Leveraging audit log message for anomaly detection

IHE Symposium 2019

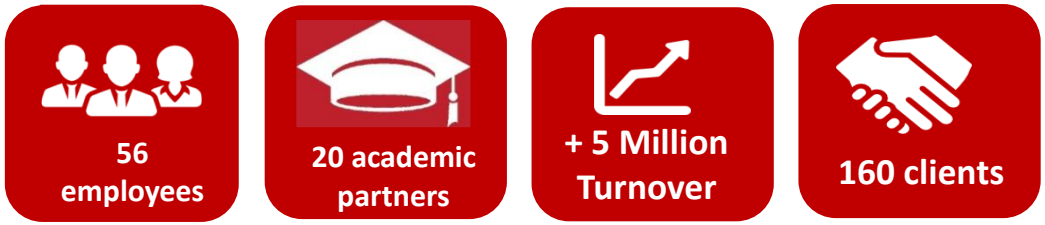
Yacine TAMOUDI
yacine.tamoudi@kereval.com
Présentation date 09-04-2019

Agenda

- Introduction
- Cybersecurity in healthcare key figures
- Cybersecurity measures
- Kereval audit based anomaly detection
- Conclusion

Who are we?

- French software testing laboratory
  - ISO/IEC 17025 Accredited
- Created in 2002
- Key figures :



- Our activity : Healthcare interoperability, Automotive Industry, Cybersecurity, …

Cybersecurity and healthcare

- Number of connected healthcare systems is increasing
  - EHR, ePrescription, cross border eHealth
- Number of incidents and breaches is increasing:
  - From 15% in 2017 to 24% in 2018 of breaches affected healthcare organizations
- Affected a total of 5.579 million patient records in the US in 2017
- The cost of a breach by patient record average around $380

Verizon : 2018 Data Breach Investigations Report; Welchallyn : healthcare-cybersecurity-statistics

Most Common Attack Vectors

Incidents by attack vectors	All sectors (Public, Retail, Financial,…)
Denial of Service	36,2%
Crimeware (Ransomware,…)	16,7%
Web Applications	10,9%
Point of Sale	10,6%
Privilege Misuse	5,1%
Miscellaneous Errors	3,1%
Cyber-Espionage (Phishing,…)	2,9%
Lost and Stolen Assets	2,1%
Everything Else	12,3%

