BACHELOR

How Age in Software Development is perceived on Reddit

van der Elsen, Joris

Award date:
2022

Awarding institution:
Tilburg University

Eindhoven University of Technology
Department of Mathematics and Computer Science

How Age in Software Development is perceived on Reddit

Bachelor Thesis

Joris van der Elsen - 1014570

Supervisors:
Alexander Serebrenik
Sebastian Baltes

Eindhoven, March 2022

Abstract

Ageism is seen as a major issue in the software development industry, as evidenced by multiple recent lawsuits. Previous research on age in software development was centered around the effects of aging and age stereotypes. Most of the studies focused on how age software developers are perceived in their workplace. It is less well researched how the general public perceives age in software development. Though the social news website Reddit, this study uses the topic modelling method LDA to determine what the general public's perception is on age in the software development industry. It finds that a LDA analysis on the gathered Reddit data does not present in coherent, useful topics regarding age in software development.

Contents

Introduction

The world's population is ageing. Almost all countries see an increase in the number and proportion of older people in their population. The United Nations labels ageing as "one of the most significant social transformations of the twenty-first century" [15]. According to Krampe and Charness, "declining birth rates and continued increases in life expectancy in industrialized countries have prompted the rediscovery of older adults in their 60s and even 70s as valuable participants in the workforce" [11]. As a result, the proportion of software developers aged 55 years or older in the USA has grown from 8.3% to 14.0% of all software developers in 2020 [17].

While older workers are becoming more prominent in the information technology industry, several big tech companies are facing lawsuits against alleged age discrimination. These companies reportedly have performed illegal actions against older employees, often the firing of the latter [20][12]. Previous research has found that older IT workers, such as software developers, face similar challenges, like being pressured to retire or become a manager [4]. This ageism is seen as a major barrier in the process of hiring developers in large software companies [2].

These events raise the question how older workers are perceived in the software development industry. Prior studies on older developers in software development have mainly focused on the effects of ageing on the abilities of developers in the workplace. Research has also been conducted on how software developers themselves view older developers. Baltes, Park, and Serebrenik have investigated how older software developers are perceived in the public discourse. They studied a number of popular online articles about ageing and software development, as well as discussions about those articles on the social news website Hacker News [2].

Another social news website on which matters around software development are discussed is Reddit [21]. Reddit attracts a broader, more diverse audience compared to Hacker News, implying that the people discussing, viewing and sharing the discourse on Reddit concerning age in software development are less likely to be software developers themselves. This could give a good perception of how older developers are perceived by the general public.

Research Questions

Therefore, the research questions for this study are as follows:

To better understand how the general public perceives the age of developers in the software industry, the following question must be answered:

> **RQ1 Which age-related topics concerning software development are discussed on Reddit?**

On top of that, to determine the most prevalent of these perceptions of age in software development, the following question must be answered:

> **RQ2 Which topics concerning age(ism) in software development are most frequently discussed on Reddit?**

To assess whether the found perceptions can support or contradict the conclusions of previous literature, and to possibly find novel perceptions not yet mentioned in previous literature, the following question must be answered:

> **RQ3 How do the topics concerning age(ism) in software development discussed on Reddit compare to the topics concerning age(ism) in software development found in literature?**

Answering these research questions can be used to inform employers and managers in the software development industry about the issues their older developers may encounter. They can also contribute to understanding these challenges and how to potentially counter them.

Related Work

Previous literature in the field on age(ism) in the workplace has primarily focused on the effects of ageing and age stereotypes. Furthermore, Ródriguez-Pérez et al. performed a systematic literature review on diversity within the software engineering field [22].

Age(ism) in the Workplace

Krampe and Charness found that general cognitive abilities of aging experts, as measured through psychometric tests, may get worse over time. However, this does not necessarily have a negative effect on the experts' performance in their field of expertise. Moreover, older experts do not have to rely on the success of initial learning, as they can maintain specific skills through deliberate practice [11].

Van Dalen et al. found that the productivity of older workers is perceived to be substantially lower than that of younger workers by both employers and employees [6]. These perceptions of productivity are based on stereotypes about hard qualities and soft qualities of workers. Hard qualities include flexibility, physical and mental capacity, and the willingness to learn new technology skills. Soft qualities include commitment to the organization, reliability, and social skills. Van Dalen et al. noted that "the comparative advantage of older workers (aged 50+) lies primarily in their soft qualities, whereas the comparative advantage of younger workers lies primarily in their hard qualities."

