

Semantic MediaWiki Farm Infrastructure for Multi-Organizational Sharing

Dr. Cindy Cicalese

The MITRE Corporation

SMWCon Spring 2014

May 22, 2014

MITRE Gestalt Framework



Collaborative analytics applications with custom semantic models



Patterns and extensions



Wiki Farms

MITRE

MPN

Customers

MW1.13/SMW1.5

MW1.16/SMW1.5.4

MW1.17/SMW1.8

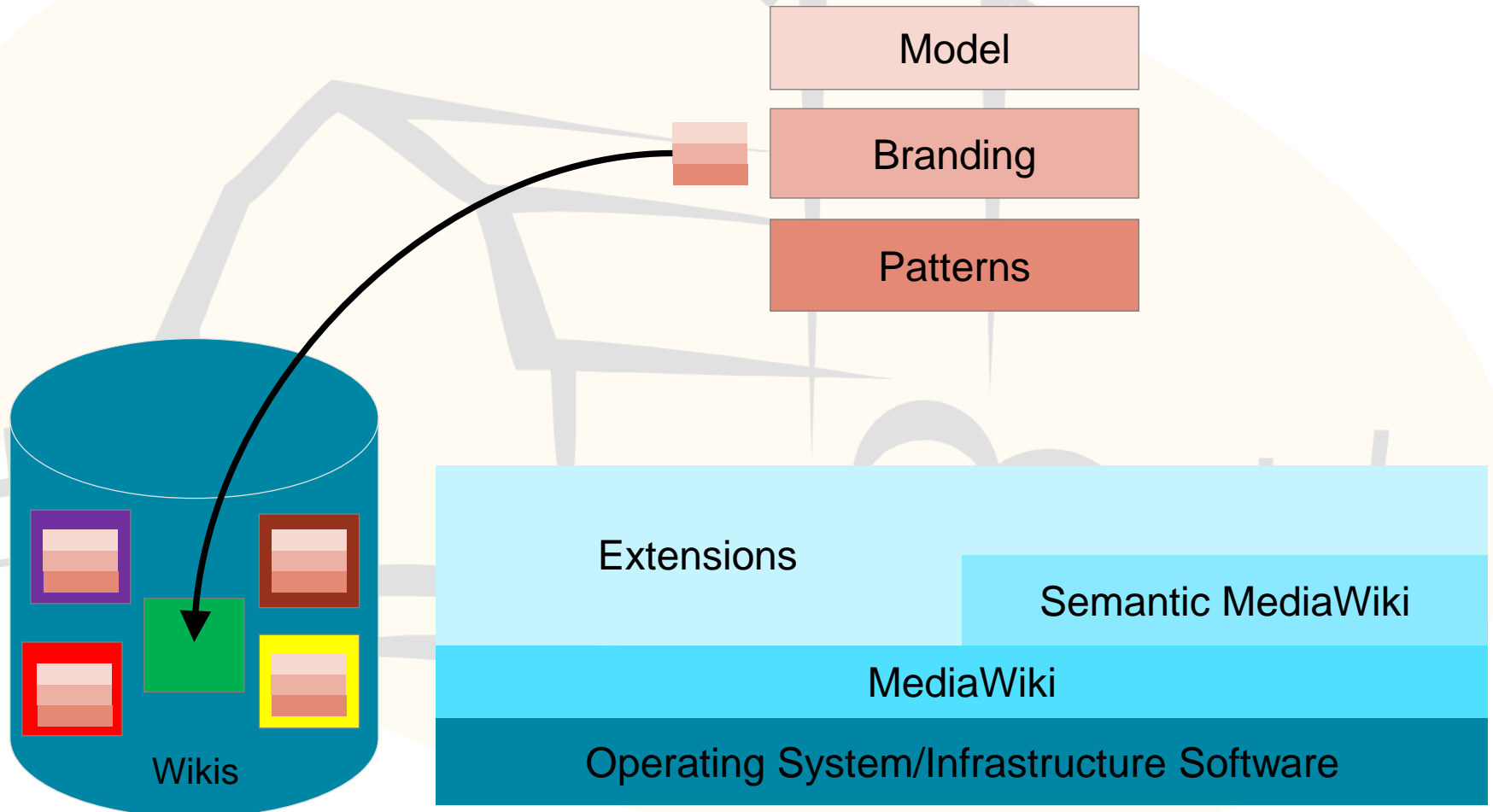
MW1.21.1/SMW1.8.0.5

MW1.22/SMW1.9.1

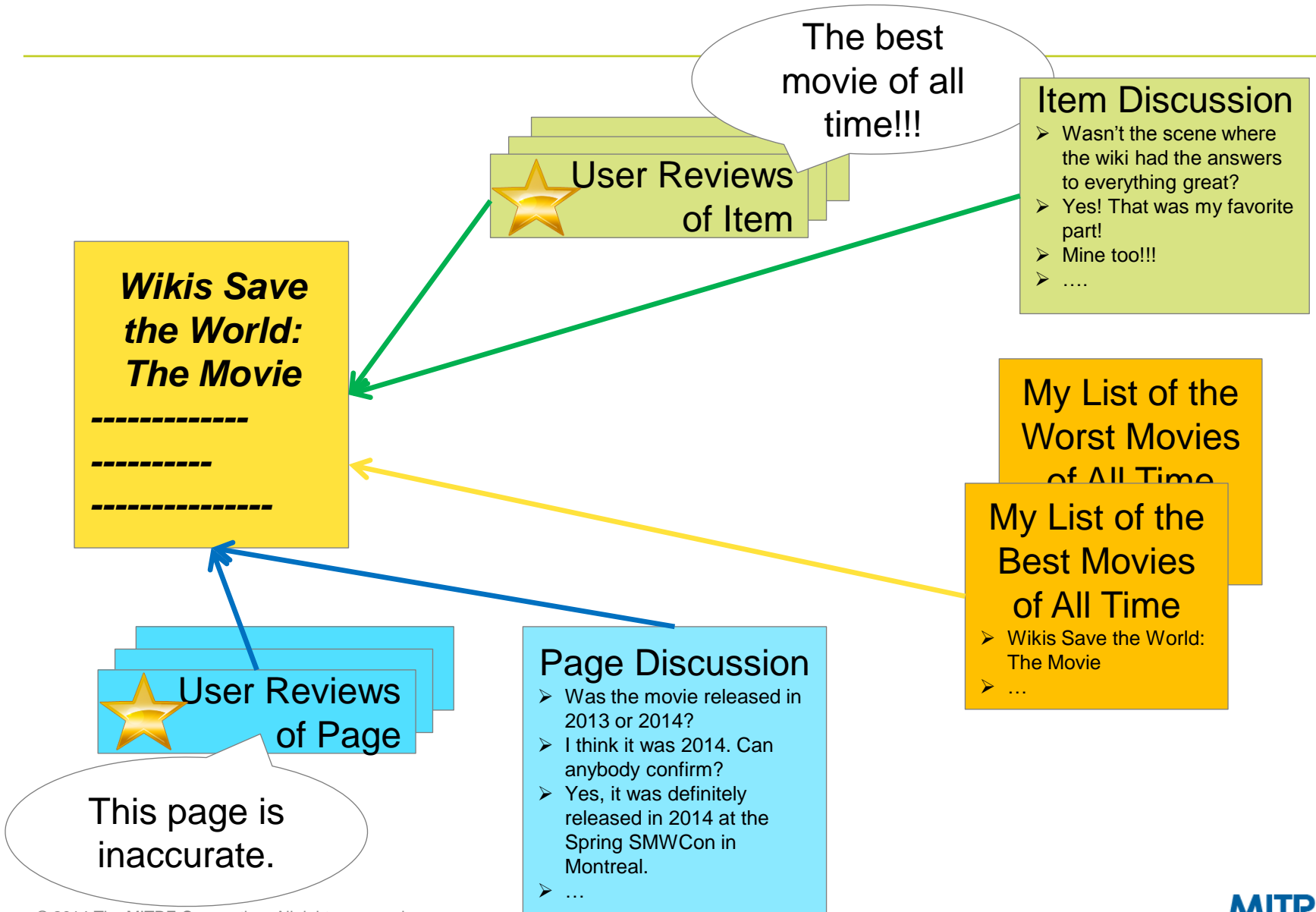
Sustained Community Knowledge Building Facilitated by Wiki Farms