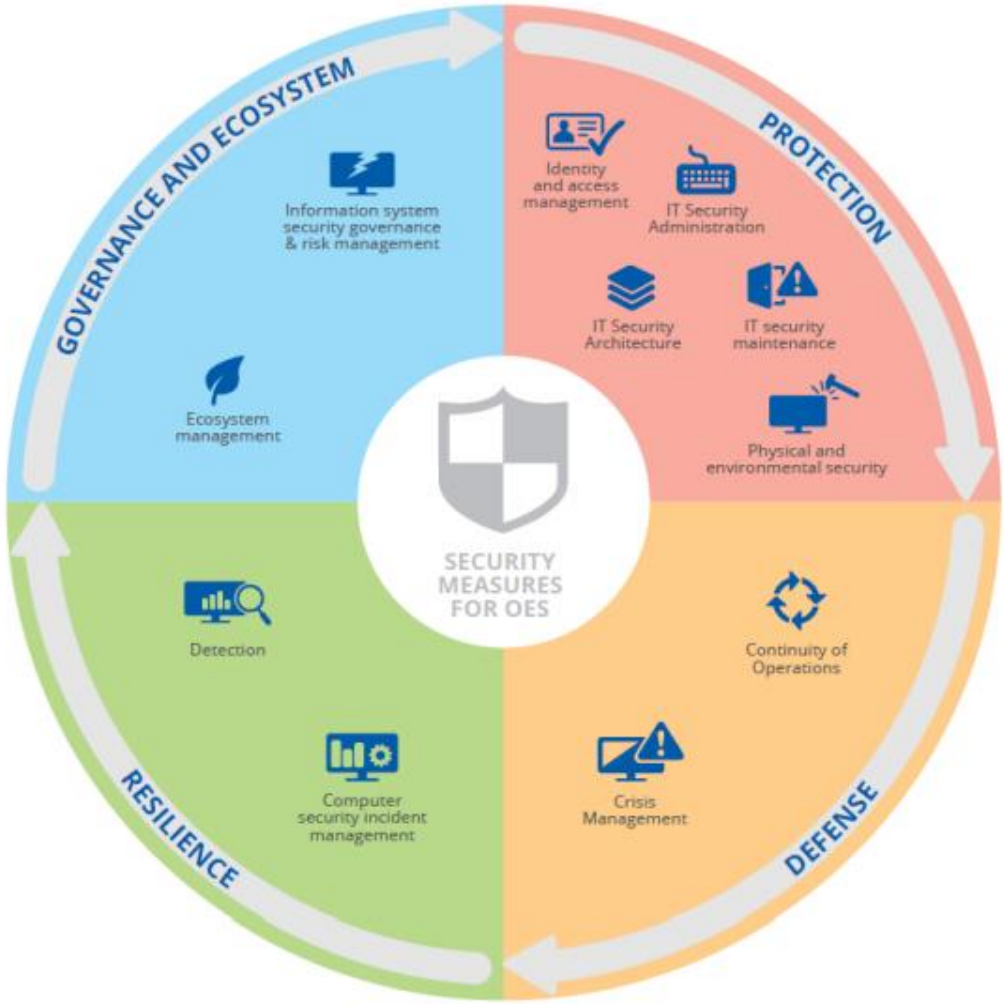
Healthcare Attack Vectors

Incidents by attack vectors	Healthcare	Other (Public, Retail, Financial,…)
Miscellaneous Errors	**24,1%**	**3,1%**
Crimeware (Ransomware,…)	20,5%	16,7%
Cyber-Espionage (Phishing,…)	**18,4%**	**2,9%**
Lost and Stolen Assets	**12,8%**	**2,1%**
Web Applications	11,7%	10,9%
Privilege Misuse	3,2%	5,1%
Denial of Service	0,1%	36,2%
Point of Sale	0,1%	10,6%
Everything Else	9,1%	12,3%

➔ 56% of incidents involving internal actors compared to 21% on average in other sectors

Security requirements

- Many sources : national, NIS, …
- Organisational
  - Asset identification
  - Risk analysis
  - Training
  - Continuity of operations
  - Crisis management
  - Incident management
  - …
- Technical
  - Cryptography
  - Traffic Filtering
  - Authentication and Identification
  - Access Rights
  - Logging
  - Detection
  - Logs Correlation and Analysis
  - …



ENISA - Guidelines on assessing DSP and OES compliance to the NISD security requirements

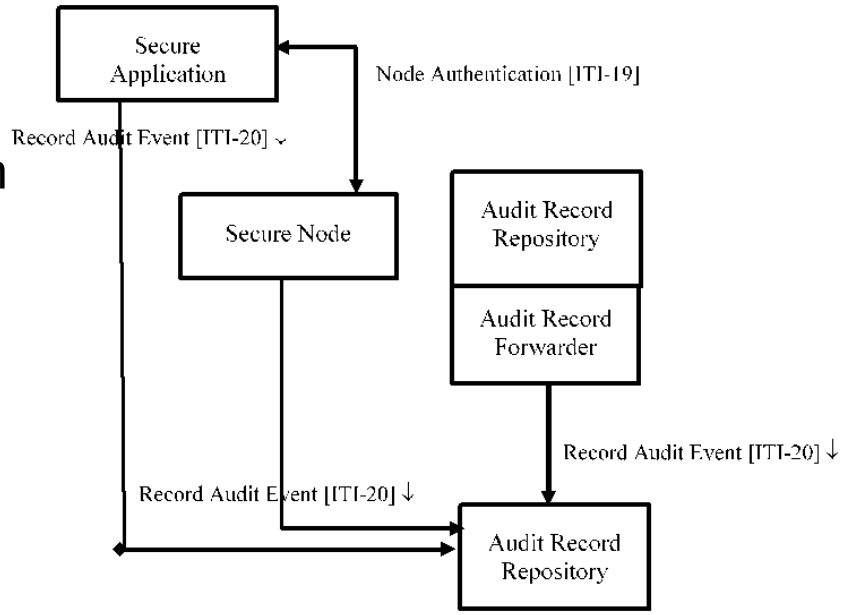
Cybersecurity and IHE

- IHE integration profiles related to security
  - Standardization of cybersecurity measures with compatibility among vendors
  - Respond to basic technical security requirements

