Nortal

DeepScan

Step towards GDPR compliance

Lauri Ilison
Head of Big Data and Machine Learning
2.11.2017

The Toughest Data Protection Law. Ever.

GDPR
EU General Data Protection Regulation

New principles

- Right to Be Forgotten

- Data Portability

- Data Protected by Default

- Mandatory Breach Notification

- ....

Fines are up to €20 million

4% 4% of global annual turnover

146 working days left

MAY 25, 2016

The world's most stringent data protection law enters into force in the EU

MAY 25, 2018

Timeline

All companies must comply under the threat of sizeable fines for leakage and non-compliance

Are governments ready?

Live test

PSII

Search for personal data

Search

Or use our dashboard

**Garage48 Baltic Open
& Big Data 2017**

SAVE THE DATE
19- 21 May 2017

VENUE
LIFT99, Telliskivi 60a Tallinn

Original document

Previous Next Page: /

KORRALDUS

Hoolduse seadmine ja hooldajatoetuse määramine

võttes aluseks Sotsiaalhoolekande seaduse paragrahvi 26 lõike 1 punkti 7, mille alusel kohalik omavalitsus seab puuetega isikutele hoolduse, märtsi 2005 määruse nr 12 „Hooldajatoetuse maksmise kord paragrahvi 4, mille kohaselt vallavalitsusel on isiku kirjaliku avalduse alusel õigus määrata või määramata jätta hooldajatoetus, läbi vaadanud annab järgmise

korralduse:

1. Määrata raske puudega hooldajaks mõlema elukoht

3. Korraldust on võimalik vaidlustada päeva jooksul korralduse

---

Metadata

Registry:
Topic: Sotsiaalhoolekanne
Date:
Title: *Hooduse seadmine ja hooldajatoetuse määramine*
[Download link]
Original filename:
Document number:

Extracted texts

Garage48 Baltic Open & Big Data 2017 Winners

May 21, 2017

#bigdataopendata2017
#TELE2EESTI

SHARE ON FACEBOOK

Latitude Pitching – early stage winner!

Palju õnne!

PSII (Personal Sensitive Identifiable Information) võitis eGovernance Solutions võistluse! Nende platvorm leiab avalikest registritest personaalset tundlikku informatsiooni.

Võitjameeskonda premeeriti 5000 euroga, et ideed edasi arendada ja eksperte kaasata. Lisaks võimaldatakse neile 2. päevane mentortugi Eesti e-Riigi tippekspertide poolt.... See More

See Translation

— 🤨 feeling accomplished at 📍 KultuuriKatel.

