THE EMERGENCE OF BIG CODE

A 2020 Survey of Software Professionals

dimensional research

There has been a distinct shift in the way development organizations create code. Enter the era of Big Code. Big Code is all about how code is growing in:

- Volume: Exponential increases in the amount of code.
- Variety: Way more complexity in the languages, tools, and processes used for delivering software.
- Velocity: Accelerated delivery cycles that mean code is changing faster and being shipped virtually every day.
- Value: The reimagination of business models and practices through high-quality software.

Similar to how Big Data disrupted data teams, Big Code is disrupting development teams across all industries in challenging ways.

This report is based on a 2020 survey of more than 500 software development professionals across North America. This research examines the state of Big Code to quantify its complexity, understand its real impact on development and business outcomes, and identify what needs to be changed for enterprises to succeed. All survey participants had direct responsibility for software development at a company with more than 200 development staff.

Key Findings

- **Volume, variety, velocity, and value are all increasing**
  - 94% report they are affected by Big Code
  - Volume: 51% have more than 100 times the volume of code they had 10 years ago
  - Variety: More than 60% report a significant or dramatic increase in each of a number of different development dimensions including architecture, supported devices, use of open source, number of platforms, etc.
  - Velocity: 92% say pressure to release code faster has increased in the past 10 years
  - Value: 90% report the software their teams deliver has become more critical

- **Big Code creates big problems**
  - Challenges caused by Big Code include time for new hires to be productive (62%), code breaking because of lack of understanding of dependencies (57%), difficulties managing changes (50%), and more
  - 99% report Big Code has a direct impact on business outcomes of software development efforts
  - 74% report their teams avoid updating code due to fear of code changes breaking dependencies

- **Development teams need new tools to scale to Big Code environments**
  - 85% agree that existing tools were not designed for the era of Big Code
  - 99% would benefit from additional capabilities for searching enterprise code

dimensional research

Sponsored by

sourcegraph

Detailed Findings: Volume, variety, velocity, and value are all increasing
### Development teams across all industries are impacted by Big Code

In today's large and complex codebases, it's invariably problematic for developers to discover, understand, and fix code because of the significant increase in the volume and complexity of code. Big Code is often compared to Big Data because both are disrupting teams, depleting productivity, diminishing quality, and decreasing competitiveness in the marketplace.

Big Code

Definition: For this survey, **Big Code** refers to the dramatic growth in the volume and complexity of code. This includes increases in the variety of development environments, platforms, and tools; the velocity of delivery schedules; and the expected business value. Big Code is impacting development teams the way Big Data impacted data teams.

Big Code is one of the most pressing issues facing software organizations today. At its core, Big Code is all about how code is expanding significantly in volume, variety, velocity, and value.

Because software is a major part of every industry, Big Code is impacting almost all organizations with significant development efforts (94%).


Is your organization affected by Big Code?
No 6%
Yes 94%
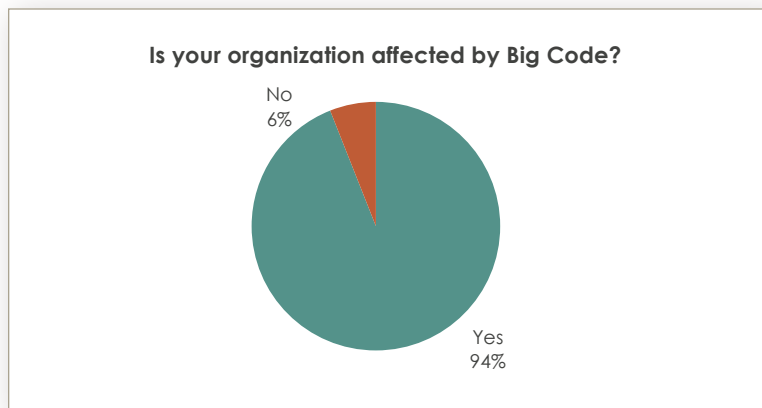
This response level is nearly identical across all organizations and industries no matter their size or the specific number of developers. Interestingly, industries outside of software report they are affected a bit more, with 95% saying they are impacted by Big Code compared to 92% of software companies. This data supports the current reality that Big Code is not just a software industry problem.
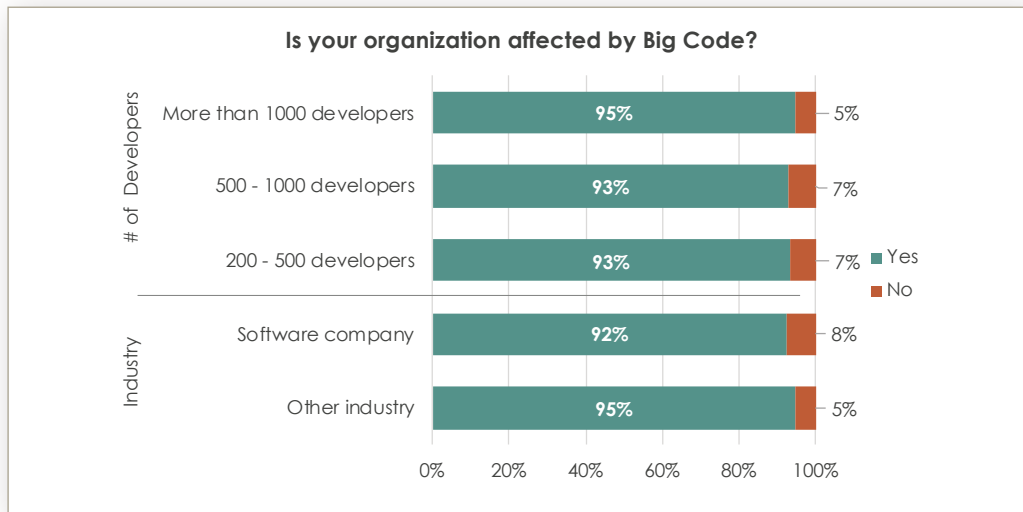
Volume is increasing