Posthuma and Campion also found several negative stereotypes towards older employees, including that they have lower ability, are less motivated, and are less productive than younger workers [18]. Another stereotype is that older workers are harder to train, less adaptable, less flexible, and more resistant. Therefore, they would provide a lower return on investments such as training. Posthuma and Campion also found stereotypes that older employees have a lower ability to learn and that they are more costly. A positive stereotype, on the other hand, is that older employees are as more stable, dependable, honest, trustworthy, loyal, committed to the job, and less likely to miss work or turnover quickly. These age stereotypes are found to be stronger in the information technology/computing industry, among other industries [18].

Age(ism) in Software Development

Schloegel et al. [23] found several age stereotypes towards older employees in the software development industry, like being less creative and less productive, although this 'poor performance stereotype' is refuted by Posthuma and Campion, as they noted that employees' job performance is affected more by individual skill and health than by employee age [18]. Murakami et al. found that the age of code reviewers does not affect the efficiency and correctness of their reviews [14]. On the contrary, in a study conducted by Baltes and Diehl, developers noticed a decline of programming performance by themselves [1].

Not all stereotypes are negative though. As found by Davidson et al., developers associate ten unique benefits with older adults participating in communities around free/open source software such as: having a wealth of software development and professional experience, having seen and understand technology trends, having life experience as a user, parent, and spouse, and having general wisdom and maturity [7]. Other age stereotypes towards older employees in the software development industry are being more reliable and having more know-how [23]. These stereotypes are supported by the findings of Morrison et al., that suggest veterans have broader knowledge of

software development in general and of old technologies that can be reinvented into newer technologies in particular [13].

In addition to stereotypes, several previous studies have focused on the motivations for older software developers. These motivations are perceived to be different from their younger peers. Using data from the website StackOverflow, the findings of Morrison et al. suggest that veteran developers are less motivated by social interactions than their younger colleagues [13]. Davidson et al. found that older developers contributing to free/open source software were mostly motivated by altruism, community identification and intrinsic motivation [8]. Older adults that were first-time contributors to free/open source software had similar motivations (intrinsic motivation, altruism, and internal values) [7].

Nevertheless, IT workers in their fifties and sixties were experiencing pressure to retire [4]. Whereas younger workers were characterized by technical specialisation, older workers were often expected to switch to broader management roles. They may prefer to participate as consultants because they face some confidence barriers [10]. Older developers were perceived to have obsolescent skills and to be unable to keep up with new developments [4]. Older workers experienced conflict between learning and applying new skills on one hand, and family responsibilities on the other [5].

The action to switch to a management role was one of the employability strategies for older developers identified by Baltes, Park, and Serebrenik on the social news website Hacker News. Some other examples of the mentioned strategies include specialization, networking, and mastering modern technologies [2]. The findings of Comeau and Kemp imply that ageist stereotypes are likely to hinder the longevity and value of older workers in these fields. As they state, "The retreat of the experienced workers, the IT 'veterans' who have lived through the industry's downturn and recovery, is a considerable loss of human resources to the industry and for the future development of computing technology." [5]

Methodology

Data Extraction

Reddit is a popular social news website. It is a forum on which a large variety of subjects are discussed, among which is software development. For this reason, Reddit attracts a broader, more diverse audience than for instance Hacker News. This implies that Reddit users engaging with discourse about age in software development are less likely to be software developers themselves. Thence, data from Reddit could give a good perception of how age in software development is perceived by a more general public.

Reddit consists of almost three million sub-fora, which are called 'subreddits'. Each of these communities often has a common subject that is discussed in its submissions. To identify relevant subreddits in which topics around 'age in software development' are discussed, two queries were performed using the default search bar on Reddit's website. These queries were "software development" and "software developer". The results of these queries were then filtered; only English-speaking subreddits directly related to software development were included. Furthermore, humorous subreddits (like 'programmingcirclejerk' and 'programmerhumor') were excluded. Only publicly available data were used, as some subreddits (like 'programming' and 'coding') did not allow data to be automatically extracted. This resulted in a total of 20 subreddits that were used for the analysis. The list of included subreddits can be found in Appendix A.