👍 Like   💬 Comment   ↪ Share

👍 8
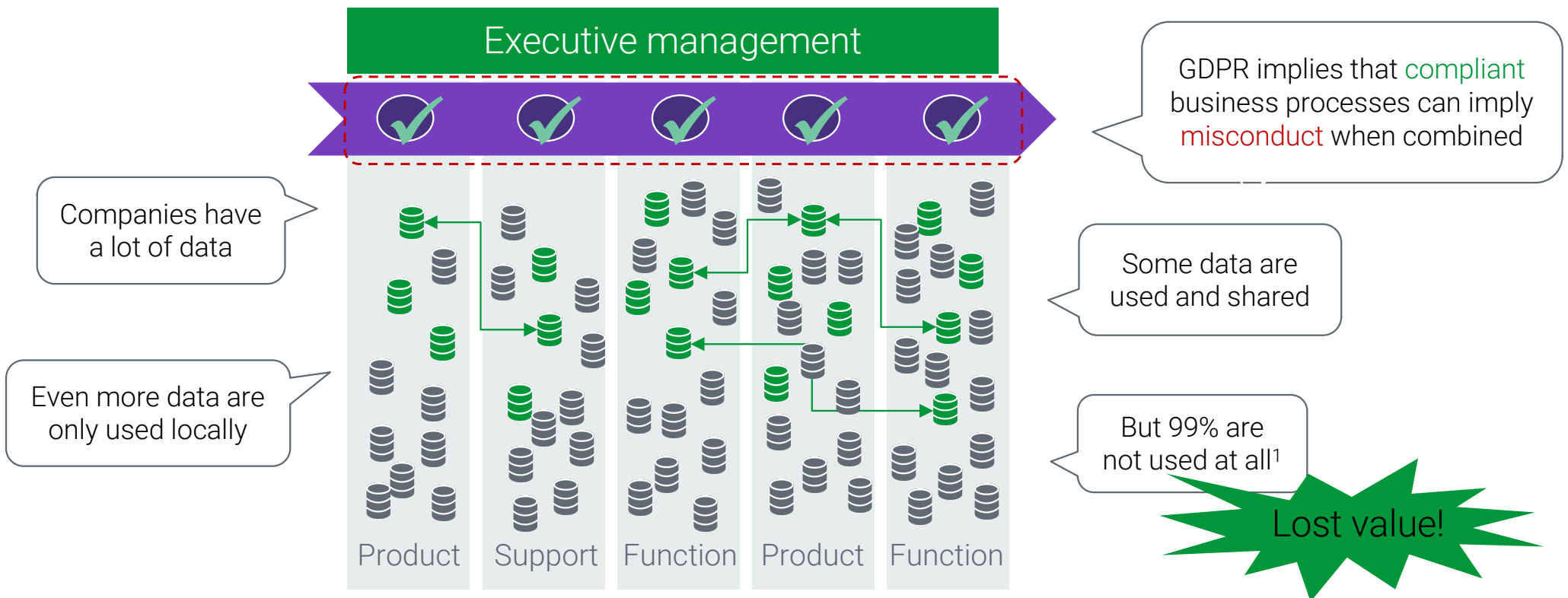
Write a comment...   ☺ 📷 GIF 🦊

Press Enter to post.

DeepScan

Platform that helps the governments and enterprises to know the data they have.
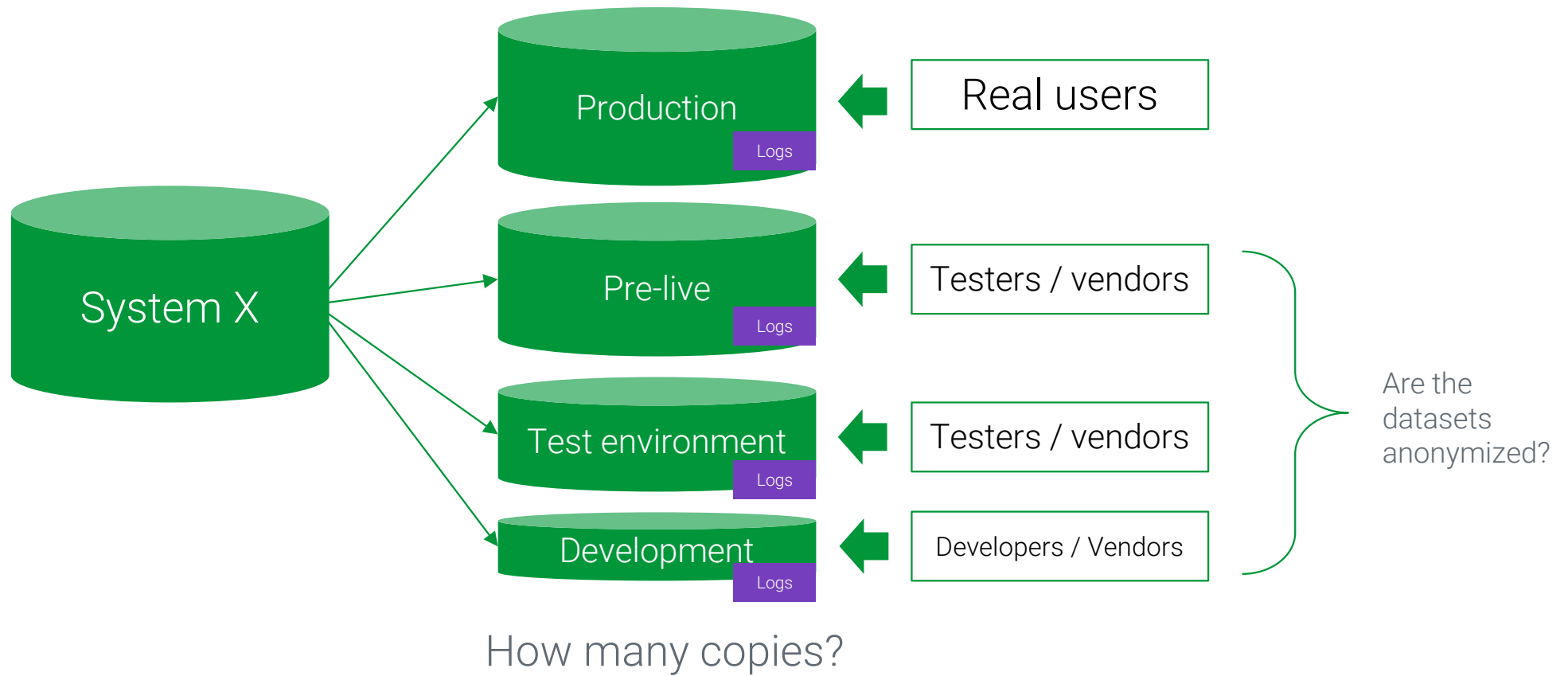
To **protect** the data you need to know the **data** you **have**