To better understand what's at stake, it's important to first quantify how big Big Code is. In other words, what volumes of code are organizations working with today compared to 10 years ago? When asked how the size of the codebase across their entire company, measured in megabytes and the number of repositories, has changed in the past decade, over half (51%) of software development stakeholders report they have more than 100 times the volume of code they had 10 years ago. And a staggering 18% say they have 500 times more code.
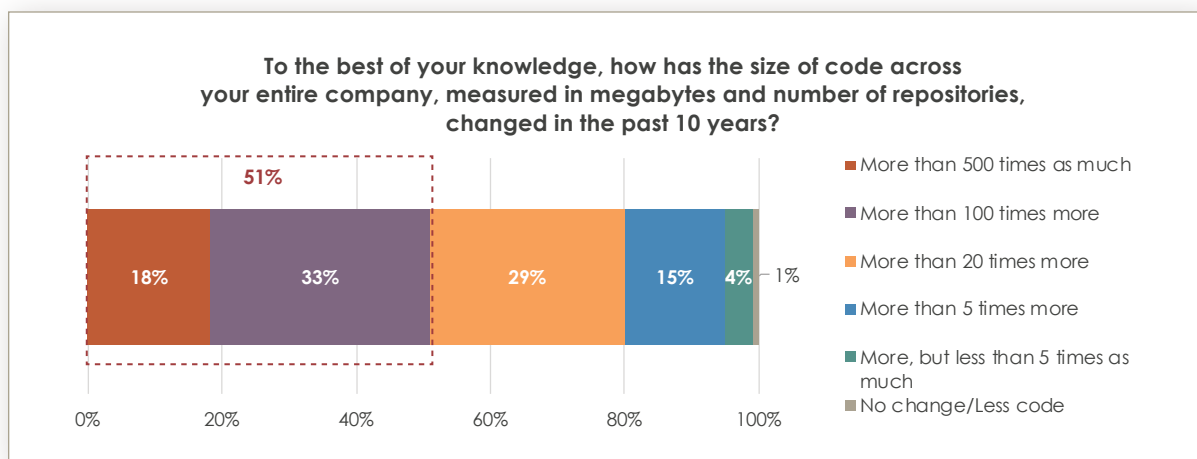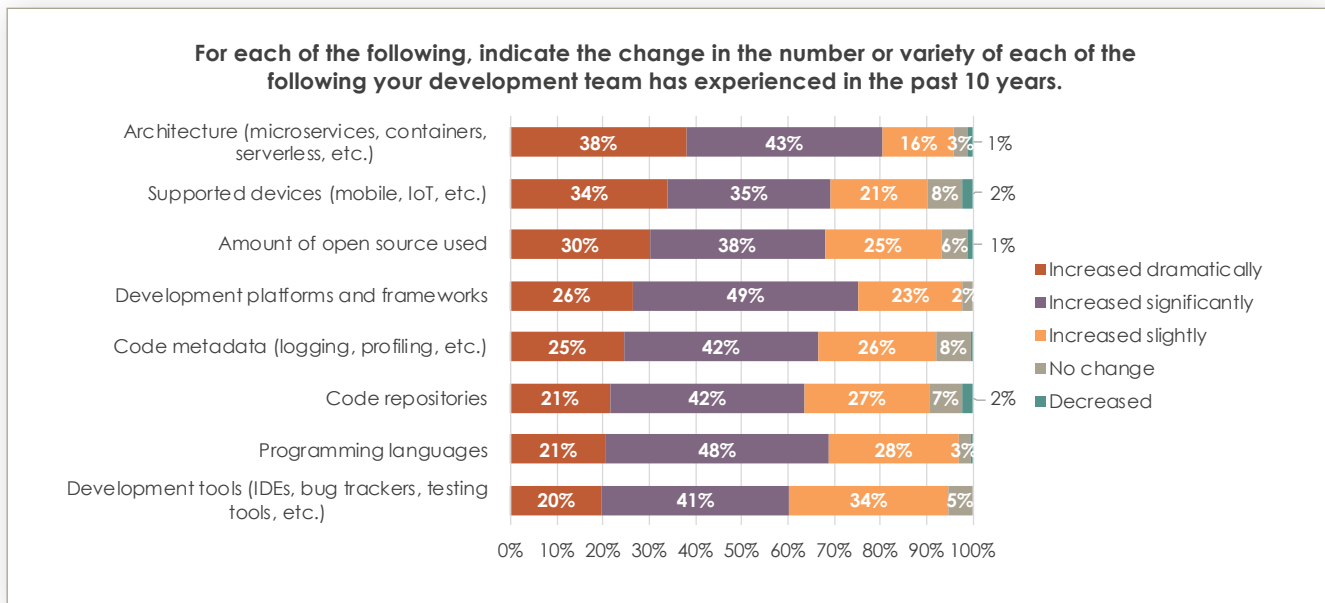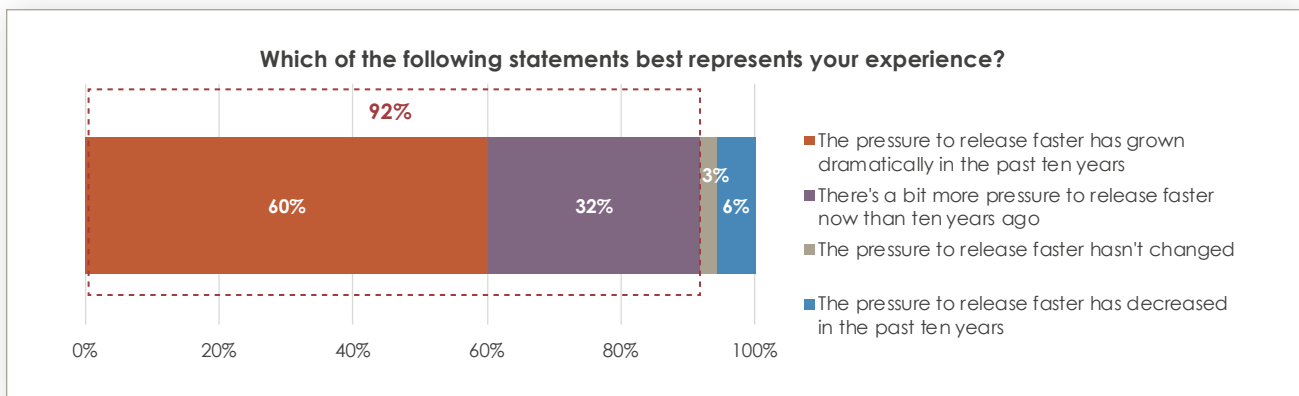
Variety is increasing

When examining current software, it's critical to note the profusion of different programming languages, code hosts, repositories, version control systems, architectures, services, supported devices, tools, APIs, and more that are contributing to code complexity and volume. Specifically, more than six in 10 software professionals report a significant increase across a wide range of different development dimensions in the past decade.

