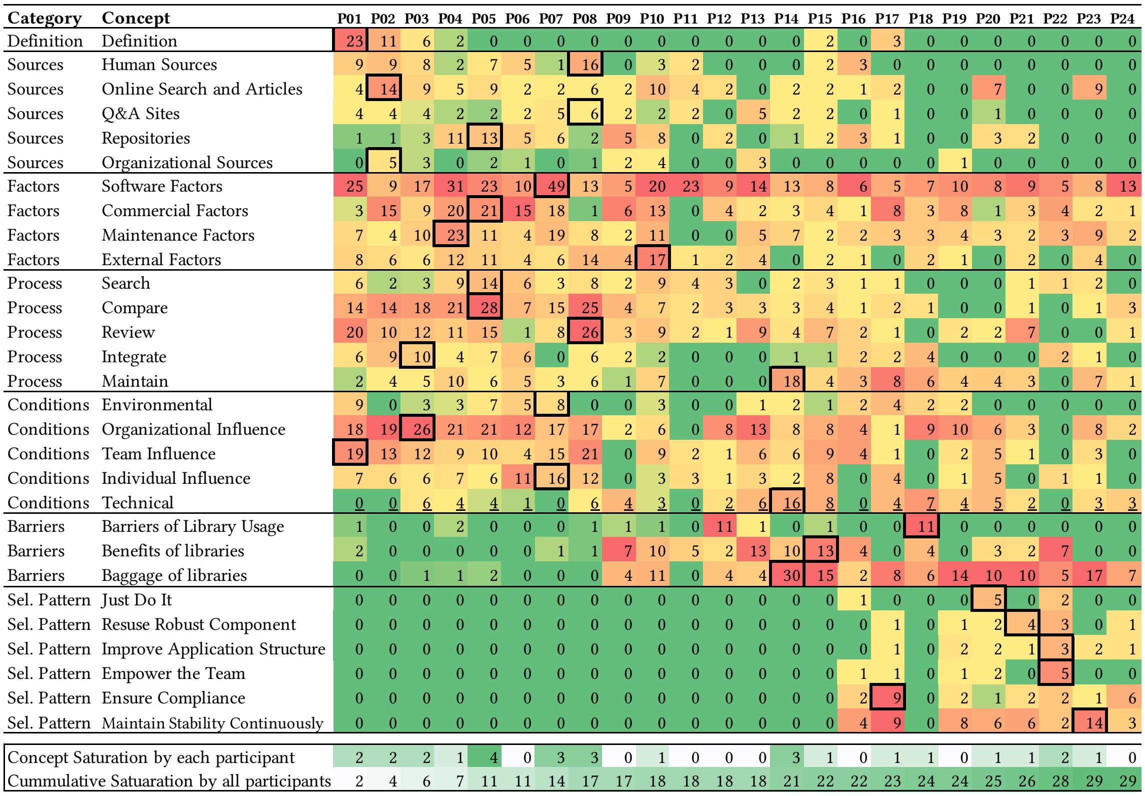
"How do people decide?": Rules for Third-Party Library Selection in Companies

ANONYMOUS AUTHOR(S)

ACM Reference Format:

Anonymous Author(s). 2024. *"How do people decide?"*: Rules for Third-Party Library Selection in Companies. In *CHI 2024: CHI Conference on Human Factors in Computing Systems, May 11–16, 2024, Honolulu, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/XXXXXXX.XXXXXXX



