

# A Big Data a jogtudományban és a jogalkalmazásban

Zódi Zsolt

Egyetemi docens,  
Infokommunikációs tanszék



# Big Data - a jelenlegi jogi szabályozás szemszögéből

## Személyes-e az adat, vagy sem?

- Itt a szabályozás élesen elválnak, mert míg a személyes adatok esetén igen szigorú szabályok vannak,
  - főszabály szerint az érintett hozzájárulása, vagy jogszabályi felhatalmazás kell a gyűjtésükhöz, tárolásukhoz, és az érintett hozzájárulásával sem lehet eltérő célra használni és másnak engedély nélkül átadni őket, addig
- Ha nem személyes az adat a jogot **lényegében nem érdekli**. Mindenki olyan adatot gyűjt és tárol, amelyet csak akar

# Lehetőségek

- Új szolgáltatásokat tesz lehetővé, munkahelyeket teremt
- Automatizált döntéshozatal és következtések
- Predikciót tesz lehetővé
- Egyediesítést, testre szabott szolgáltatásokat tesz lehetővé
- Alapja lesz, lehet a robotikának, a mesterséges intelligenciáknak
- A teljesítmények jobb mérését teszi lehetővé (pl. olyan területeken is, ahol ez eddig nehézkes volt - közigazgatás, közszolgáltatások)
- Kisebb mintákon lehetséges a kísérletezés, a szimuláció és a visszamérése pl. intézkedéseknek
- A policyk (pl. új jogszabályok) kialakításánál hatásvizsgálatokra lehet használni.

# A BD-ben rejlő lehetőségek: a BD alapú eszközök, mint a jogalkotást támogató eszköz

ISSN 2207-1130 (Online)  
<http://www.jksii.or.kr>

## Designing a Crime-Prevention System by Converging Big Data and IoT

Jin-ho Jeon<sup>1</sup>


Seung-Ryul Jeong<sup>1\*</sup>

Articles

## Big data in the policy cycle: Policy decision making in the digital era

Johann Höchtel, Peter Parycek  & Ralph Schöllhammer

Pages 147-169 | Accepted author version posted online: 02 Dec 2015, Published online: 02 Dec 2015

 Download citation  <http://dx.doi.org/10.1080/10919392.2015.1125187>  Crossmark

„Hidden patterns and correlations will be identified to support common-sense experience or received wisdom. Predictive analytics applied on top will increase the quality of scenario planning and result in true evidence-based policymaking.”

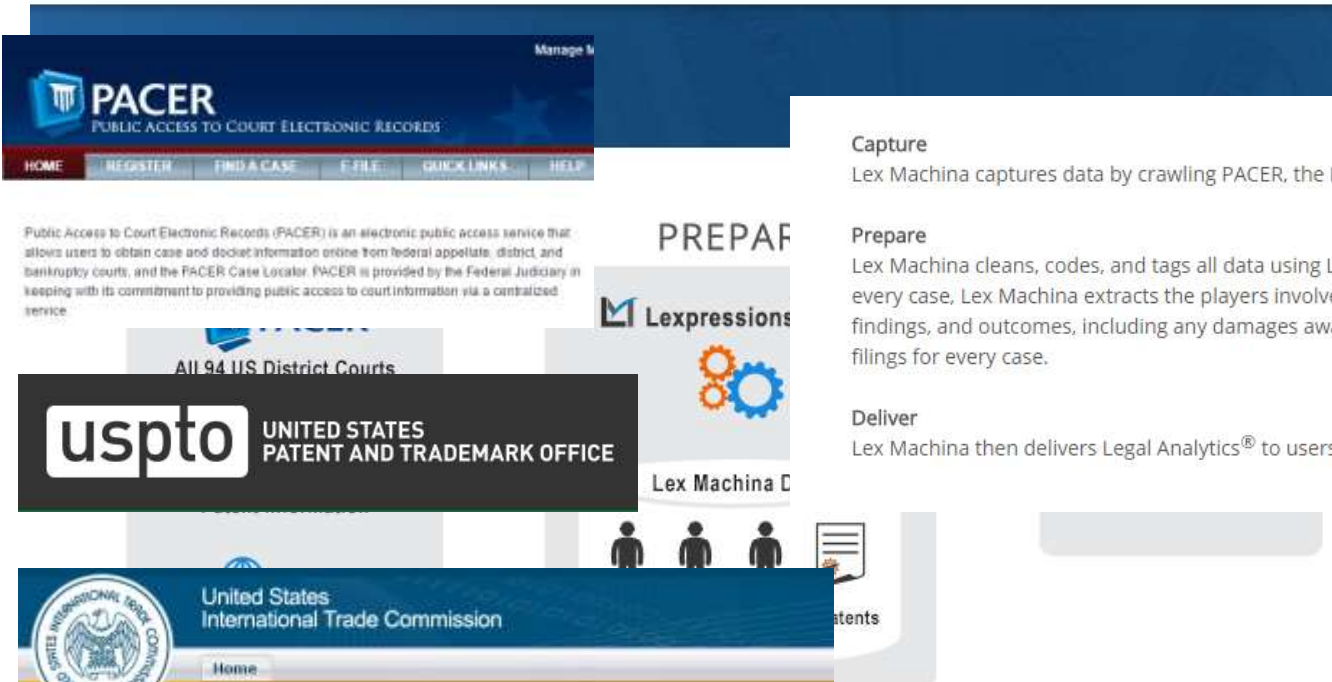
Journal of Organizational Computing and Electronic Commerce