For each of the following, indicate the change in the number or variety of each of the following your development team has experienced in the past 10 years.

	Increased dramatically	Increased significantly	Increased slightly	No change	Decreased
Architecture (microservices, containers, serverless, etc.)	38%	43%	16%	3%	1%
Supported devices (mobile, IoT, etc.)	34%	35%	21%	8%	2%
Amount of open source used	30%	38%	25%	6%	1%
Development platforms and frameworks	26%	49%	23%	2%	
Code metadata (logging, profiling, etc.)	25%	42%	26%	8%	
Code repositories	21%	42%	27%	7%	2%
Programming languages	21%	48%	28%	3%	
Development tools (IDEs, bug trackers, testing tools, etc.)	20%	41%	34%	5%	

Velocity is increasing

In addition to the volume and variety of expansion within Big Code, the velocity of software releases is also rising rapidly. Through the implementation of such development approaches and practices as Agile, scrum, and DevOps, organizations can complete faster builds and swiftly deploy new releases. Almost all development stakeholders (92%) say there is increasing pressure over the past 10 years to release code faster.

Which of the following statements best represents your experience?

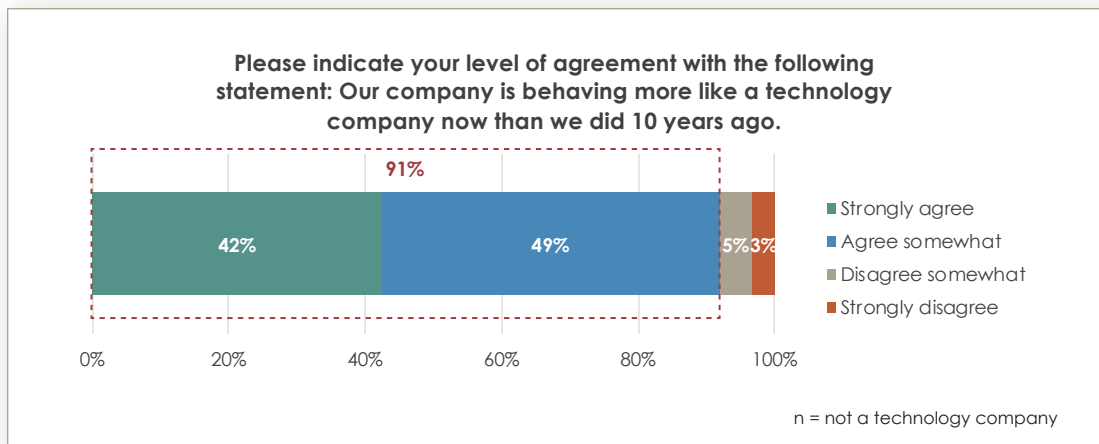| The pressure to release faster has grown dramatically in the past ten years	There's a bit more pressure to release faster now than ten years ago	The pressure to release faster hasn't changed	The pressure to release faster has decreased in the past ten years
60%	32%	3%	6%

92%

Value is increasing

Software is among many companies' most valued assets, propelling critical internal operations or generating direct licensing revenue. A bright finding of this research is the growing importance of software and how it is shaping the business. A whopping 90% of stakeholders say the software they develop has become more critical in the past decade.

Has the importance of the software your team develops changed in the past ten years?

No change 10%
Software is less critical 0%
90%
Software is slightly more critical 31%
Software is much more critical 59%
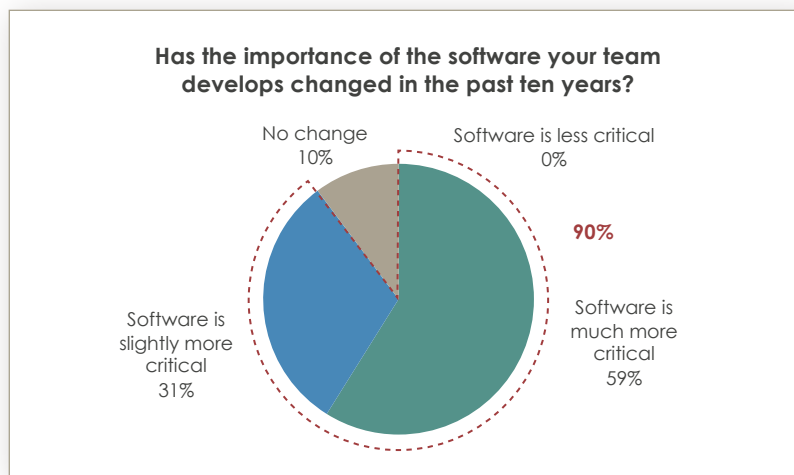
One of the interesting takeaways from the data is the extent to which companies outside of technology are now behaving more like technology companies. Software development stakeholders couldn't agree more, with 91% reporting their business is acting more like a tech company than it did 10 years ago.