- Technical profiles
  - Cross-Enterprise User Assertion (XUA)
    - Identity and authorization
  - Document Digital Signature (DSG) & Document Encryption (DEN)
    - Cryptography
  - Basic Patient Privacy Consents (BPPC), Advanced Patient Privacy Consents (APPC) & Internet User Authorization (IUA)
    - Authorization
  - Enterprise User Authentication (EUA)
    - Authentication
  - Audit Trails and Node Authentication (ATNA)
    - Authentication
    - Cryptography
    - Auditing

ATNA Profile

- Motives
  - Channel encryption and authentication
  - Specification of audit records
- Actors
  - Secure Node
  - Secure Application
  - Audit Record Repository
- Mitigation against unauthorized use
  - Actor authentication
  - A posteriori Audit log investigation for patterns and behavior outside policy
- Audit Record Repository
  - Can filter and auto-forward
  - Restful ATNA to access audit records



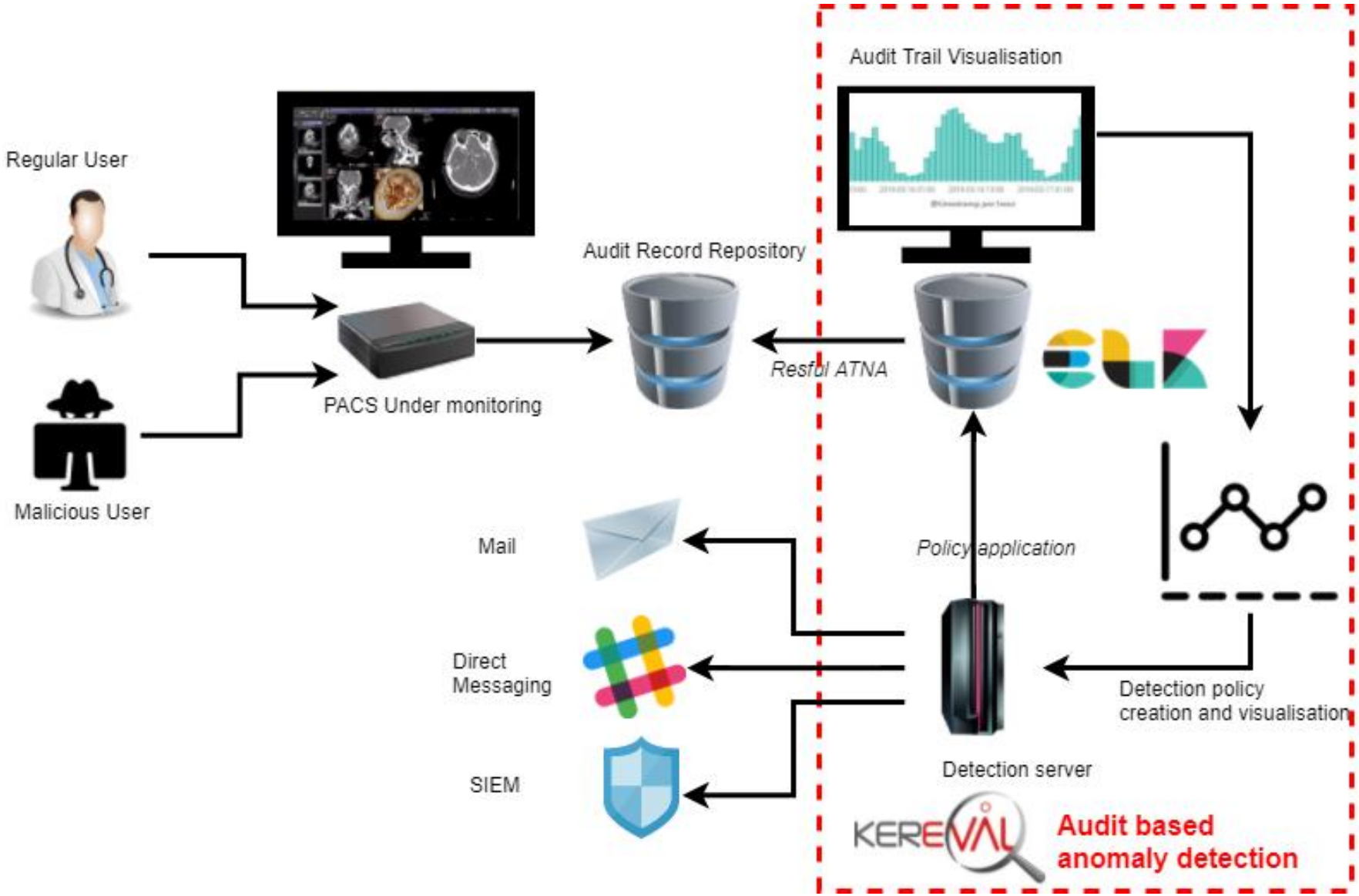
Why anomaly detection ?