Only a fraction of companies' data are known, and even less are put to effective use – implying lost value for the owners

Executive management

Companies have a lot of data

Even more data are only used locally

Product   Support   Function   Product   Function

GDPR implies that compliant business processes can imply misconduct when combined

Some data are used and shared

But 99% are not used at all[1]

Lost value!

Note: 1) According to VentureNet, only about 28% of the available data have any value to it
Source: NG data/ Lucidworks

Why this is important?

System X

Production — Logs ← Real users

Pre-live — Logs ← Testers / vendors

Test environment — Logs ← Testers / vendors

Development — Logs ← Developers / Vendors

Are the datasets anonymized?

How many copies?

The solution

Personal information discovery and classification
with *DeepScan* in any data-source

- Scan the data and automatically identify personal information

- Build holistic, truth-based *datamap* from scattered data

- Keep *datamap* up-to-date by running scans periodically

Technical setup
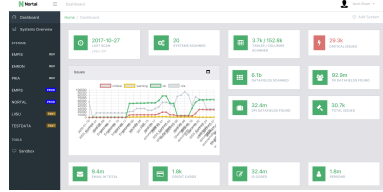
Datasources
- ORACLE
- Microsoft SQL Server
- MySQL
- PostgreSQL
- IBM DB2
- Exchange
- SharePoint
- Word, Excel, PowerPoint, PDF Adobe
- Files
- Logfiles

