Patterns for New Software Engineering:
Machine Learning, IoT and Security Patterns

Hironori Washizaki
Professor at Waseda University, Tokyo, Japan
washizaki@waseda.jp

The Seventh International Conference on Fundamentals and Advances in Software Systems Integration (FASSI 2021)
November 14-18, 2021 Athens, Greece
Prof. Dr. Hironori Washizaki

- Professor and the Associate Dean of the Research Promotion Division at Waseda University in Tokyo
- Visiting Professor at the National Institute of Informatics
- Outside Directors of SYSTEM INFORMATION and eXmotion
- Research and education projects
  - Leading a large-scale grant at MEXT enPiT-Pro Smart SE
  - Leading framework team of JST MIRAI eAI project
- Professional contributions
  - IEEE Computer Society Vice President for Professional and Educational Activities
  - Editorial Board Member of MDPI Education Sciences
  - Steering Committee Member of the IEEE Conference on Software Engineering Education and Training (CSEE&T)
  - Associate Editor of IEEE Transactions on Emerging Topics in Computing
  - Advisory Committee Member of the IEEE-CS COMPSAC
  - Steering Committee Member of Asia-Pacific Software Engineering Conference (APSEC)
  - Convener of ISO/IEC/JTC1 SC7/WG20
Patterns for New Software Engineering: Machine Learning, IoT and Security Patterns

Hironori Washizaki
Professor at Waseda University, Tokyo, Japan

washizaki@waseda.jp  http://www.washi.cs.waseda.ac.jp/


Supported by JST MIRAI eAI JPMJMI20B8 https://www.jst.go.jp/mirai/jp/program/super-smart/JPMJMI20B8.html
Supported by JSPS Bilateral Joint Research Project
Agenda

• Paradigm shifts in new software engineering
• Pattern language
• Machine learning patterns
• IoT patterns
• Security patterns
What is software engineering?

• “Application of systematic, disciplined, quantifiable approach to development, operation, and maintenance of software” – SWEBOK 2014

• Guide to the Software Engineering Body of Knowledge (SWEBOK)

- Software Requirements
- Software Design
- Software Construction
- Software Testing
- Software Maintenance
- Software Configuration Management
- Software Engineering Management
- Software Engineering Process

- Software Engineering Tools and Methods
- Software Quality
- Software Engineering Professional Practice
- Software Engineering Economics

- Computing Foundations
- Mathematical Foundations
- Engineering Foundations
Vision of SWEBOK 2022 (subject to change)

(Evolution lead: Hironori Washizaki, since 2018-)

https://www.computer.org/volunteering/boards-and-committees/professional-educational-activities/software-engineering-committee/swebok-evolution

- Expansion of SE
  - AI/Machine Learning Engineering
  - Restructuring foundation areas incl. Internet of Things (IoT)
- Value in SE
  - Value proposition
- Dependable SE
  - Architecture
  - Security
- Modern SE
  - Agile
  - DevOps
Paradigm shifts in “new” software engineering

<table>
<thead>
<tr>
<th></th>
<th>Current</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scope and perspective</strong></td>
<td>Software systems</td>
<td>Software systems, <strong>business, society and related disciplines</strong></td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>Planned, static, common, and closed</td>
<td>Adaptive, dynamic, diverse, and open</td>
</tr>
<tr>
<td><strong>Focus</strong></td>
<td>Specification</td>
<td><strong>Value, data, and speed</strong></td>
</tr>
<tr>
<td><strong>Thinking</strong></td>
<td>Cognitive (logical) or affective (design)</td>
<td>Cognitive (logical), affective (design), and <strong>conative (conceptual)</strong></td>
</tr>
<tr>
<td><strong>Inference</strong></td>
<td>Deduction and analogy</td>
<td>Deduction, analogy, <strong>induction, and abduction</strong></td>
</tr>
</tbody>
</table>

Example: Deduction and induction

Conventional software: Deduction

Goal  →  Model  →  Behavior  →  Data

ML-based software: Induction

Goal  →  Model  →  Behavior  →  Data

H. Maruyama, “Machine Learning Engineering and Reuse of AI Work Products,” The First International Workshop on Sharing and Reuse of AI Work Products, 2017

Problem and goal

• ML and IoT are key enablers of digital transformations.
• Patterns in ML and IoT software design are not well classified and studied.
• Security must be a critical cross-cutting concern in ML and IoT software.
• We are conducting a systematic literature review to reveal landscapes of ML, IoT, and security software engineering patterns.
Agenda

• Paradigm shifts in new software engineering
• Pattern language
• Machine learning patterns
• IoT patterns
• Security patterns
Street Cafe

**Problem**: Needs to have a place where people can sit lazily, legitimately, be on view, and watch the world go by…

**Solution**: Encourage local cafes to spring up in each neighborhood. Make them intimate places, with several rooms, open to a busy path …

Towards a pattern language

... OK, so, to attract many people to our city, **Small Public Square**s should be located in the center. At the **Small Public Square**, make **Street Cafes** be **Opening to the Street** ...
New SE needs pattern (language)!

