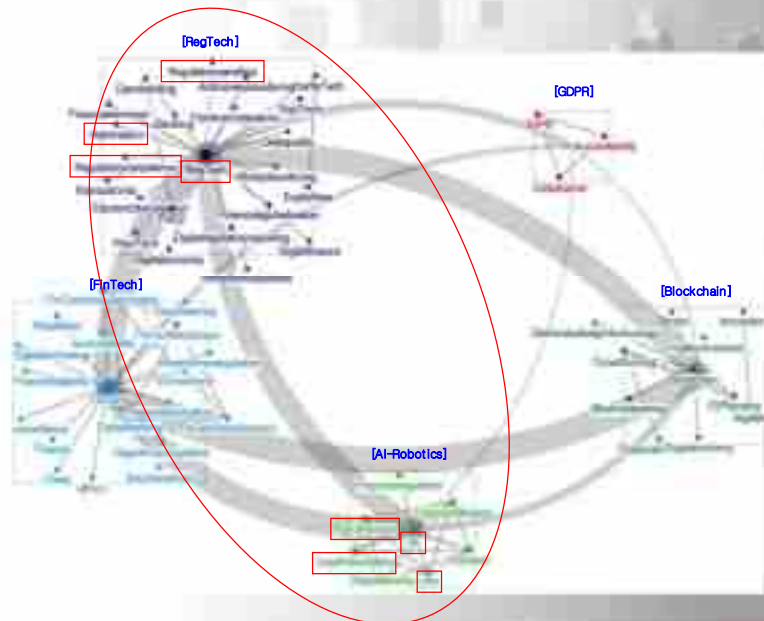
레그테크란?

신기술 사용을 통해 규제 효율성 및 효과성을 향상시키는 것을 총칭

- 기업입장에서는 규제순응을 보다 쉽게 만들
- 감독기관 입장에서는 기업의 규제순응과 위험관리를 모니터링
- 기타 등등...

3

레그테크 관련 키워드 네트워크 분석



4

한국형 규제 샌드박스



5

규제란?

광의,
분산적
개념

Steering the flow of events (Braithwaite and Parker)

사전에 설정한 목표를 달성하기 위해 위험을 관리하거나 행위를 수정/변경하려는 의도적인 시도 (Black, Hood)

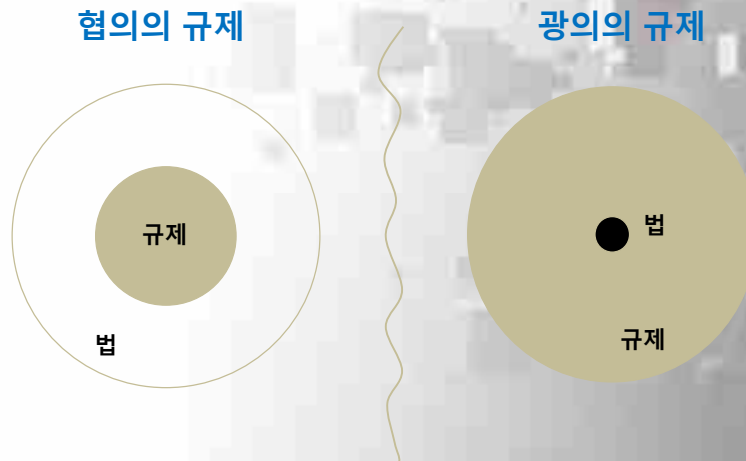
협의,
국가중심
개념

정부에 의한 규칙 부여 (OECD)

국가나 지방자치단체가 특정한 행정 목적을 실현하기 위하여 국민(국내법을 적용받는 외국인을 포함한다)의 권리를 제한하거나 의무를 부과하는 것으로서 법령등이나 조례·규칙에 규정되는 사항(행정규제기본법 제2조)