- **Small, tight knit communities driven by a shared vocabulary**
 - A wiki for each community
 - Each wiki has tens to hundreds of users
 - Do not have the benefits of mega-scale crowdsourced authoring
 - But, the community for each wiki is tightly cohesive
 - Each wiki has a domain-specific semantic model
 - 7+/- 2 classes (template/category/form)
 - Each class with a manageable number of properties
- **Enterprise wiki farms need:**
 - Manageable software infrastructure
 - Tools to support consistent branding (e.g. common templates)
 - Tools to assess and maintain wiki quality
 - Tools to support cross-wiki findability and user identity

Wiki Farm Deployment and Maintenance



Pattern Example: Reviews and Lists



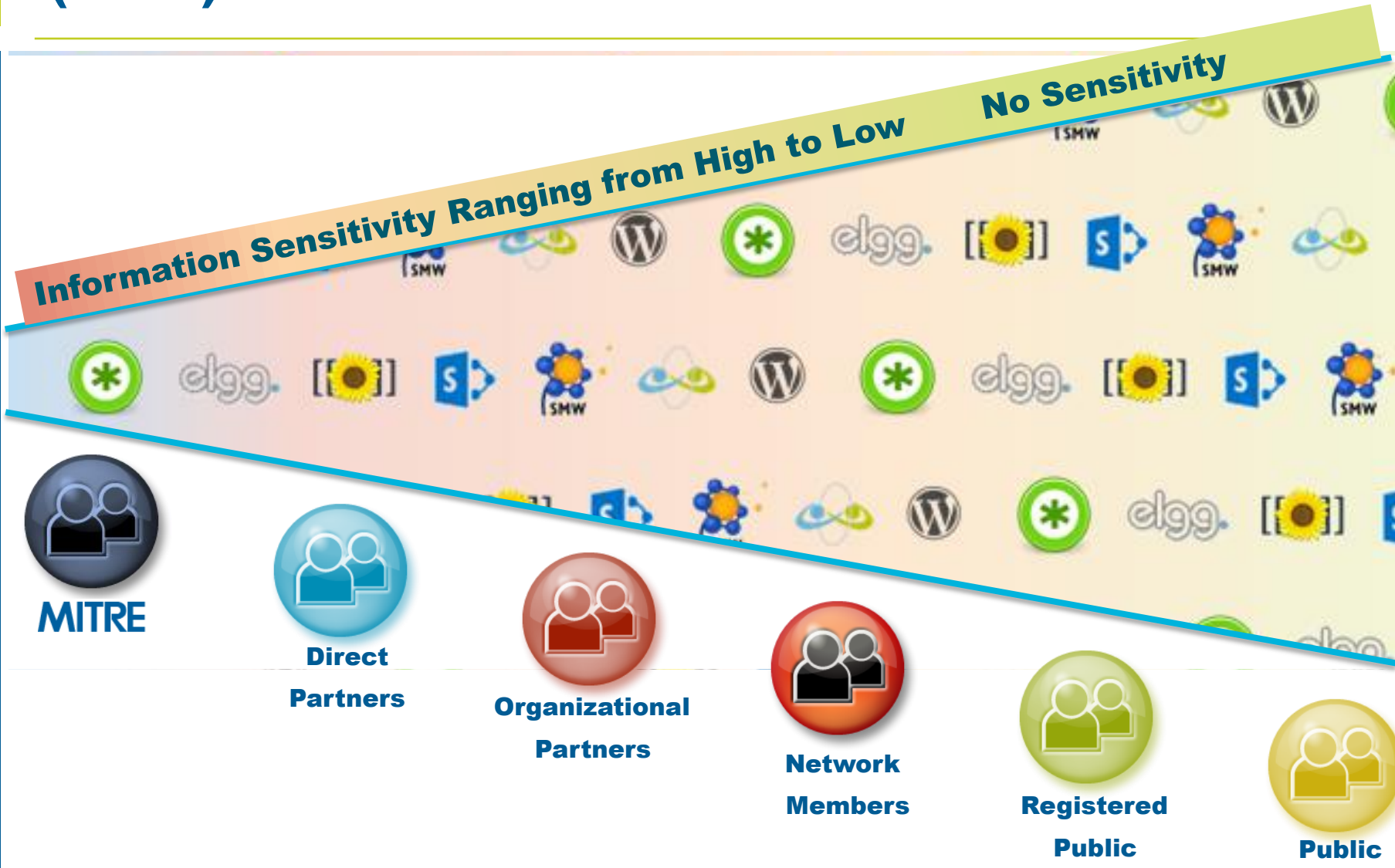
Wiki Farms for Multi-Organizational Sharing