- **Bridge** between abstract paradigms and concrete cases/tools
  - Verbalizing and documenting Know-Why (context), What (problem) and How (solution)
  - Reusing solutions and problems
  - Getting consistent architecture

- **Common language** among stakeholders
  - Software engineers, hardware engineers, network engineers, domain experts, data analyst ...
Agenda

- Paradigm shifts in new software engineering
- Pattern language
- Machine learning patterns
- IoT patterns
- Security patterns
Practices and patterns in ML-SE

- Researchers and practitioners studying best practices strive to design Machine Learning (ML) systems and software.
- Some practices are formalized as patterns.

(NOTE: NOT handle ML model patterns.)

Different Workloads in Different Computing Environments (e.g., Facebook)

K. M. Hazelwood, et al., Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective, HPCA 2018
Research questions

• RQ1. Does academic and gray literature address the design of ML systems and software?
  – 19 scholarly and 19 gray documents identified
  – 15 SE patterns for ML applications extracted

• RQ2. Can ML patterns be classified?
  – Categories of scopes: Topology, programming and model

• RQ3. How do practitioners perceive ML patterns?
  – Questionnaire-based survey for 600+ developers
  – Developers were unfamiliar with most ML patterns, although there were several major patterns used by 20%
RQ1. Does academic and gray literature address the design of ML systems and software?

- Systematic Literature Review (SLR)
  - Scholar papers: Engineering Village
  - Gray documents: Google

- 19 scholarly papers and 19 gray documents identified
- 15 patterns extracted

**Engineering Village**

```plaintext
(((system) OR (software)) AND (machine learning) AND (implementation pattern) OR (pattern) OR (architecture pattern) OR (design pattern) OR (anti-pattern) OR (recipe) OR (workflow) OR (practice) OR (issue) OR (template))) WN ALL) + ((cpx OR ins OR kna) WN DB) AND ({{ca} OR {ja} OR {ip} OR {ch}} WN DT)
```

**Google**

```plaintext
(system OR software) "Machine learning" (pattern OR "implementation pattern" OR "architecture pattern" OR "design pattern“ OR anti-pattern OR recipe OR workflow OR practice OR issue OR template)
```

"machine implementation pattern" OR "architecture pattern" OR "design pattern“ OR anti-pattern OR recipe OR workflow OR practice OR issue OR template
Numbers of Documents per Year

- ML application systems have recently become popular due to the promotion of artificial intelligence.
- Since 2008, academic and gray documents have discussed good (bad) practices of ML application systems design.

RQ2. Can ML patterns be classified?

- Model operation patterns that focus on ML models
- Programming patterns that define the design of a particular component
- Topology patterns that define the entire system architecture.
## Topology patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Problem</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Different Workloads in Different Computing Environments</strong></td>
<td>It is necessary to separate and quickly change the ML data workload ...</td>
<td>Physically isolate different workloads to separate machines...</td>
</tr>
<tr>
<td><strong>Distinguish Business Logic from ML Models</strong></td>
<td>The overall business logic should be isolated from the ML models ...</td>
<td>Separate the business logic and the inference engine, loosely coupling the business logic and ML-specific dataflows.</td>
</tr>
<tr>
<td><strong>ML Gateway Routing Architecture</strong></td>
<td>Difficult to set up and manage individual endpoints for each service...</td>
<td>Install a gateway before a set of applications ...</td>
</tr>
<tr>
<td><strong>Microservice Architecture for ML</strong></td>
<td>ML applications may be confined to some “known” ML frameworks ...</td>
<td>Provide well-defined services to use for ML frameworks....</td>
</tr>
<tr>
<td><strong>Lambda Architecture for ML</strong></td>
<td>Real-time data processing requires scalability, fault tolerance, predictability ...</td>
<td>The batch layer keeps producing views while the speed layer creates the relevant real-time views ...</td>
</tr>
<tr>
<td><strong>Kappa Architecture for ML</strong></td>
<td>It is necessary to deal with huge amount of data with less code resource ...</td>
<td>Support both real-time data processing and continuous reprocessing with a single stream processing engine ...</td>
</tr>
</tbody>
</table>

**Distinguish Business Logic from ML Models**

- **Problem:** Business logic should be isolated from ML models so that they can be changed without impacting rest of business logic.
- **Solution:** Separate the business logic and the inference engine, loosely coupling the business logic and ML-specific dataflows.