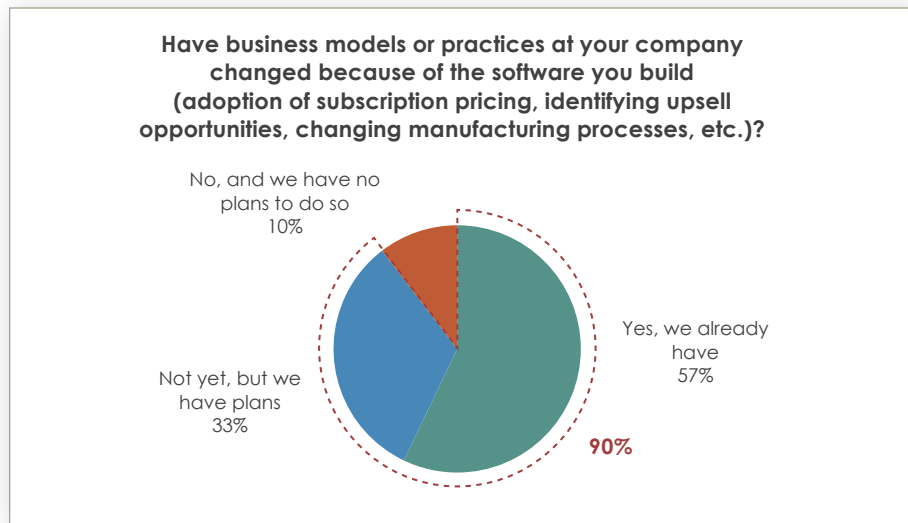**Please indicate your level of agreement with the following statement: Our company is behaving more like a technology company now than we did 10 years ago.**

91%
42% Strongly agree
49% Agree somewhat
5% Disagree somewhat
3% Strongly disagree

n = not a technology company

Furthermore, the increased value of software is not simply in attitude alone. Software is changing the business in meaningful ways, with 90% of stakeholders revealing how business models and practices have changed because of the software their teams build.

Have business models or practices at your company changed because of the software you build (adoption of subscription pricing, identifying upsell opportunities, changing manufacturing processes, etc.)?

No, and we have no plans to do so
10%

Not yet, but we have plans
33%

Yes, we already have
57%

90%

Detailed Findings: Big Code causes big problems
Development teams face many difficulties as a result of increasing complexity

Software supports the business. If the business grows increasingly complex, then its codebase will follow suit, since development mirrors the business and its key processes. When we asked software development stakeholders about the intricacy of their software, 77% say it is dramatically more complex now than 10 years ago.
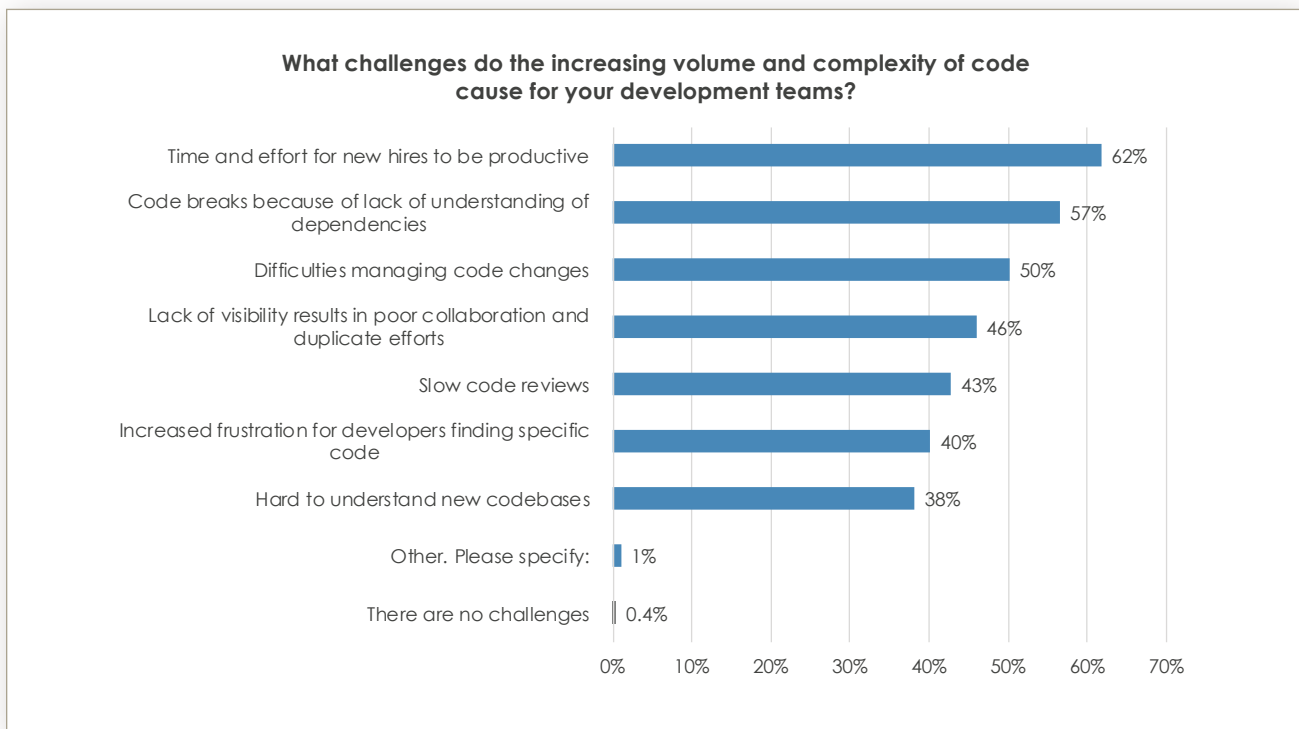
"Development is dramatically more complex now than it was 10 years ago."

Disagree
23%

Agree
77%
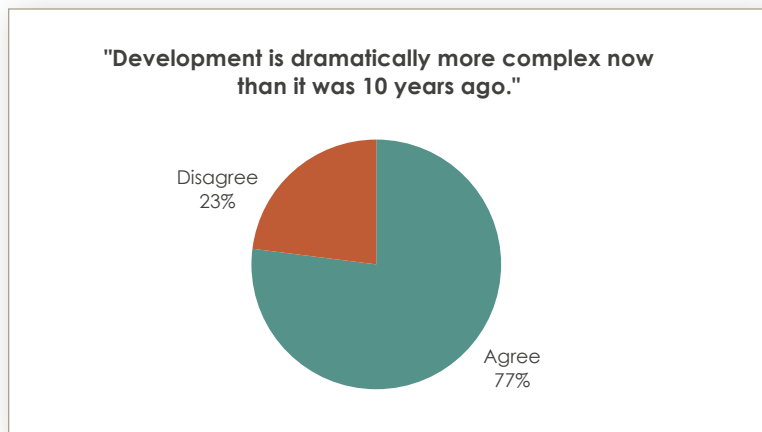
As expected, this dramatic rise in code complexity presents a wide range of challenges for development teams. The top challenge cited is the time and effort for new hires to be productive (62%), followed by code breaking because of a lack of understanding dependencies (57%), difficulties managing code changes (50%), lack of visibility (46%), slow code reviews (43%), increased frustration for developers finding specific code (40%), problems understanding new codebases (38%), and much more. A few individuals wrote in "other" challenges including maintaining quality, lack of testing knowledge, the learning curve of new technologies, and processing time.