6

규제와 법의 관계



7

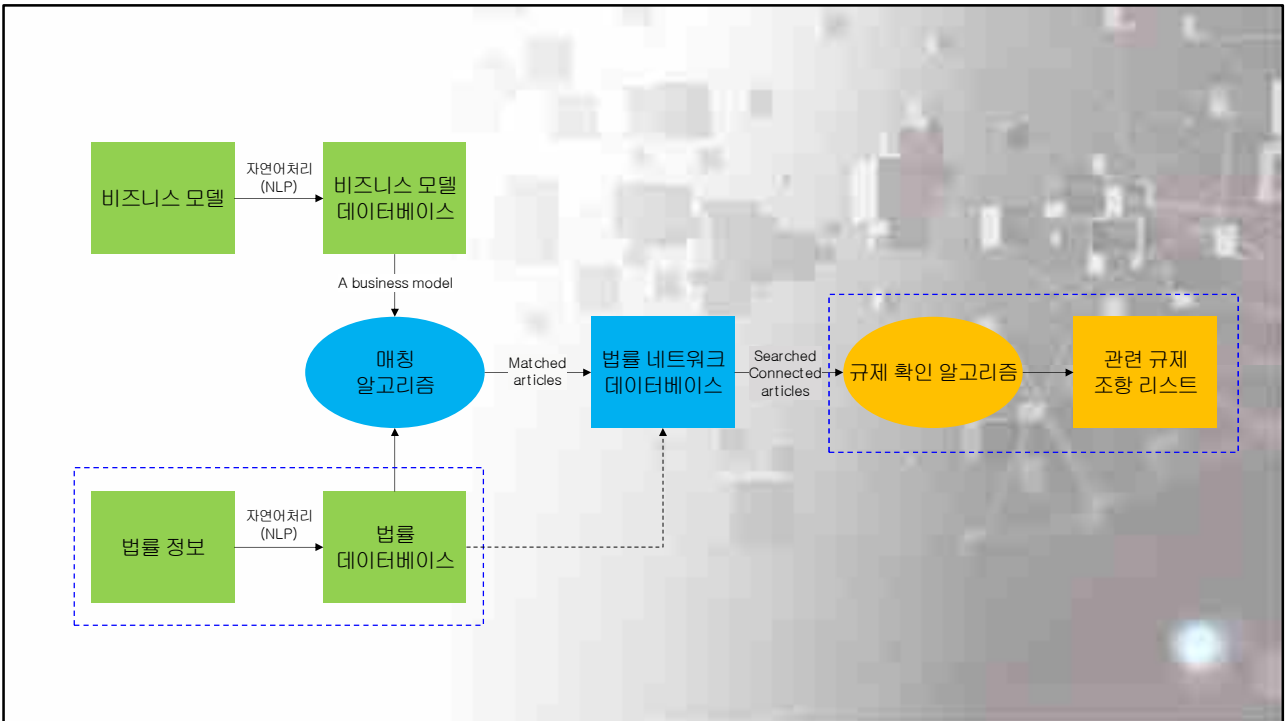
규제등록제도

- 정부는 규제관리를 위해 규제의 숫자를 세기 시작
- 조문 단위로 규제를 등록
- 규제등록 담당자가 특정 조문이 규제인지 여부를 판단
- 규제를 구분하는 범정부적 가이드라인이 존재하지만, 애매한 경우 규제로 등록하도록 권고하고 있음

8



9



10

규제 샌드박스에서 가장 많이 언급되는 스마트 교통(Smart Mobility), 바이오헬스 (Bio-health), 핀테크(FinTech)와 관련된 11개 법률 데이터베이스 구축

분류	법 이름	조항 수		
		규제	비규제	합계
스마트 교통 (Smart Mobility)	도로교통법	77	121	198
	여객자동차운수사업법	75	60	135
	자동차관리법	77	89	166
	택시운송사업의 발전에 관한 법률	24	0	24
바이오헬스 (Bio-health)	생명윤리 및 안전에 관한 법률	48	27	75
	의료기기법	40	45	85
	의료법	63	73	136
핀테크 (FinTech)	보험업법	106	129	235
	외국환거래법	19	18	37
	자본시장과 금융투자업에 관한 법률	357	205	562
	전자금융거래법	37	25	62
Total		923	792	1715

11

언어모델 구축을 위한 전처리



12

규제확인 언어모델

keywords dtm ($p * k$)					keywords weighted matrix ($k * k$)					weighted keywords dtm ($p * k$)						
	word 1	word 2	...	word k		word 1	word 2	...	word k		word 1	word 2	...	word k		
Article 1	0	0	...	1	×	word1	1	$w_{1,2}$...	$w_{1,k}$	=	Article 1	$\hat{w}_{p,1,1}$	$\hat{w}_{p,1,2}$...	$\hat{w}_{p,1,k}$
Article 2	1	0	...	0		word 2	$w_{2,1}$	1	...	$w_{2,k}$		Article 2	$\hat{w}_{p,2,1}$	$\hat{w}_{p,2,2}$...	$\hat{w}_{p,2,k}$
Article 3	1	1	...	0		word 3	$w_{3,1}$	$w_{3,2}$...	$w_{3,k}$		Article 3	$\hat{w}_{p,3,1}$	$\hat{w}_{p,3,2}$...	$\hat{w}_{p,3,k}$
Article 4	0	1	...	1		word 4	$w_{4,1}$	$w_{4,2}$...	$w_{4,k}$		Article 4	$\hat{w}_{p,4,1}$	$\hat{w}_{p,4,2}$...	$\hat{w}_{p,4,k}$
Article 5	0	0	...	1		word 5	$w_{5,1}$	$w_{5,2}$...	$w_{5,k}$		Article 5	$\hat{w}_{p,5,1}$	$\hat{w}_{p,5,2}$...	$\hat{w}_{p,5,k}$
...
Article p	1	1	...	0	word k	$w_{k,1}$	$w_{k,2}$...	1	Article p	$\hat{w}_{p,p,1}$	$\hat{w}_{p,p,2}$...	$\hat{w}_{p,p,k}$		