- Healthcare is much prone to misuse, and errors and breach involving internal actors than the other sectors
  - 56% of attacks involve internal users
- All these attack can bypass efficient access control
  - Stolen credential
    - Phishing
    - Negligence
    - Breach
    - Vulnerability(CSRF,…)
    - Lost / stolen credential
  - Malicious neighbour doctor
- The access is valid but the behaviour pattern is different
- How can we detect anomaly in behaviour pattern?
  - Through the audit logs (ATNA)
    - Critical actions should be monitored and audited
    - Common logging format for different applications
    - Interoperability of the audit structure

Case Study

- Kereval and local universities
- Experimentation
  - With a PACS
    - Web portal
    - DICOM access point
    - HL7 access point
  - Test bench for
    - Simulating regular data
    - Model and simulate misuse scenarios
- Threats addressed :
  - Practitioners credential abuse
- PACS under test:
  - Data :
    - Patient personal and medical data
    - Imaging
    - Practitioners ID
  - Actions
    - Create/Delete/Modify Patient
    - Create/Delete/Modify Studies

Methodology



ATNA audit message structure

- DICOM PS 3.15
  - IHE add requirements for critical transactions
  - Can also be use for internal message depending on the application

- Useful data for intrusion detection
  - EventActionCode
    - Execute, Create, Read,…
  - EventIdentification
    - EventDateTime
    - EventID
    - EventTypeCode
  - ActiveParticipant
    - RoleIDCode
    - NetworkAccessPointID
  - ParticipantObjectIdentification
    - Transaction subject
    - Document / Patient / Media type / Request