Gateway
- nifi

Results

Scan + Artificial Intelligence
- Apache Spark

Preprocessing
- Apache Spark

Storage
- Apache HBASE

Storage
- Hadoop HDFS

Nortal

☰ Dashboard

👤 lauri.ilison ▾

⊕ Add System

- 🕐 Dashboard
- 📊 Systems Overview

SYSTEMS

EMPIS	DEV
ENRON	DEV
PRIA	DEV
EMPIS	PROD
NORTAL	PROD
LIISU	TEST
TESTDATA	TEST

TOOLS

- 🖥 Sandbox

2017-10-27
LAST SCAN
LIISU-ZIP

20
SYSTEMS SCANNED

3.7k / 152.8k
TABLES / COLUMNS SCANNED

29.3k
CRITICAL ISSUES

Issues

🔲

6.1b
DATAFIELDS SCANNED

92.9m
PII DATAFIELDS FOUND

32.4m
SPI DATAFIELDS FOUND

30.7k
TOTAL ISSUES

Legend: ■ critical ■ warning ■ ok ■ n/a

Y-axis: 100000, 90000, 80000, 70000, 60000, 50000, 40000, 30000, 20000, 10000, 0
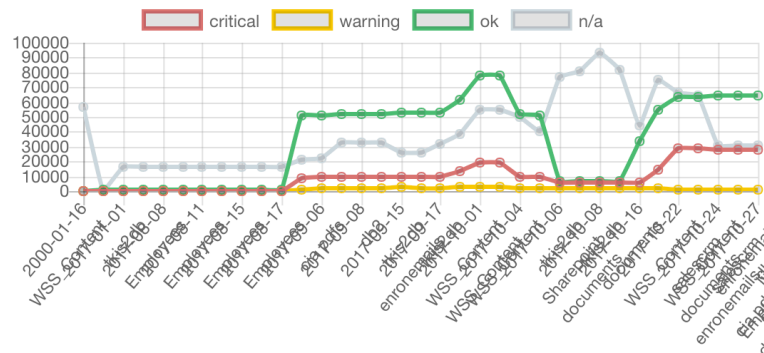
9.4m
EMAIL IN TOTAL

1.8k
CREDIT CARDS

32.4m
ID CODES

1.8m
PERSONS

Nortal

Nortal ☰ Dashboard

👤 lauri.ilison ▾

	CRITICAL	WARNING	OK	N/A	CREDIT CARD	BIRTH DATE	EMAIL	GENDER	PERSON	PERSONAL CODE	PHONE NUMBER	POSTAL ADDRESS
	⬤OFF	⬤OFF	⬤OFF	OFF	💳 OFF	📅 OFF	✉ OFF	⚥ OFF	👤 OFF	🪪 OFF	📞 OFF	📍 OFF

- ⊙ Dashboard
- ᴵᴵᴵ Systems Overview

SYSTEMS

- EMPIS — DEV
- ENRON — DEV
- 💗 ENRONEMAILS
- PRIA — DEV
- EMPIS — PROD
- NORTAL — PROD
- LIISU — TEST
- TESTDATA — TEST

TOOLS

- 🖥 Sandbox

👎	emails Items: 7 Rows: 8936	Detected data	💳 Credit card	📅 Birth Date	✉ Email	⚥ Gender	👤 Person	🪪 Personal code	📞 Phone number	📍 Postal address
✉	**bcc** text, min: 16, avg: 109, max: 2,470	email	0 / **0%**	0 / **0%**	4783 / **100%**	0 / **0%**	0 / **0%**	0 / **0%**	0 / **0%**	0 / **0%**
🔥	**body** text, min: 6, avg: 1,688, max: 10,240	mixed	4 / **0.03%**	0 / **0%**	5963 / **43.14%**	967 / **7%**	12124 / **87.72%**	0 / **0%**	1311 / **9.48%**	1 / **0.01%**
✉	**cc** text, min: 16, avg: 109, max: 2,470	email	0 / **0%**	0 / **0%**	4783 / **100%**	0 / **0%**	0 / **0%**	0 / **0%**	0 / **0%**	0 / **0%**
✉	**from** text, min: 15, avg: 26, max: 62	email	0 / **0%**	0 / **0%**	17993 / **100%**	0 / **0%**	0 / **0%**	0 / **0%**	3 / **0.02%**	0 / **0%**
👤	**attachment_count** text, min: 28, avg: 28, max: 29	person	0 / **0%**	0 / **0%**	0 / **0%**	0 / **0%**	1802 / **100%**	0 / **0%**	0 / **0%**	0 / **0%**
🔥	**subject** text, min: 6, avg: 42, max: 193	mixed	0 / **0%**	0 / **0%**	95 / **4.51%**	73 / **3.46%**	1095 / **51.94%**	0 / **0%**	24 / **1.14%**	0 / **0%**
✉	**to** text, min: 14, avg: 165, max: 24,160	email	0 / **0%**	0 / **0%**	17258 / **100%**	0 / **0%**	0 / **0%**	0 / **0%**	2 / **0.01%**	0 / **0%**

GDPR is more than compliance investment

Being smart ... you can turn GDPR from legal-driven costly compliance initiative into a seamless enterprise-wide better customer service initiative

- Compliance investments do not get postponed, even in hard times

- GDPR-investment can be the "hidden hammer" forcing enterprises to bridge their internal silos

- GDPR = 360-degree view of the customers, the the Holy Grail of better sales & marketing

146

Nortal

DeepScan

Platform that helps the governments and enterprises to know the data they have.

Lauri.Ilison@nortal.com