Volume 26, 2016 - Issue 1-2:  
Open Data: Value Creation for Policy Makers and Analysts

# A BD, mint a jogalkalmazást támogató eszköz: Lex Machina, Ravel Law



What We Do Product Law Firms Companies



## Capture

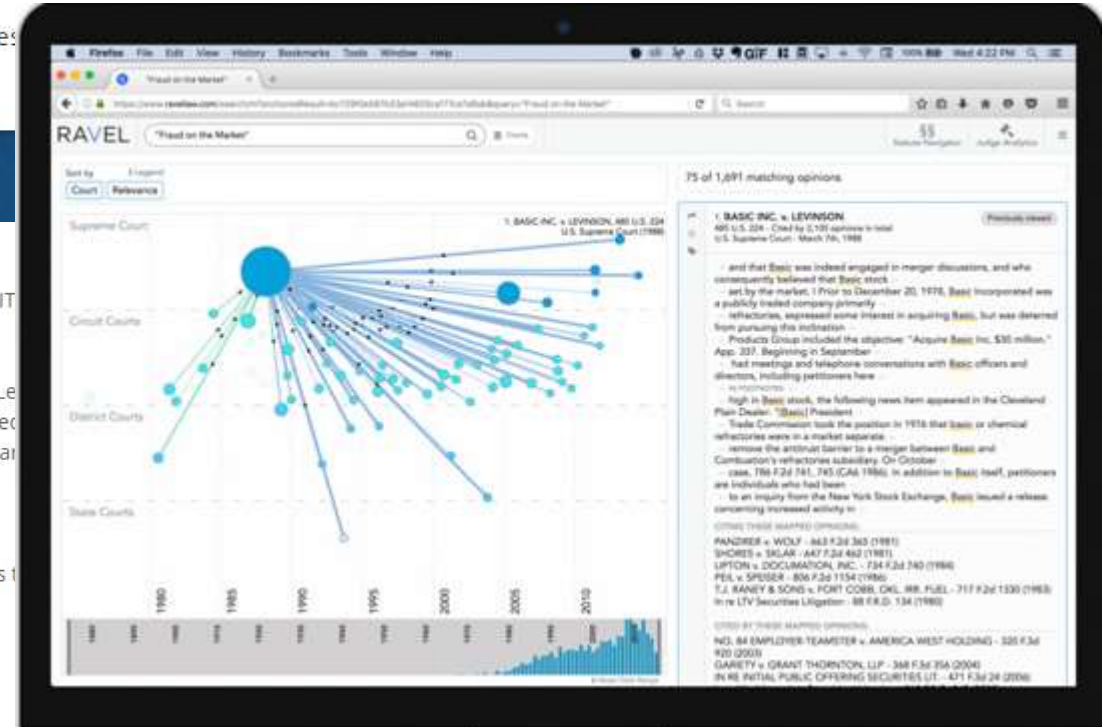
Lex Machina captures data by crawling PACER, the IT

## Prepare

Lex Machina cleans, codes, and tags all data using Lex every case. Lex Machina extracts the players involved findings, and outcomes, including any damages awarded filings for every case.