- Content release requests are more complex for knowledge collections when they are moved from behind an enterprise firewall to an extended partnership network
- How can we preserve partner trust while providing increased value through broadened knowledge sharing?



- We are working to create tools in harmony with appropriate content release policies and procedures that make this capability a reality

MITRE Partnership Network (MPN)



Major Technical Requirements for MPN Deployment

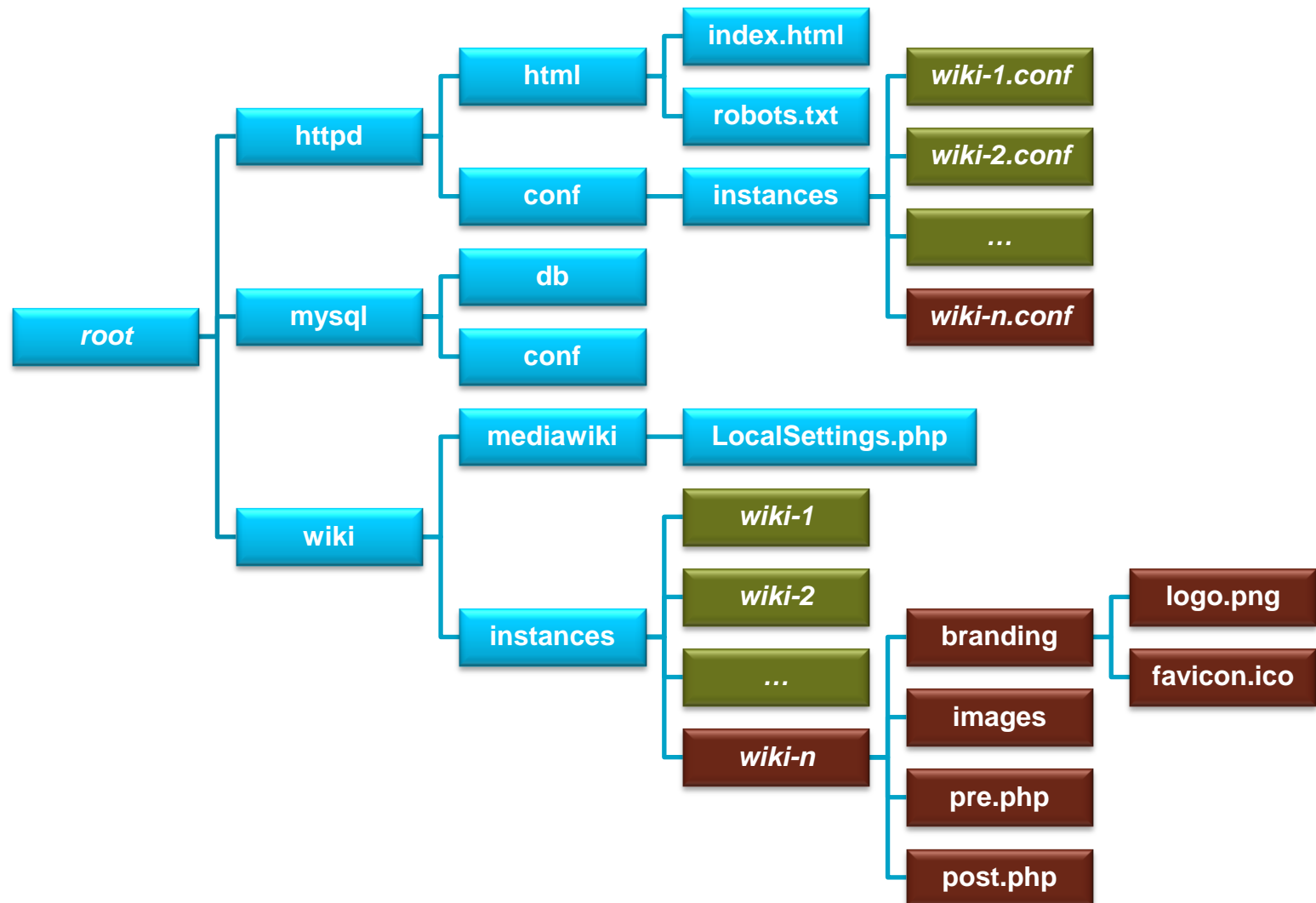
- **Internet-accessible wiki farm**
 - Must pass information security review
 - Must be easy to add new wikis
 - Authentication using OpenID Connect
- **Staging and development wiki farms behind the firewall**
 - Must be able easily to grab entire wiki farm (wiki software, configuration, and database) to create identical copy
 - All new extension development occurs on similar development wiki farm
 - Test all MediaWiki and extension upgrades on identical staging wiki farm before deploying to external wiki farm



OpenID Connect

- **Based upon OAuth 2.0**
- **Relies on JSON/REST**
- **Not wire compatible with OpenID**
- **Several Improvements over OpenID including:**
 - email address rather than URI as identifier
 - TLS for encryption
- **Final OpenID Connect specs launched February 26, 2014 by OpenID Foundation (<http://openid.net/>)**