What challenges do the increasing volume and complexity of code cause for your development teams?

| Challenge	Percentage
Time and effort for new hires to be productive	62%
Code breaks because of lack of understanding of dependencies	57%
Difficulties managing code changes	50%
Lack of visibility results in poor collaboration and duplicate efforts	46%
Slow code reviews	43%
Increased frustration for developers finding specific code	40%
Hard to understand new codebases	38%
Other. Please specify:	1%
There are no challenges	0.4%

Big Code complexity directly impacts the business

With billions of lines of code, multiple system interfaces, and complex requirements, the complexity of software can test the development team's control and result in unwanted business consequences. Big Code complexity is a substantial concern for those teams managing numerous technologies and applications. In this study, 99% of development stakeholders correlate code complexity directly to business impact. According to stakeholders, the primary negative business outcomes related to Big Code are difficulty maintaining high quality standards (57%), high security risks (51%), delayed release schedules (49%), increased costs (47%), less agility (46%), and increasing difficulty to innovate (36%).

How does increasing volume and complexity of code impact development and business outcomes?

Category	Value
Hard to maintain high quality standards	57%
High security risks	51%
Delayed release schedules	49%
Increased costs	47%
Decreased agility	46%
Difficult to innovate	36%
Hard to scale successes	34%
Low user and customer satisfaction	20%
Other. Please specify:	0.4%
There is no impact on outcomes	1%

Big Code complexity personally impacts developers

On a more personal level, nearly every new release of complex software stirs unwanted feelings, with 88% of software development teams admitting that every release causes some anxiety.
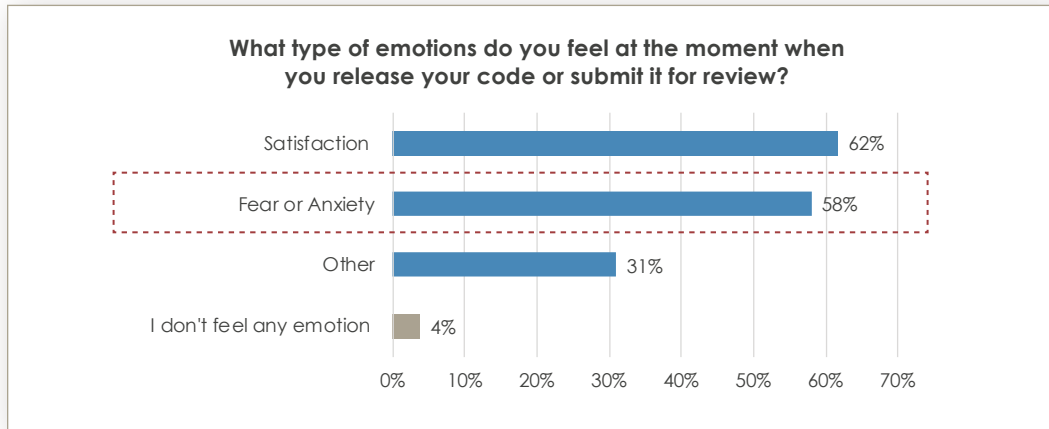
"Every release causes some anxiety."

| Response	Value
Disagree	12%
Agree	88%
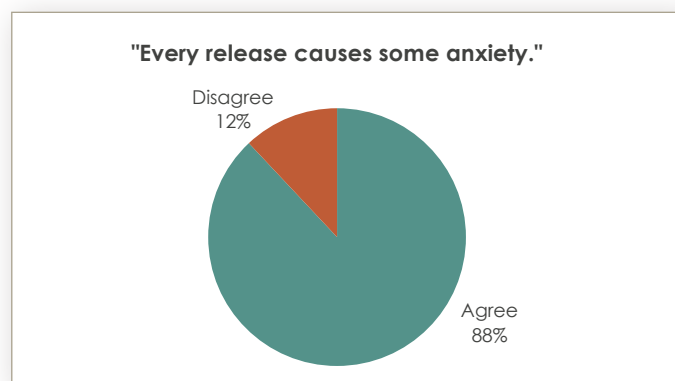
When we dive deeper, developers experience a wide range of emotions when releasing code. Almost all development teams (96%) reveal that code releases are emotional. While many report positive feelings like satisfaction, it is alarming to note that well over half (58%) say they feel negative emotions, including fear and anxiety, at the moment they release code or submit it for review.

What type of emotions do you feel at the moment when you release your code or submit it for review?

Emotion	Percentage
Satisfaction	62%
Fear or Anxiety	58%
Other	31%
I don't feel any emotion	4%

Teams avoid updating code because they fear breaking dependencies

One of the most somber revelations of this survey is how this fear can affect development progress. Three-quarters (74%) of development stakeholders say their teams avoid updating code because they are not sure of the dependencies and fear they might break something.