Category	Concept	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24
Definition	Definition	23	11	6	2	0	0	0	0	0	0	0	0	0	0	2	0	3	0	0	0	0	0	0	0
Sources	Human Sources	9	9	8	2	7	5	1	16	0	3	2	0	0	0	2	3	0	0	0	0	0	0	0	0
Sources	Online Search and Articles	4	14	9	5	9	2	2	6	2	10	4	2	0	2	2	1	2	0	0	7	0	0	9	0
Sources	Q&A Sites	4	4	4	2	2	2	5	6	2	2	2	0	5	2	2	0	1	0	0	1	0	0	0	0
Sources	Repositories	1	1	3	11	13	5	6	2	5	8	0	2	0	1	2	3	1	0	0	3	2	0	0	0
Sources	Organizational Sources	0	5	3	0	2	1	0	1	2	4	0	0	3	0	0	0	0	0	1	0	0	0	0	0
Factors	Software Factors	25	9	17	31	23	10	49	13	5	20	23	9	14	13	8	6	5	7	10	8	9	5	8	13
Factors	Commercial Factors	3	15	9	20	21	15	18	1	6	13	0	4	2	3	4	1	8	3	8	1	3	4	2	1
Factors	Maintenance Factors	7	4	10	23	11	4	19	8	2	11	0	0	5	7	2	2	3	3	4	3	2	3	9	2
Factors	External Factors	8	6	6	12	11	4	6	14	4	17	1	2	4	0	2	1	0	2	1	0	2	0	4	0
Process	Search	6	2	3	9	14	6	3	8	2	9	4	3	0	2	3	1	1	0	0	0	1	1	2	0
Process	Compare	14	14	18	21	28	7	15	25	4	7	2	3	3	3	4	1	2	1	0	0	1	0	1	3
Process	Review	20	10	12	11	15	1	8	26	3	9	2	1	9	4	7	2	1	0	2	2	7	0	0	1
Process	Integrate	6	9	10	4	7	6	0	6	2	2	0	0	0	1	1	2	2	4	0	0	0	2	1	0
Process	Maintain	2	4	5	10	6	5	3	6	1	7	0	0	0	18	4	3	8	6	4	4	3	0	7	1
Conditions	Environmental	9	0	3	3	7	5	8	0	0	3	0	0	1	2	1	2	4	2	2	0	0	0	0	0
Conditions	Organizational Influence	18	19	26	21	21	12	17	17	2	6	0	8	13	8	8	4	1	9	10	6	3	0	8	2
Conditions	Team Influence	19	13	12	9	10	4	15	21	0	9	2	1	6	6	9	4	1	0	2	5	1	0	3	0
Conditions	Individual Influence	7	6	6	7	6	11	16	12	0	3	3	1	3	2	8	0	4	0	1	5	0	1	1	0
Conditions	Technical	0	0	6	4	4	1	0	6	4	3	0	2	6	16	8	0	4	7	4	5	2	0	3	3
Barriers	Barriers of Library Usage	1	0	0	2	0	0	0	1	1	1	0	11	1	0	1	0	0	11	0	0	0	0	0	0
Barriers	Benefits of libraries	2	0	0	0	0	0	1	1	7	10	5	2	13	10	13	4	0	4	0	3	2	7	0	0
Barriers	Baggage of libraries	0	0	1	1	2	0	0	0	4	11	0	4	4	30	15	2	8	6	14	10	10	5	17	7
Sel. Pattern	Just Do It	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5	0	2	0	0
Sel. Pattern	Resuse Robust Component	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	4	3	0	1
Sel. Pattern	Improve Application Structure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	2	1	3	2	1
Sel. Pattern	Empower the Team	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	2	0	5	0	0
Sel. Pattern	Ensure Compliance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	9	0	2	1	2	2	1	6
Sel. Pattern	Maintain Stability Continuously	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	9	0	8	6	6	2	14	3
Concept Saturation by each participant		2	2	2	1	4	0	3	3	0	1	0	0	0	3	1	0	1	1	0	1	1	2	1	0
Cummulative Satuaration by all participants		2	4	6	7	11	11	14	17	17	18	18	18	18	21	22	22	23	24	24	25	26	28	29	29

Fig. 1. Heatmap of concept saturation over the interviews. Red refers to the higher discussion of a concept by an interviewee, the number refers to how many times the interviewee has discussed a concept. Green refers to the lower discussion of concepts by an interviewee. Thick black-bordered cells refer to the most discussion reference of a concept among all participants. For example, the review concept under the selection process category was discussed 26 times in P08's interview and hence the corresponding cell is marked by the thick border. The last two rows show the number of concept saturation by each participant and the cumulative number of concepts saturated till that participant. For example, the interview with P08 alone saturated 3 concepts (denoted by the thick boxes in the P08 column) and after the P08 interview, total of 17 concepts were saturated.