## Deliver

Lex Machina then delivers Legal Analytics® to users t



<https://lexmachina.com/legal-analytics/>

<http://ravellaw.com/products/#eluid7fcb4be2>

Electronic Document Information System (EDIS)

Welcome to EDIS. You are currently not logged in.

EDIS is the repository for all documents filed in relation to an investigation conducted by the United States International Trade Commission (USITC). EDIS provides the capability to file documents for an investigation as well as search for documents which have been submitted to the USITC.

# A BD, mint a jogtudományt segítő eszköz

- Módszertani viták vannak azzal kapcsolatban, hogy mi is az „igazi jogtudomány”
  - Az ún. „black letter”, vagy „dogmatikai” tudomány - a jogi rendelkezések közötti értelem-összefüggések és fogalmak vizsgálata, rendszerbe foglalása
- VAGY
- A „társadalomtudományi” jogtudomány: empirikus, szociológiai módszerek használata. A jognak, mint egy társadalmi jelenségnek, alrendszernek, szabályrendszer *a sok közül*.
- Az „adatalapú” jogtudomány hídakat képes verni a kétféle álláspont között - jogi dokumentumokat, szövegeket elemez, de empirikus, tartalomelemzési, szövegbányászati módszerekkel.
- Milyen mintázatokat mutat egy kifejezés használata, egy fogalom előfordulása, egy érvelési sablon, stb.

# A BD alapú jogtudomány

	Módszer, használt fogalmi háló	Vizsgált sokaság	Vizsgált tárgy	Lehetséges-e a következtetés és a predikció?
Hagyományos dogmatikai	Desktop, létező (normatív) fogalmi hálókat használ	Autoritások által determinált, „fontos”	Normatív jelenség megjelenései - szövegek	Nem
Empirikus	Társadalom-tudományi (jogon kívüli) kategóriák	Esetleges, véletlenszerű	Társadalmi jelenség megjelenései, és/vagy szövegek	Nem
Adataalapú	Társadalom-tudományi (jogon kívüli) kategóriák	Esetleges, véletlenszerű	Társadalmi jelenségről felvett adatok, és/vagy szövegekről felvett adatok	Nem
Statisztikai	Társadalom-tudományi (jogon kívüli) kategóriák	Reprezentatív	Társadalmi jelenségről felvett reprezentatív adatok, és/vagy szövegekről felvett reprezentatív adatok.	Részben
BD alapú	Utólagos, előre megjósolhatatlan narratívák	Teljes	Adathalmazok	Igen

# BD alapú döntések a jogban

- BD alapú döntések már ma is vannak a jog területén kívül: üzleti élet, marketing-kampányok, banki döntések (scoring) stb.
- Mérlegelési kategóriák pontosabb „belövése”: ésszerű idő, méltányos haszonérmérséklés
- Büntetéskiszabás - pontosabb büntetéskiszabás
- Lehetséges következmények:
  - Target variability: a célnak megfelelő mintázatokat csak akkor tudom kiválasztani, ha pontosan definiálom a célokat - pl. büntetéskiszabás - speciális vagy generális prevenció?
  - Széttartó gyakorlatok - jogegység (formális igazságosság) helyett („hasonló eseteket hasonlóképp”) minden esetet egyediként - a gép sokkal több változóval tud dolgozni, mint az ember.
  - Indokohatatlanság - egy statisztikai alapon hozott gépi döntést hogyan lehet megindokolni? (Indokolás = morális okozatosságon alapuló indokolás - nem vagyok felelőssé tehető olyan dolgokért, amelyekről nem tehetek.)



# Az EU és a Big Data

- Az új adatvédelmi rendeletet még a Big Data korszak előtt kezdték megszövegezni...
- A rendelet célja: egyszerre fokozni a személyes-adat védelmi szintet, egységesíteni az EU-s szabályozást és csökkenteni az adminisztratív terheket
- Ez nem megy egyszerre, ezek szembevető követelmények
- A R. azzal a trükkös érveléssel oldja ezt fel, hogy a magasabb adatvédelmi szint a magasabb vevői bizalom-szint miatt versenyelőnyt fog biztosítani az európai vállalkozásoknak. Ezt SEMMI nem bizonyítja, sőt, az adatvédelmi aggályok miatt bizonyos vállalkozások eleve létre sem jöhetnek
- A R. nem említi a Big Datát, csak két ponton, közvetve: az „automatizált döntéshozatalnál”, és a „profilalkotásnál”. Az EU mindkettőt kifejezetten rossz kontextusba helyezi
- Az automatikus döntéshozatal alól az érintett