![Architectural Layers Diagram]

- **Legend**
  - **Architectural Layers**
  - **Deployed as ML System**
  - **Business Logic Data Flow**
  - **ML Runtime Data Flow**
  - **ML Development Data Flow**

H. Yokoyama, Machine Learning System Architectural Pattern for Improving Operational Stability, ICSA-C, 2019
H. Washizaki, et al., Software Engineering Patterns for Machine Learning Applications (SEP4MLA), AsianPLoP 2020
Usage of **Distinguish Business Logic from ML Models**

**Presentation Layer**
- User Interface (Chatbot UI)
  - Input: Web App Front-end
  - Output: Slack

**Logic Layer**
- Business Logic (Chatbot Logic)
  - Input: Web App Back-end
  - Output: Slack

**Data Layer**
- Database (Previous Q&A Store)
  - Input: DB Server
  - Output: (None)

**Data Collection (Dataset)**
- Input: Datasets
  - Output: Nagoya Univ. Conversation Corpus

**Data Processing (Text to Vector Transformer)**
- Input: Datasets
  - Output: NN Model pre- and post-processing
  - Input: TensorFlow

**Inference Engine (Language Model)**
- Input: NN Model
  - Output: TensorFlow

**Data Lake (Vectorized Corpus)**
- Input: Word Vector
  - Output: TensorFlow (Text)

**Legend**
- Architectural Elements (Example Role as Chatbot)
  - What
  - How

**Data Flow**
- Business Logic Data Flow
- ML Runtime Data Flow
- ML Development Data Flow
# Programming patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Problem</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Lake for ML</td>
<td>We cannot foresee the kind of analyses that will be performed on the data ...</td>
<td>Store data, which range from structured to unstructured, as “raw” as possible into a data storage ...</td>
</tr>
<tr>
<td>Separation of Concerns and Modularization of ML Components</td>
<td>ML applications must accommodate regular and frequent changes to their ML components ...</td>
<td>Decouple at different levels of complexity from the simplest to the most complex ...</td>
</tr>
<tr>
<td>Encapsulate ML Models within Rule-based Safeguards</td>
<td>ML models are known to be unstable and vulnerable to adversarial attacks, drifts, ...</td>
<td>Encapsulate functionality in the containing system using deterministic and verifiable rules ...</td>
</tr>
<tr>
<td>Discard PoC Code</td>
<td>The code created for Proof of Concept (PoC) often includes code that sacrifices maintainability ...</td>
<td>Discard the code created for the PoC and rebuild maintainable code ...</td>
</tr>
</tbody>
</table>
Encapsulate ML Models within Rule-based Safeguards

- **Problem:** ML models are known to be unstable and vulnerable to adversarial attacks, noise, and data drift.
- **Solution:** Encapsulate functionality provided by ML models and deal with the inherent uncertainty in the containing system using deterministic and verifiable rules.
- **Know usage:** E.g. Apollos’s object detection [Peng20]
## Model operation patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Problem</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter-Server Abstraction</strong></td>
<td>For distributed learning, widely accepted abstractions are lacking ...</td>
<td>Distribute both data and workloads over worker nodes, while the server nodes maintain globally shared parameters ...</td>
</tr>
<tr>
<td><strong>Data Flows Up, Model Flows Down</strong></td>
<td>Standard ML approaches require centralizing the training data on one machine ...</td>
<td>Enable mobile devices to collaboratively learn while keeping all the training data on the device as federated learning ...</td>
</tr>
<tr>
<td><strong>Secure Aggregation</strong></td>
<td>The system needs to communicate and aggregate model updates in a secure and scalable way ...</td>
<td>Encrypt data from each device and calculate totals and averages without individual examination ...</td>
</tr>
<tr>
<td><strong>Deployable Canary Model</strong></td>
<td>A surrogate ML that approximates the behavior of best model must be built to provide explainability ...</td>
<td>Run the explainable inference pipeline in parallel to monitor prediction differences ...</td>
</tr>
<tr>
<td><strong>ML Versioning</strong></td>
<td>ML models and their different versions may change the behavior of the overall ML applications ...</td>
<td>Record the ML model, dataset, and code to ensure a reproducible training and inference processes ...</td>
</tr>
</tbody>
</table>