1 CONCEPT SATURATION OVER THE INTERVIEWS

An integral part of grounded theory is concept saturation when further data collection and analysis does not provide additional information about a concept. After each interview, we analyzed the transcript and performed open coding to identify all the concepts discussed in that interview. We also compared the discussed concepts and their dimensions with previous interviews. If we found that no new dimension of a concept was discussed consecutively for few interviews, we considered the concept saturated. Still, for the sake of uniform interview outline, we kept discussing the concepts in case any new corner case came up. Different concepts saturated over different periods of the interviews. For example, library selection factors were saturated by P10 (those concepts were most discussed in earlier interviews and gradually declined in later interviews). By the time we interviewed P24, no new concepts were discussed and all previous concepts were saturated.

Figure 1 shows how the concepts were discussed during each interview. The number denote how many times a concept was discussed by one particular interview. The more a participant discussed about a particular concept, the more red the corresponding cell is. For example, library search and analysis process was most discussed by P5 and after their interview, subsequently we did not have much to discuss about the search process. After P8, the concept almost saturated and we discussed very little about this concept with subsequent interviewees.

2 THEORETICAL SAMPLING FOR RECRUITMENT

We started with an architect (participant P1) from our professional network who had twelve years of experience including designing and developing from scratch a payment system in a 260 human-year project spanning over six years. We knew that they had to select a huge number of libraries throughout this development. Our first interview lasted 115 minutes, as we were developing initial concepts around all of our research questions, while subsequent interviews averaged 64 minutes. After analyzing the initial interview data, we realized that we needed more information about commercial factors (particularly licensing) and the integration process of a library. Hence, we chose our next participant (P2) who was working in a large organization (#engineers $\geq$ 80K), and who had experience with licensing concerns about software libraries. In this way, we continued the theoretical sampling to saturate all the relevant concepts. Though for every interviewee, we had particular target of concepts to explore deeply, we still discussed all the library selection process and concepts in general with all interviewees. We started selecting and evaluating the interviewees through convenience sampling from the direct professional networks. In the later phase of the interviews, we reached out to experts through snowballing and our extended professional networks. Before any interview, we conducted a screening call or email conversation to confirm the relevance of the interviewee regarding the specific concept we wanted to explore. The motivation for selecting every participant is provided in the Table 1.

3 SELECTION PATTERNS

In this section, we present the six full patterns associated with the library selection steps.

Table 1. Recruitment of interview participants following theoretical sampling for Concept Saturation. (Could not enhance targeted concepts from the (*) participants, rather enriched other important concepts.)