# Az EU és a Big Data II.

- „Profilalkotás”: személyes adatok automatizált kezelésének bármely olyan formája, amelynek során a személyes adatokat valamely természetes személyhez fűződő bizonyos személyes jellemzők értékelésére, különösen a munkahelyi teljesítményhez, gazdasági helyzetéhez, egészségi állapothoz, személyes preferenciákhoz, érdeklődéshez, megbízhatósághoz, viselkedéshez, tartózkodási helyhez vagy mozgáshoz kapcsolódó jellemzők elemzésére vagy előrejelzésére használják
- Automatizált döntéshozatal. Nincsen definíciója, kizárólag gépi eszközökkel, korábban gyűjtött adatokon alapuló döntéshozatal. „Az érintett jogosult arra, hogy ne terjedjen ki rá az olyan, kizárólag automatizált adatkezelésen - ideértve a profilalkotást is - alapuló döntés hatálya, amely rá nézve joghatással járna vagy őt hasonlóképpen jelentős mértékben érintené.” kivétel: az érintett és az adatkezelő közötti szerződés megkötése vagy teljesítése érdekében szükséges; - ez egy gumiszabály lesz, már most látható.

# Problémák, veszélyek

Buried in a list of 20 released on the Intecompany to protect No. 4417749 conduct topics

on eve  
No. 44  
in Lib  
shado  
tigin  
lives in  
a repo  
three  
spid  
ri

<http://www.nytimes.com/2012/02/19/magazine/shopping>

FEB 16, 2012 @ 11:02 AM 3,073,794 VIEWS

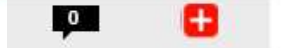
## How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did



The commission's study of a cross-section of nine data brokers revealed that data brokers aggregate online and offline data from disparate sources to make inferences about consumers' ethnicity, income, religion, age, and health conditions, among other characteristics. As the FTC and others have found, some brokers create segments or clusters of consumers with high concentrations of minorities or low income individuals.

How Netflix is turning viewers  
The New York Times Magazine

The Little Black Book of Billionaire



BY, MOVIES, NETFLIX,

Big Data



[x is turning](#)

# Kockázatok

- Kockázatok, amelyeket a Big Data megnövelt, és speciális új kockázatok
- Parttalan adatgyűjtés - Citizen Score stb, nagyobb csábítás
- Az érintettek nem képesek követni a szabályokat - még kevésbé
- Profilalkotás - pontosabb, mélyebb
- Automatikus döntéshozatal - még elmagyarázhatatlanabb szabályok. Még kisebb esély a kijelentkezésre
- Visszaszemélyesítés lehetősége
- „Predikciós hatalom” - „információs túlhatalom” - „azt is tudják, amit mi magunk sem tudunk magunkról”
- Diszkriminatív gyakorlatokat erősíthet meg - a „rossz vevők” között több a kisebbségi, a „jó vevők” között a többségi
- Manipulációs potenciál

# BD etika

- Jogi szabályozás még nincsen, de egyesek már arról beszélnek, milyen elvek mentén kellene szabályozni
- A magánszemély képtelen a személyes adatai feletti kontrollt gyakorolni. > a szigorúan beleegyezés-alapú szabályozást fel kellene oldani. Ha megadja sem lehessen mindent, ha nem adja meg is lehessen bizonyos dolgokat.
- A középpontban az „önmeghatározás szabadsága” kell álljon, és nem az, hogy senki semmilyen adatot nem kezelhet, ha nem egyeztem bele.
- A „privacy” binaritását fel kell oldani - pl. az, hogy 1. más tud az adatomról 2. Ez alapján döntést hoz, 3. nyilvánosságra hozza nagyon különböző konzekvenciákkal jár rám nézve.
- Eleve az adatvédelem technikai szabályai helyett a „privacy” védelme?
- Két szélsőséges eset: a privacy nem sérül, ha valaki kezeli az adatot, de annak nincsen rám semmilyen hatása vs. ha hozzájárultam sem lehet mindent