Deployable Canary Model

- **Problem:** A surrogate ML that approximates the behavior of the best ML model must be built to provide explainability.
- **Solution:** Run the explainable inference pipeline in parallel with the primary inference pipeline to monitor prediction differences.
- **Known usage:** Image-based anomaly detection at factory

---

S. Ghanta et al., Interpretability and reproducibility in production machine learning applications, ICMLA 2018
<table>
<thead>
<tr>
<th>Pattern</th>
<th>Performance</th>
<th>Compatibility</th>
<th>Reliability</th>
<th>Security</th>
<th>Maintainability</th>
<th>Portability</th>
<th>Robustness</th>
<th>Explainability</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different Workloads in Different Computing Environments</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distinguish Business Logic from ML Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML Gateway Routing Architecture</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microservice Architecture for ML</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Lambda Architecture for ML</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappa Architecture for ML</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Lake for ML</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separation of Concerns and Modularization of ML Components</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encapsulate ML Models within Rule-based Safeguards</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Discard PoC Code</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter-Server Abstraction</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Flows Up, Model Flows Down</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Secure Aggregation</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Deployable Canary Model</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML Versioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
RQ3. Practitioners’ insights on quality

- Surveyed 300+ developers, 46 answered in ML development

What product quality attributes considered?
- Maintainability, reliability, security, and usability

What model and prediction quality attributes?
- Robustness, accuracy, and explainability

Maintainability, reliability, robustness and accuracy are well handled by ML patterns. There are demands for having ML patterns addressing security, usability, and explainability, which are not handled well now.
Practitioners’ insights on ML design patterns

- Surveyed 600+ developers, 118 answered
- Have you ever referred to ML patterns?
  - Major: ML Versioning, Microservice Architecture for ML
  - None: Secure Aggregation, Data Flows Up (aka. Federated Learning)
Practitioners’ insights on ML design patterns

• Have you ever referred to ML patterns?
  – Developers were unfamiliar with most ML patterns, although there were several major patterns used by 20+% of the respondents.
  – For all patterns, most respondents indicated that they would consider using them in future designs.
  – Promoting existing ML patterns will increase their utilization

• How do you solve and share design challenges of ML application systems?
  – 37 (i.e., 31%) organized design patterns and past design results.
  – As respondents become more organized in their approach to design problems by reuse, the pattern usage ratio increased.
  – Development teams and organizations will reuse more ML patterns as they become more consistent in their reuse approach.

<table>
<thead>
<tr>
<th>Design solution and reuse practice</th>
<th>#Respondents</th>
<th>#Patterns used</th>
<th>Pattern usage ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lv3. Organizing, reusing patterns (and past results)</td>
<td>37</td>
<td>64</td>
<td>11.5%</td>
</tr>
<tr>
<td>Lv2. Reusing externally documented patterns</td>
<td>31</td>
<td>50</td>
<td>10.8%</td>
</tr>
<tr>
<td>Lv1. Resolving problems in an ad-hoc way</td>
<td>37</td>
<td>35</td>
<td>6.3%</td>
</tr>
<tr>
<td>Others</td>
<td>13</td>
<td>3</td>
<td>1.5%</td>
</tr>
</tbody>
</table>
Conclusion and future work

• Literature review of academic and gray literature
  – 15 SE patterns for ML applications extracted.
  – Patterns at https://eai-transfer.github.io/ml-design-pattern/en/

• Survey of practitioners’ insights
  – Developers were unfamiliar with most ML patterns, although there were several major patterns used by 20% (such as ML Versioning and Microservice Architecture for ML)

• Identify ML patterns addressing specific quality attributes that are not handled well now
  – Security, usability, and explainability

• Future work
  – Investigate the impact of patterns on quality attributes of systems
  – Analyze relationships among patterns including related ones towards a pattern language
  – Integration into framework to handle from requirements on implementations and testing/debugging
Agenda

• Paradigm shifts in new software engineering
• Pattern language
• Machine learning patterns
• IoT patterns
• Security patterns
Executive summary

• IoT architecture and design patterns at different abstraction levels are not well classified and studied.