P#	Concept to enhance	Why we selected this Participant	Concepts they enriched significantly
P1	Initial process and factors	Architect of a large system	Library definition, factors, influences
P2	Licensing and Security Issues	Working in a large structured company	License, company technology
P3	Mobile development Factors	12+ years experienced in mobile application	Cost, company tech, comparison
P4	Long term maintenance concerns	Being a CEO, takes decisions considering long term impact	Company application domain, active development of library
P5	Decision making processes	Stablishing the processes in a startup team	Information search, company culture
P6	Open Source factors	Has experience regarding OSS contribution and research	Open source, Personal motivation
P7	Factors for a startup	Being a startup CTO may share different priorities	Flexibility, Ease of Installation, Community Support
P8	Performance factors	Working in a cloud company that may require high performing libraries	Familiarity, Team Discussion, Library Migration
P9	Migration scenarios	Experienced to migrate company tech stack as architect	Legal risks, Lack of Stability, Less prefered than native support
P10	Visualization and front end libraries	Working as web developer for over a decade	Customer support, flexibility, existing repository
P11	Machine learning libraries	Experienced in machine learning in gradudate research studies and in industry	Talk to people, Performance, Outstanding library selection
P12*	DevOps Process for Library Security Issues	Consulted dozens of companies in DevOps process establishment	Barriers of library usage, Baggage of libraries
P13	Selection process in large organizations for legal and security risks	Has been an architect in a large team for 10+ years	Consent Process, Benefits of libraries, Tech Expert Opinion
P14	Library migration scenarios	Experienced in managing mobile apps with large user base in all platforms	Make life easy, Life long maintenance, Migration to other library
P15	Organizational process and motivation for libraries	Experienced in organization process since increased dev team from 3 to 300	Delivery Deadline, Don't Reinvent the wheel, Feature criticality
P16*	Process of security concerns	Cerified security professional actively developing security products	License issues, Data Transfer Security, Geographic Impact
P17	Security Process	Delivers custom software to customers and maintains SecOps in CI/CD	Post Integration Maintenance for Security
P18	C++ libraries in large scale long term products	Leads development of a 30 year old product written in C++ with 2M lines of code	Lifelong Maintenance Burden, Compatibility, Uniform Coding Style
P19	Company Culture, Open Source, Concept Saturation	Experienced working in start-up and large organizations who open source libraries	Standard practices in large organizations, Considerations in open source
P20	Challenges in mobile application libraries, Concept Saturation	Full career in mobile app development, mostly in iOS which requires more maintenance	Lifelong Maintenance Burden, Abandoned Libraries, Migration
P21	Company Culture, Open Source, Concept Saturation	Works full-time in a prominent open core company	Company policies, Guiding Principles
P22	Guiding Principles, Open Source	Experienced in persuing large corporation for open source library adoption	Guiding Principles
P23	ML libraries	Working in South America in ML domain	ML Library Dependency Issues
P24	Company Culture, Industry, ML Libraries	Working in health sector using ML libraries extensively.	ML Library deployment and upgrade issues

Selection Pattern 1: Just Do It

Actor: Developers

Barrier: Lack of company policy or lack of developer experience can promote this short-term focused pattern.

Condition: Faster go to market is critical. Developers rewarded for delivering minimum feature on time.

Concern: How the developers can meet the deadline with relatively less effort?

Solution: Use a third-party library that reduces the work load and delivery can be done on time

Consideration: Easy to use, easy to install, popular, and familiar library.

Steps: Find libraries, compare them, and choose one according to the consideration. Take support from known people or use search engine or Stack Overflow. May skip team discussion, code review or maintenance plan.

Consequence: Can lead to future maintenance burden such as performance bottleneck or security vulnerability.

Example Trace: *"For startups, a lot of it [priority] is just speed to market and how much resources is gonna eat up using any specific library."* (P07)

Selection Pattern 2: Reuse Robust Component

Actor: Developers, Senior Developers, Architects

Barrier: Lack of learning culture or lack of developer experience can hinder adopting this pattern.

Condition: Mature organization, stable code. Developers concerned of quality, performance & maintainability.

Concern: How can the application avoid boiler plate code and follow best design principles?

Solution: Use a trusted proven third-party library that will keep the code clean and manageable.

Consideration: Open source, community trusted, stable and high-performing library.

Steps: Find and compare libraries, review thoroughly by more than one developer. Look into the library's source code repository to analyze the stability, quality of the library, and also consider reputed technical blogs.

Consequence: Stable, maintainable code. Too much exploration for small features can cause delivery delay.