Does your development team ever avoid updating code because they're not sure of the dependencies and fear they might break something?
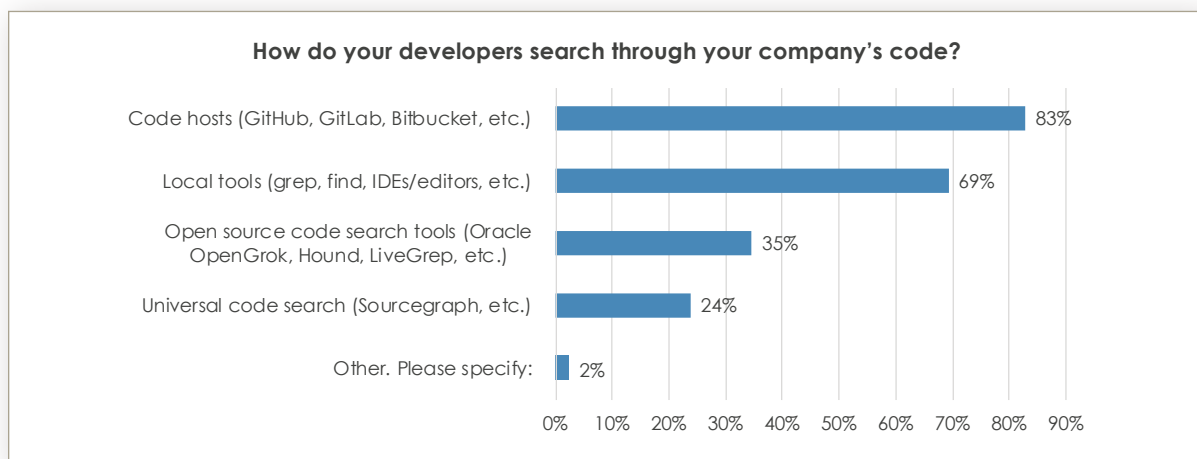
- No 26%
- Yes 74%
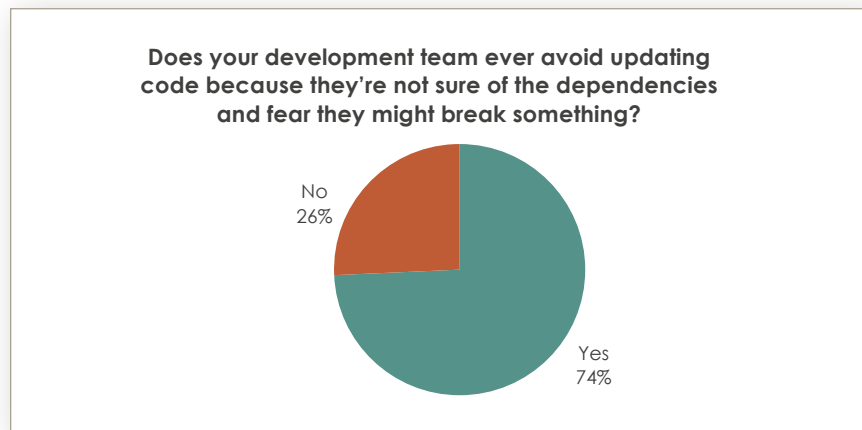
Detailed Findings: Development teams need new tools to scale to Big Code environments
Existing tools were not designed for the era of Big Code

Editors and IDEs were built for individual developers working in small teams on a sole repository. But in today's environment, developers are now regularly collaborating on code at a massive scale. To meet the requirements of the business and remain competitive, they must move fast, even in big codebases. But can their current tools scale to Big Code environments? An overwhelming majority of software developers (85%) are in agreement that existing tools are not designed for working with large codebases at scale.

"Existing tools were not designed for the era of Big Code."

Disagree
15%

Agree
85%
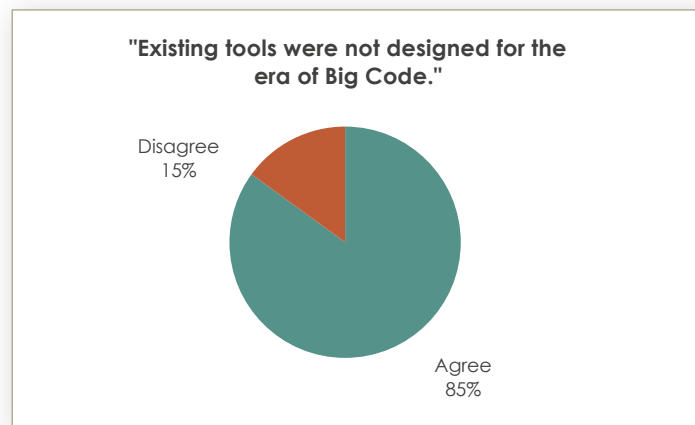
In particular, how do developers search through massive codebases to efficiently write and make changes to their enterprise's code while meeting tight deadlines under strict quality and security requirements? When asked about the technology used to search through company code, software teams say they rely on code hosts (83%), local tools (69%), open source code search tools (35%), universal code search (24%), and more. Several individuals took the time to write in other responses including Team Foundation Server (TFS), Azure DevOps, ClearCase, homemade solutions, internal code hosts, Jira, Resharper, SCCS, and Codeproject.

How do your developers search through your company's code?

Response	Percentage
Code hosts (GitHub, GitLab, Bitbucket, etc.)	83%
Local tools (grep, find, IDEs/editors, etc.)	69%
Open source code search tools (Oracle OpenGrok, Hound, LiveGrep, etc.)	35%
Universal code search (Sourcegraph, etc.)	24%
Other. Please specify:	2%

Upon closer examination, there is an alarming gap between what management thinks their developers use for code search and what developers report using. Among executives, 50% say their developers use universal code search, the most sophisticated approach to searching code, while only 13% of the developers using tools report the same.

How do your developers search through your company's code?