13

weighted keywords dtm ($p * k$)					Reg. label ($p * 1$)	Reg. Language Modeling Frame ($p * (k + 1)$)							
	word 1	word 2	...	word k	Reg		word 1	word 2	...	word k	Reg		
Article 1	$\hat{w}_{p,1,1}$	$\hat{w}_{p,1,2}$...	$\hat{w}_{p,1,k}$	+	0	=	Article 1	$\hat{w}_{p,1,1}$	$\hat{w}_{p,1,2}$...	$\hat{w}_{p,1,k}$	0
Article 2	$\hat{w}_{p,2,1}$	$\hat{w}_{p,2,2}$...	$\hat{w}_{p,2,k}$		0		Article 2	$\hat{w}_{p,2,1}$	$\hat{w}_{p,2,2}$...	$\hat{w}_{p,2,k}$	0
Article 3	$\hat{w}_{p,3,1}$	$\hat{w}_{p,3,2}$...	$\hat{w}_{p,3,k}$		1		Article 3	$\hat{w}_{p,3,1}$	$\hat{w}_{p,3,2}$...	$\hat{w}_{p,3,k}$	1
Article 4	$\hat{w}_{p,4,1}$	$\hat{w}_{p,4,2}$...	$\hat{w}_{p,4,k}$		0		Article 4	$\hat{w}_{p,4,1}$	$\hat{w}_{p,4,2}$...	$\hat{w}_{p,4,k}$	0
Article 5	$\hat{w}_{p,5,1}$	$\hat{w}_{p,5,2}$...	$\hat{w}_{p,5,k}$		0		Article 5	$\hat{w}_{p,5,1}$	$\hat{w}_{p,5,2}$...	$\hat{w}_{p,5,k}$	0
...
Article p	$\hat{w}_{p,p,1}$	$\hat{w}_{p,p,2}$...	$\hat{w}_{p,p,k}$	1	Article p	$\hat{w}_{p,p,1}$	$\hat{w}_{p,p,2}$...	$\hat{w}_{p,p,k}$	1		

14

모델 결과

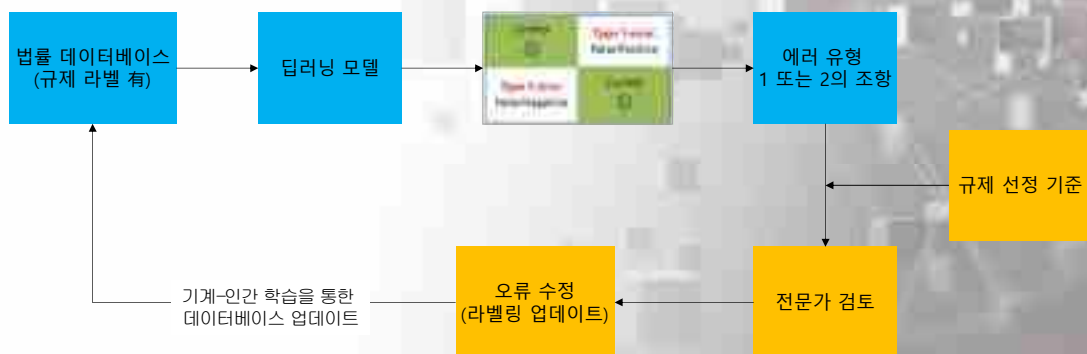
Auto ML (Stacked Ensemble)							
confusion matrix on validation data						Metrics	
Predicted \ Actual	0	1	Error rate				
0	151	94	0.3837	(94.0/245.0)	AUC	0.8099	
1	37	232	0.1375	(37.0/269.0)	Recall	0.8625	
Total	188	326	0.2549	(131.0/514.0)	Precision	0.7117	

Deep Learning (Neural Network)							
confusion matrix on validation data						Metrics	
Predicted \ Actual	0	1	Error rate				
0	137	108	0.4408	(108.0/245.0)	AUC	0.8034	
1	28	241	0.1041	(28.0/269.0)	Recall	0.8959	
Total	165	349	0.2646	(136.0/514.0)	Precision	0.6905	

Generalized Linear Model							
confusion matrix on validation data						Metrics	
Predicted \ Actual	0	1	Error rate				
0	144	101	0.4122	(101.0/245.0)	AUC	0.7993	
1	32	237	0.119	(32.0/269.0)	Recall	0.8810	
Total	176	338	0.2588	(133.0/514.0)	Precision	0.7012	

15

인간의 학습 보조를 통한 고도화



16

베타시스템 운영을 통한 고도화

- 3차년도 최종 목표는 규제신속확인 앱의 베타 시스템 출시
- 이후 베타 시스템 운영을 위한 후속 사업 발굴 계획 예정



17

기대효과

정부의 입장:

- 규제신속확인제도의 자동화를 통해 규제샌드박스 고도화
- 규제확인을 위한 행정비용 절감(3차년도 시범 분석 예정, 후속 연구 계획 중)
- 규제개혁 만족도 및 체감도 향상(베타서비스 운영을 통한 시범분석 예정, 후속 연구 계획 중)

기업 및 시민의 입장:

- 규제신속확인 앱 사용을 통해 혁신서비스 출시를 위한 기업의 비용과 시간 절약 (거시경제모형 등을 활용한 Impact Analysis를 위한 후속 연구 구상 중)
- 규제에 대한 기업 및 대중의 지식 향상(기업이 비즈니스 모델 개발 시 참고)

18