Example Trace: *"So [this large corporation] as a whole is actually built on the open source libraries that are suitable for our use cases. But that actually has been one of our primary focus as well. If you find a library, use it; only build if you can't find anything."* (P13)

Selection Pattern 3: Avoid Structural Change

Actor: Developers, Architects
Barrier: Architectural limitation may create blocker against libraries and push for this conservative pattern.
Condition: Large scale software with custom framework. Library fits well and improves current architecture.
Concern: How can the architecture be improved or be protected from unwanted rigidity while using a library?
Solution: Flexible library, small sized, wrapping the library in API to allow future replacement.
Consideration: The library should be flexible and should not be too large in size compared to the system.
Steps: Find, compare, and review libraries. Conduct design review to assess impact on architecture. Review internal organizational content for design principles.
Consequence: Ensures long-term maintenance, less dependent on external changes. Avoids libraries if possible.
Example Trace: *"The moment you have to bring something in because there are new requirements is the time to assess how you've structured your application and does it still serve you and your customers or the business requirements you have. So that's the opportunity to look at the structural aspect of the application and make sure you do want to avoid changing it or maybe it's the time to change it."* (P22)

Selection Pattern 4: Empower the Team

Actor: Developers, Team Leaders
Barrier: Lack of inclusivity, lack of learning culture, or change-averse mindset can hinder adopting this pattern.
Condition: Strong company culture to promote transferable skill and provide learning space for developers. The development team may have limitations or strengths in certain technologies.
Concern: Does the library fit well within the capability of the dev team? Will it provide any transferable skills?
Solution: Use a library appreciated by the developers by providing transafarable skills.
Consideration: Well documented, popular, easy to use library with customer support.
Steps: Besides finding a library that fits well with the technology, thoroughly discuss with developers about their opinion and acceptance of the library. Look into official documentation of the library.
Consequence: May compromise the optimum technology for dev team's limitation. Can keep teams motivated.
Example Trace: *"So looking at community popularity helps because then it helps to hire people. It helps to retain people. They like to use technologies that are transferable."* (P19)

Selection Pattern 5: Ensure Compliance

Actor: Developers, Information Security Experts, Legal Experts, Open Source Program Office
Barrier: Lack of supporting process, required legal or security resources can hinder adopting this pattern.
Condition: Matured, regulated org or industry (health, finance). Presence of dedicated security/legal experts.
Concern: Any penalty or legal complication arising from using a library? How to protect the organization?
Solution: Use a library which is compliant with the application security standards and legal requirements
Consideration: License compatible with existing code and business. Secure, no known vulnerability.
Steps: Reach out to specialists in the organization for taking their expert consent before adopting the library. See the license and security declarations in the library documentation in the source code or package repository.
Consequence: Protects companies from security and legal crisis. May add approval delay during selection.
Example Trace: *"We had a very bad experience with this. With the legacy system, we were using so many different libraries and there is a licensing issue and we had to replace half of the library. Otherwise we had to pay lots of money. So that's why we are now very, very concerned about adding any external library, because if we don't comply with the license, it will be a legal problem."* (P09)

Selection Pattern 6: Maintain Continuous Stability

Actor: Developers, Operation teams, DevOps
Barrier: Lack of supporting process and lack of maintenance tool/process are barriers for adopting this pattern.
Condition: Long term application. Critical library update (vulnerability fix) or unmaintained library.
Concern: How to ensure a library is maintained in foreseeable future and developers can use smoothly?
Solution: Library with good history of maintenance and prepare to continuously upgrade the library in future
Consideration: Actively maintained library, supported by reputed organizations, and has larger community.
Steps: Analyze the maintenance and issue history of the library to assess the active development practices or the library. Establish a process for software bill of materials to document all third-party library dependencies and their upgrade plan in conjunction with DevOps teams. Look into source code commit and issue history from source repository and download usage trend from package repository.
Consequence: Requires dedicated resource upgradation for smooth operation. Otherwise can break system.
Example Trace: *"When we integrated the updated version our whole interface broke. And we had to change a lot of code, all the interceptors, interfaces, everything... This maintenance is quite hard. It's actually a full time work to always keep updated, to always stay updated."* (P14)