Category	Developer	Manager	Executive
Code hosts (GitHub, GitLab, Bitbucket, etc.)	82%	84%	83%
Local tools (grep, find, IDEs/editors, etc.)	68%	71%	68%
Open source code search tools (Oracle OpenGrok, Hound, LiveGrep, etc.)	20%	38%	55%
Universal code search (Sourcegraph, etc.)	13%	20%	50%
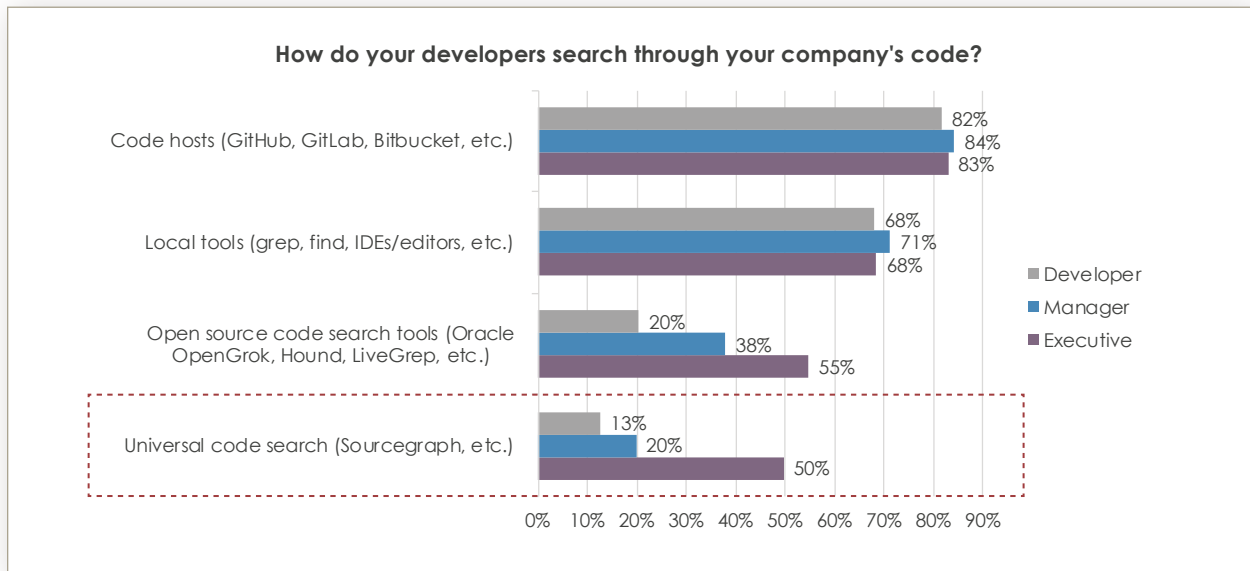
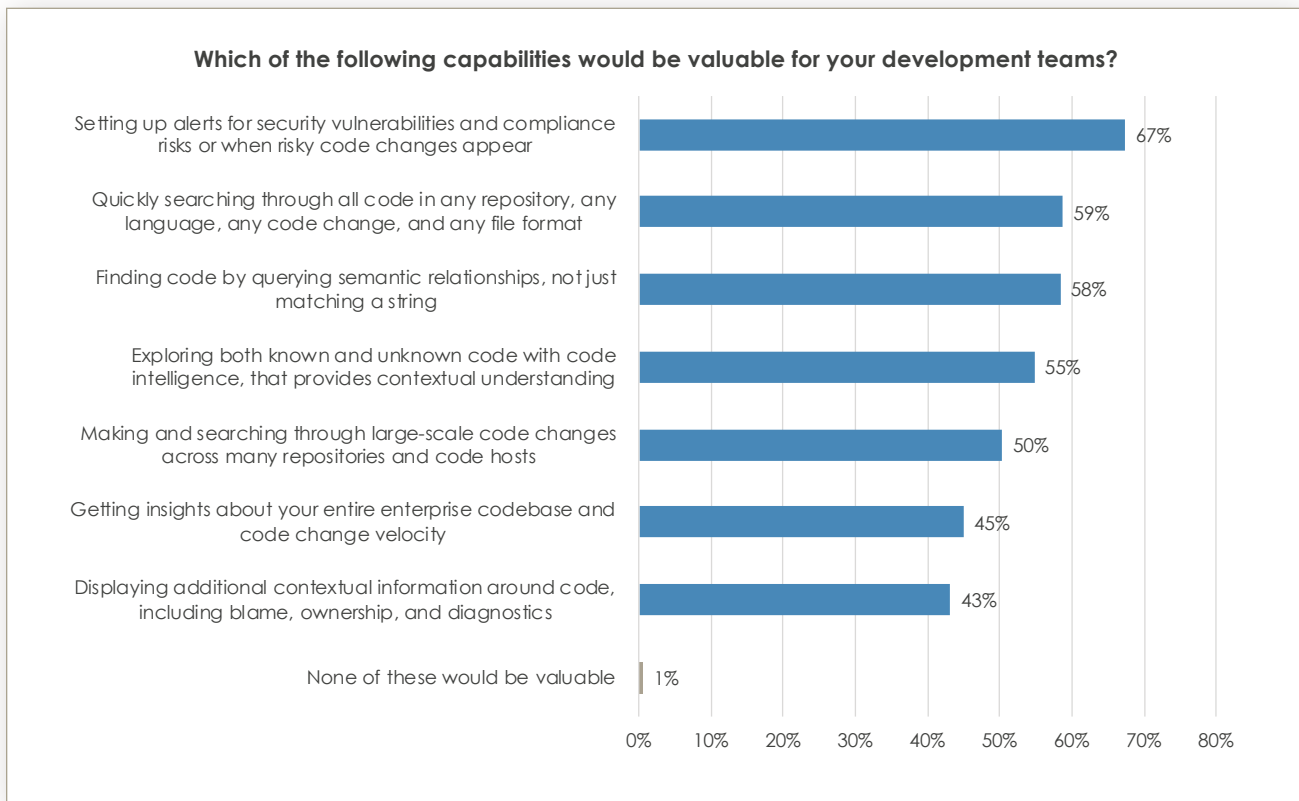## Development teams would benefit from additional capabilities for searching code

While management and developers are out of sync about which search tools they are currently using, they agree about wanting new technologies with the potential to transform business in the near term. Almost all (99%) development stakeholders say they would benefit from additional capabilities for searching code.

Two-thirds (67%) say the ability to set up alerts for security vulnerabilities and compliance risks would be the most valuable for their development teams. Other top capabilities cited by stakeholders are quickly searching through all code in any repository, any language, any code change, and any file format (59%), finding code by querying semantic relationships (58%), exploring both known and unknown code with contextual understanding (55%), and making and searching through large-scale code changes across many repositories and code hosts (50%).

Which of the following capabilities would be valuable for your development teams?