• **RQ1. How does academic literature address IoT architecture and design patterns?**
  – There are 32 academic papers related to IoT architecture and design patterns.

• **RQ2. Are all existing IoT architecture and design patterns really IoT patterns?**
  – Of the 143 extracted patterns, 57% are non-IoT patterns.

• **RQ3. Can IoT architecture and design patterns be classified?**
  – Patterns can be divided along three main characteristics: abstraction level, domain specificity, and quality attributes.

• **RQ4. What IoT architecture and design patterns exist?**
  – Many IoT patterns address interoperability, security, and maintainability.
  – Many IoT architecture patterns are domain-specific.
Systematic literature review (SLR)

• Initial Search: 63 papers 2014–2018 in Scopus
  – “IoT” AND (“design pattern” OR “architecture pattern”)
• Impurity Removal: 56
• Inclusion and Exclusion Criteria: 32
  – Inclusion: Addressing patterns for designing IoT systems and software, and papers written in English
• Data Extraction
  – Publication title, publication year, publication venue
  – Types of patterns proposed or used, pattern names
  – Domain names in the case of Specific IoT patterns
  – Quality attributes addressed
RQ1. How does academic literature address IoT architecture and design patterns?

• 32 academic papers related to IoT architecture and design patterns
  – Most are conference papers followed by journal publications.
• The high number of conference papers indicates that the entire topic of IoT architecture and design patterns is in its early stage
• But the presence of journal articles suggests that some types of IoT patterns are maturing.
RQ2. Are all existing IoT architecture and design patterns really IoT patterns?

- 143 patterns mentioned in 32 papers

- **82 general (non-IoT) patterns**
  - Incl. 11 non-IoT patterns appeared in multiple papers: Publish-Subscribe, Client-Server, Peer-to-Peer, REST, SOA, RBAC, MVC, Reflection, Blockchain, Strategy and Observer
  - 14 papers used such **non-IoT patterns only**.
  - **IoT systems and software are often designed via conventional architecture and design patterns.**

- 61 IoT patterns in 18 papers
RQ3. Can IoT architecture and design patterns be classified?

• **Abstraction level**
  – High: Architecture styles
  – Middle: Architecture patterns
  – Low: Design patterns

• **Domain specificity**
  – Any: General architecture/design patterns those can be adopted to design IoT systems and software
  – General IoT: Applicable to any IoT systems and software.
  – Specific IoT: Addressing specific problem/technical domains

• **Quality characteristic**
  – All quality characteristics except for functional suitability in ISO/IEC 25010
    + Emerging characteristics common in IoT such as scalability and privacy
E.g.: Layered architecture for IoT applications

- IoT platform providing resource virtualization using **lightweight virtualization for multi-layer** applications

---

<table>
<thead>
<tr>
<th>Core IoT Microservices</th>
<th>Edge IoT Microservices</th>
<th>IoT Gateway Microservices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core-Cloud</td>
<td>Edge-Cloud</td>
<td>Aggregators</td>
</tr>
<tr>
<td>Cluster Master</td>
<td>Cluster Control Services</td>
<td>IoT Control Services</td>
</tr>
<tr>
<td>Macroservices</td>
<td>Cluster Workers</td>
<td>IoT Control Services</td>
</tr>
<tr>
<td>Autonomic Manager</td>
<td>IoT Manager</td>
<td>Autonomic Manager</td>
</tr>
<tr>
<td>IoT Control Services</td>
<td></td>
<td>IoT Control Services</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

E.g.: IoT Gateway Event Subscription

- Employ a **subscription** mechanism into the IoT gateway
- Allowing **asynchronous and mutual transmissions** of data obtained by sensors at the destination and the message between artifacts

RQ4. What IoT architecture and design patterns exist?

- IoT patterns are **not recognized by different author groups**
  - Only two patterns mentioned in multiple papers
  - Pattern authors are encouraged to carefully check existing IoT patterns
- Combinations of abstraction level and domain specificity
  - Most of IoT **design patterns are applicable to any** domain
  - Many IoT **architecture patterns exist for specific** domains
  - Unique nature of IoT adoption in specific domains appears at the architecture level
- Major quality attributes: **Interoperability, security and maintainability**

<table>
<thead>
<tr>
<th>Type</th>
<th>Non-IoT</th>
<th>General IoT</th>
<th>Domain-specific IoT</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture style</td>
<td>22</td>
<td>2</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Architecture pattern</td>
<td>7</td>
<td>1</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>Design pattern</td>
<td>53</td>
<td>38</td>
<td>4</td>
<td>95</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>82</td>
<td>41</td>
<td>20</td>
<td>143</td>
</tr>
</tbody>
</table>
Conclusions and future work

• We surveyed 143 patterns mentioned in 32 papers
  – Most of IoT design patterns are applicable to any domain but many **IoT architecture patterns exist for specific domains.**
  – Many IoT patterns address **interoperability, security and maintainability.** Other quality attributes remain to be researched.