Example audit message

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<AuditMessage xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
              xsi:noNamespaceSchemaLocation="http://www.dcm4che.org/DICOM/audit-message.rnc">
    <EventIdentification EventActionCode="E" EventDateTime="2019-04-04T16:04:17.912+02:00" EventOutcomeIndicator="0">
        <EventID csd-code="110112" codeSystemName="DCM" originalText="Query"/>
    </EventIdentification>
    <ActiveParticipant UserID="admin" UserIsRequestor="true" UserTypeCode="1" NetworkAccessPointID="192.168.0.53"
                       NetworkAccessPointTypeCode="2">
        <RoleIDCode csd-code="110153" codeSystemName="DCM" originalText="Source Role ID"/>
        <UserIDTypeCode csd-code="113871" codeSystemName="DCM" originalText="Person ID"/>
    </ActiveParticipant>
    <ActiveParticipant UserID="/dcm4chee-arc/aets/DCM4CHEE/rs/patients/count" AlternativeUserID="841"
                       UserIsRequestor="false" UserTypeCode="2"
    NetworkAccessPointID="127.0.0.1" NetworkAccessPointTypeCode="2">
        <RoleIDCode csd-code="110152" codeSystemName="DCM" originalText="Destination Role ID"/>
        <UserIDTypeCode csd-code="12" codeSystemName="RFC-3881" originalText="URI"/>
    </ActiveParticipant>
    <AuditSourceIdentification AuditSourceID="dcm4chee-arc">
        <AuditSourceTypeCode csd-code="4"/>
    </AuditSourceIdentification>
    <ParticipantObjectIdentification ParticipantObjectID="CountPatients" ParticipantObjectTypeCode="2"
                                     ParticipantObjectTypeCodeRole="24">
        <ParticipantObjectIDTypeCode csd-code="QIDO" originalText="QIDO Query" codeSystemName="99DCM4CHEE"/>
        <ParticipantObjectQuery>L2RjbTRjaGVlLWFyYy9hZXRzL0RDTTRDSEVFL3JzL3BhdGllbnRzL2NvdW50b3JkZXJieT0tUGF0aWVud
                                E5hbWUmcmV0dXJuZW1wdHk9ZmFsc2UmUGF0aWVudE5hbWU9QWxpY2U=</ParticipantObjectQuery>
        <ParticipantObjectDetail type="QueryEncoding" value="VVRGLTg="/>
    </ParticipantObjectIdentification>
</AuditMessage>
```

Detection rules

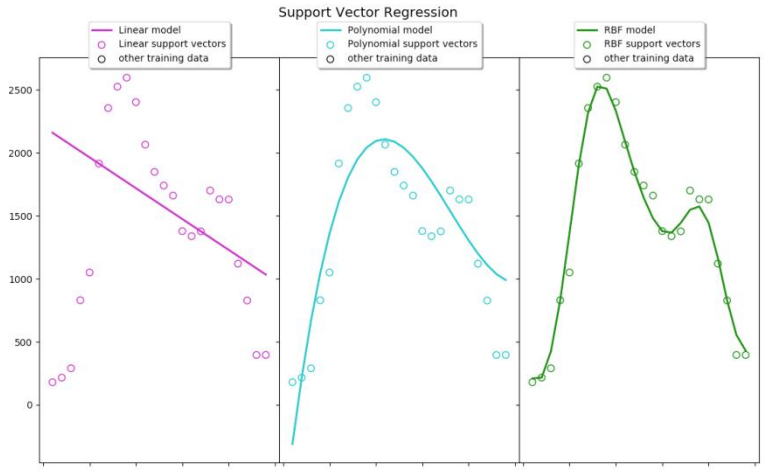
- Threshold on the frequency of action by an Active Participant
  - By Event Identification
  - By Participant
  - By time period
- EventID profile by an Active Participant
- Active Participant or Participant Object Values:
  - Dictionary
    - ParticipantObjectID
    - NetworkAccessPointID
  - % of new data
- Unusual action sequencing:
  - Read without create

Machine Learning

- Automatic rule creation from the observation of normal activity in the system

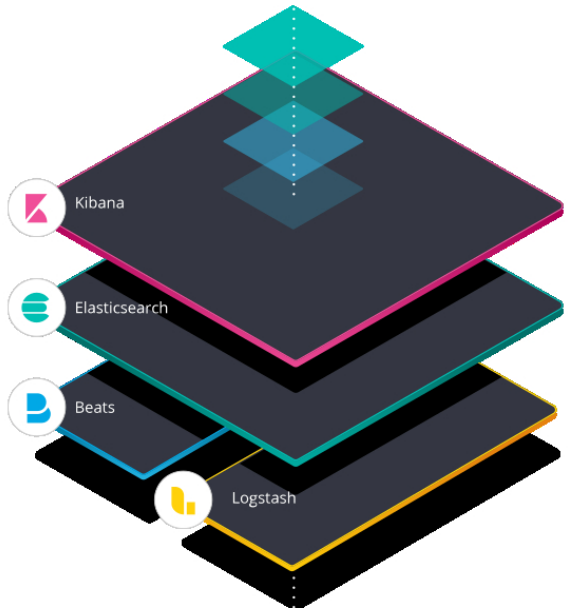
- Blackbox ML vs White box ML
  - All alert can be easily checked by an operator

- Auditable algorithms
  - Threshold detection
  - Regression
  - Pattern Mining



Architecture

- Massive quantity of audit messages
- ELK Based solution
  - ElasticSearch
  - Logstash
  - Kibana



- Possibility to adapt to various input format
- Scalability of logs :
  - Automatic sharding and replication, flexible schema
  - Used by Netflix, LinkedIn, Stackoverflow,…

Experimentation results

- The interoperability of ATNA means it is
  - Interoperability ATNA Audit Record Repository
  - Requirement for necessary items interesting for detection
  - Use of custom visualisation for specific audit messages
- Experimentation
  - Tested with a PACS and simulated data
- Next Steps
  - Experimentation with more powerful machine learning models
  - Experimentation with partners

Perspectives

- The healthcare sector is heavily impacted by cybersecurity
  - Some of the threats are specific to healthcare
- On top of traditional security measures
  - Detect abnormal user behaviour and raise alerts
  - Fit for different usecases (Cross-community access, DMP, …)
- Easily customisable approach leveraging ATNA audit message interoperability

Thank you for your attention