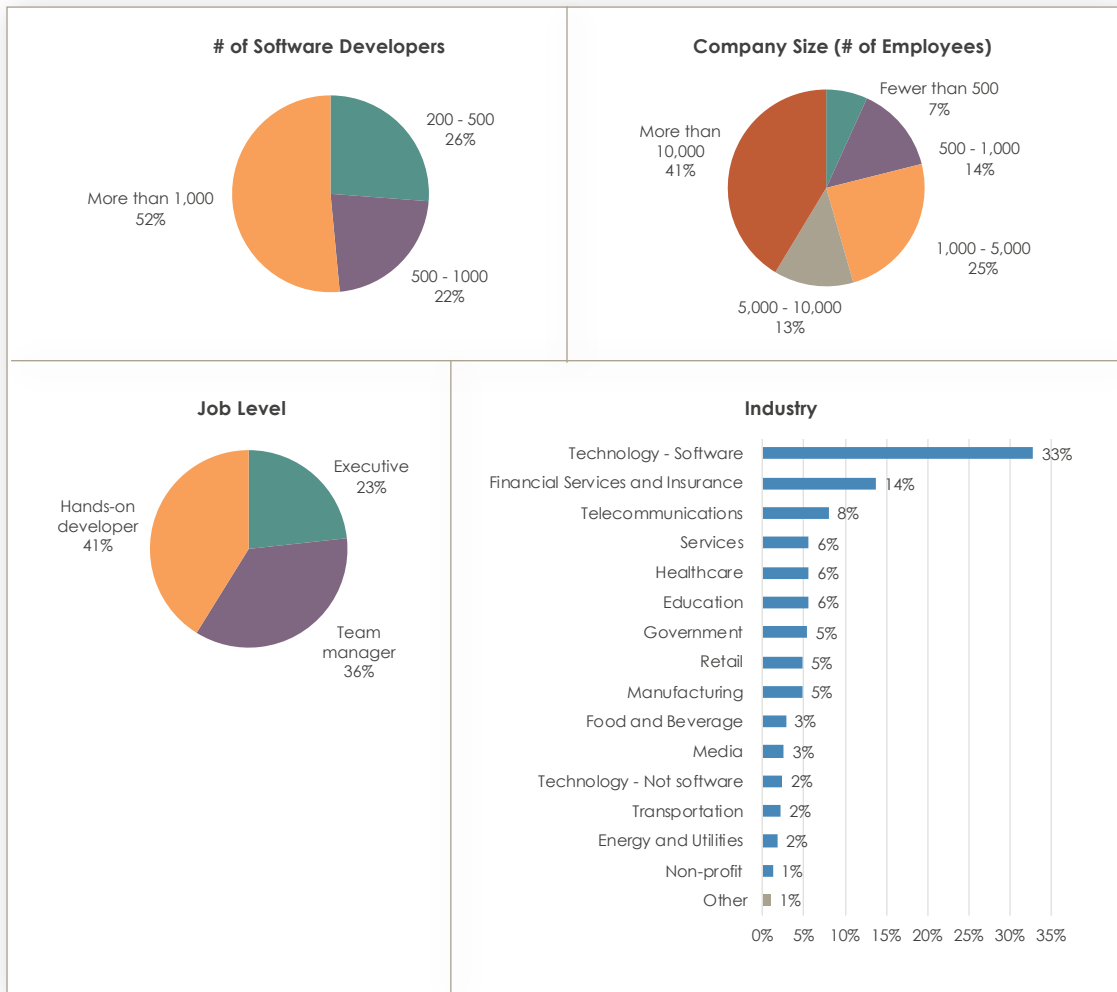| Capability	Percentage
Setting up alerts for security vulnerabilities and compliance risks or when risky code changes appear	67%
Quickly searching through all code in any repository, any language, any code change, and any file format	59%
Finding code by querying semantic relationships, not just matching a string	58%
Exploring both known and unknown code with code intelligence, that provides contextual understanding	55%
Making and searching through large-scale code changes across many repositories and code hosts	50%
Getting insights about your entire enterprise codebase and code change velocity	45%
Displaying additional contextual information around code, including blame, ownership, and diagnostics	43%
None of these would be valuable	1%

Survey Methodology and Participant Demographics

In June 2020, an online survey was sent to an independent database of software development stakeholders in the United States and Canada. A total of 503 qualified individuals completed the survey. All participants had direct responsibility for software development at a company with more than 200 development staff. Participants included a mix of job levels, company sizes, and industries.

# of Software Developers

- 200 - 500: 26%
- More than 1,000: 52%
- 500 - 1000: 22%

Company Size (# of Employees)

- More than 10,000: 41%
- Fewer than 500: 7%
- 500 - 1,000: 14%
- 1,000 - 5,000: 25%
- 5,000 - 10,000: 13%

Job Level

- Hands-on developer: 41%
- Executive: 23%
- Team manager: 36%

Industry

- Technology - Software: 33%
- Financial Services and Insurance: 14%
- Telecommunications: 8%
- Services: 6%
- Healthcare: 6%
- Education: 6%
- Government: 5%
- Retail: 5%
- Manufacturing: 5%
- Food and Beverage: 3%
- Media: 3%
- Technology - Not software: 2%
- Transportation: 2%
- Energy and Utilities: 2%
- Non-profit: 1%
- Other: 1%

About Dimensional Research

Dimensional Research® provides practical market research to help technology companies make their customers more successful. Our researchers are experts in the people, processes, and technology of corporate IT. We understand how technology organizations operate to meet the needs of their business stakeholders. We partner with our clients to deliver actionable information that reduces risks, increases customer satisfaction, and grows the business. For more information, visit dimensionalresearch.com.

About Sourcegraph

Sourcegraph Universal Code Search empowers all developers to explore, navigate, and better understand all code, everywhere, faster. Sourcegraph's mission is to make it easier and faster for developers to work on solving problems. Sourcegraph is built for all companies, from startups needing a solid foundation for growth to enterprises with complex security, scaling, and deployment needs. The all-remote company is headquartered in San Francisco, CA, and backed by Craft Ventures, Felicis Ventures, Redpoint Ventures, and Goldcrest Capital. Follow Sourcegraph on Twitter at @srcgraph or visit about.sourcegraph.com.