• Further directions
  – We opened the classification results to the public and call for comments at our Website. [http://www.washi.cs.waseda.ac.jp/iot-patterns/](http://www.washi.cs.waseda.ac.jp/iot-patterns/)
  – Only Scopus for SLR. We plan to additionally use other databases.
  – Needs to analyze relationships among IoT patterns towards a pattern language
Agenda

• Paradigm shifts in new software engineering
• Pattern language
• Machine learning patterns
• IoT patterns
• Security patterns
Security concerns must be addressed at any phase

- Patterns are **recurrent problems and solutions** under specific **contexts** from requirements to maintenance
Example of security pattern

- Name: **Role-based access control (RBAC)**
- Problem: How do we assign rights to people based on their functions or tasks?
- Solution: Assign users to roles and give rights to these roles so they can perform their tasks.
- Related patterns: **Authorization**, . . .

![Diagram showing the relationship between User, Role, Right, and Protection Object]
We categorize and analyze 240 papers to clarify state-of-the-art and future directions of security pattern research in terms of 13 facets including topics and security characteristics.

E.g., breakdown of research topics

- Application: 59
- Development methodology: 22
- Classification, Catalog, Map: 20
- Analysis: 11
- Specification: 8
- Verification: 7
- Empirical and case study: 9
- Survey: 22
- Knowledge base and repository: 20
- Selection: 20
- Detection: 11
- Pattern extraction: 9
- Pattern integration: 8

Model-driven security pattern application

[PLoP’10] Model-Driven Security Patterns Application and Validation,” 17th Conference on Pattern Languages of Programs
TESEM: Test Driven Secure Modeling Tool
[ARES’13][ARES’13][IJSSE’14][ICST’15][Information’16]

Security Design Pattern

Context
Problem
Solution

Test design as requirement

Test Script

Test case

[ARES’14] Verification of Implementing Security Design Patterns Using a Test Template, Conf. Availability, Reliability and Security
[Information’16] Implementation Support of Security Design Patterns Using Test Templates, Information 7(2)
Challenges in cloud security and privacy (S&P)

- How to consistently utilize diverse S&P knowledge? ⇒ Metamodel
- How to consider S&P over different layers? ⇒ Layered metamodel

Patterns
Guidelines
Practices

Cloud services
Ex.) User Authentication
Software Application
User Authorization
Platform
Secure Config.
OS Hardening
Infrastructure
Electronic Access
Control system
Cloud Security and Privacy Metamodel (CSPM)

SaaS (Application)  PaaS (Platform)  IaaS (Infrastructure)

Modeling vulnerability and security pattern

**Common Vulnerabilities and Exposures: CVE-2012-4394** Cross-site scripting (XSS) vulnerability in apps/files/js/filelist.js in own Cloud before 4.0.5 allows remote attackers to inject arbitrary web script or HTML via the file parameter.

![Diagram showing XSS vulnerability and countermeasure approach](image-url)

**Validator** for data-injection vulnerability such as XSS
Security and privacy development process

1. Requirement Analysis
2. Design
3. Implementation
4. Test
Conclusion and future work

**Current**
- Targeting authentication and authorization
- Many researches using UML, but independent
- Often simple case studies
- Targeting existing patterns only
- Limited education for secure development methods in IoT era

**Future**
- Address various security patterns
- Integration based on common metamodel
- Complex case studies with measurements
- New vulnerabilities and patterns
- IoT and security education program
Summary

• There are paradigm shifts in “new” software engineering.
  – ML and IoT are key enablers of digital transformations.
  – Security must be a critical cross-cutting concern in ML and IoT software.

• New software engineering needs patterns and pattern languages.
  – Bridge between abstract paradigms and concrete cases/tools
  – Common language among stakeholders

• Future
  – Classify and relate patterns across over different disciplines such as ML and IoT
  – Build pattern languages
  – Open expanded community