- **MediaWiki OpenID Connect extension relies upon OpenID Connect PHP library available at <https://github.com/jumbojett/OpenID-Connect-PHP>**

Wiki Farm Movable File Structure



MPN Deployment Process

- **Configure virtual machine infrastructure**
- **Develop repeatable process for wiki farm construction**
- **Develop repeatable process for wiki creation**
- **Develop OpenID Connect authentication extension**
- **Develop OpenID Connect authorization extension**
- **Information security review**
 - Penetration testing revealed two cross-site scripting vulnerabilities that I submitted patches for:
 - MediaWiki core: info action ([CVE-2014-2853](#))
 - SemanticTitle extension ([CVE-2014-2854](#))
- **SUCCESS!**

Now the interesting research on sustained multi-organizational community knowledge building can begin!

Work In Progress and for the Future



- Release OpenID Connect authentication and authorization extensions as open source
- Develop more flexible authorization extensions
- Develop tools to support resource release
- Deploy more wikis to MPN
- Deploy public examples wiki to MPN
- Migrate subset of current behind-the-firewall content to MPN
- Wikishake: seamless user experience for integrated dialog and content authoring
- Medley: DevOps tools for wiki farm deployment and management (composer, vagrant, puppet, chef, . . .)
- Tools for monitoring and maintaining semantic model quality