PRAW (Python Reddit API Wrapper) [19] was used to scrape the data from Reddit. PRAW enables its users to extract data from specified subreddits. Additionally, it is possible to search these subreddits for submissions relevant to key words chosen by the user. For this analysis, the query with the key words "age software developer" was performed on the selected subreddits. The result was 3485 relevant submissions.

Data Preprocessing

After the submissions were extracted, they needed to be preprocessed to prepare them for the actual analysis. Only the text of the submissions was used in the analysis.

First, a simple preprocessing was performed, which included tokenization, lowercasing the words and removing punctuation. Next, frequently used words called stop words were removed from the data. For this, the stop word list provided by *nltk* was used [16]. Some other frequently occurring words (like "would", "should", "could") appeared in results of earlier versions of the analysis. A list of these additional stop words can be found in Appendix B. Multiple words were manually added to the stop word list to get more insightful results. In order to further preprocess the data, the tokenized words were lemmatized and stemmed. The used lemmatizer was the WordNetLemmatizer and SnowballStemmer is the stemmer that was used, both provided by *nltk*. To reduce the noise in the analysis, all the tokens that appeared in more than 80% of the documents were removed.

Topic Modelling

To answer RQ1 *(Which age-related topics concerning software development are discussed on Reddit?)* and RQ2 (*Which topics concerning age(ism) in software development are most frequently discussed on Reddit?*), an analysis based on the concept of topic modelling was performed. Topic modeling is a method for representing each document in a corpus as being generated by a variety of distinct topic [9]. In this study, the corpus is the dataset with all the submissions from Reddit, in which every submission is a document. Each topic consists of a weighted set of words.

Latent Dirichlet Allocation

One of the most widely used topic modelling methods is Latent Dirichlet Allocation (LDA). LDA was first introduced in its current form by Blei, Jordan and Ng in 2003 [3]. Each document is represented as random mixtures over various latent topics, each of which refers to a probability distribution over words. LDA is a generative probabilistic model that assumes that the topic in all documents and the word in all topics are generated following the Dirichlet distribution. LDA was chosen because it is easily implementable and gives better results than more basic models like tf-idf.

Given a number of topics, the LDA model generates for each topic 10 words that contribute to that topic. Every word is weighted reflecting how important that word is to this topic. The coherence performance of the overall LDA greatly depends on the amount of topics that was chosen. Therefore it is important find the optimal number of topics. This is done by running the LDA for each number of topic up until 40. The results thereof show which number of topics gives the best coherence score, and thus is the optimal amount. Then the final LDA is executed with the topic number derived from the graph.

RQ3 (*How do the topics concerning age(ism) in software development discussed on Reddit compare to the topics concerning age(ism) in software development found in literature?*) was answered by comparing the topics found while answering RQ1 and RQ2 to the topics that are discussed in the Related Work section of this study.

Results

Optimal Number of Topics

The initial LDA ran to find the optimal number of topics resulted in the graph below. The graph shows that the LDA model with 20 topics has the highest coherence score.

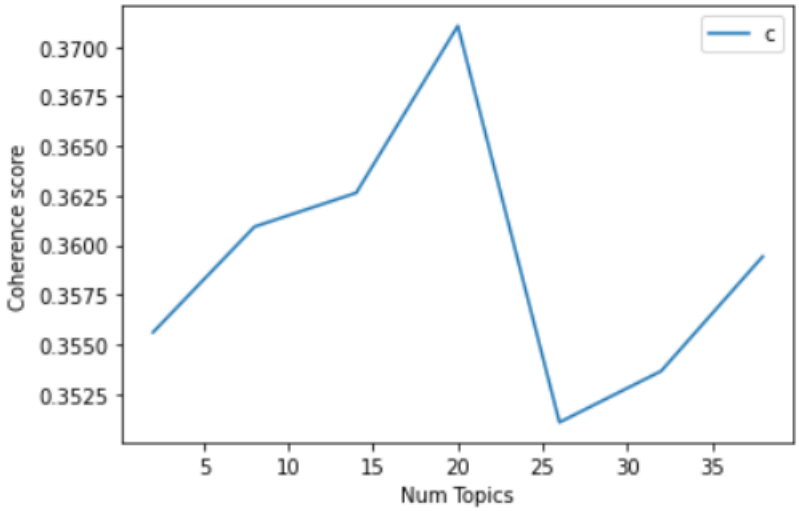

Figure 4.1: Optimal number of LDA topics

Latent Dirichlet Allocation

The results of the LDA with 20 topics are listed below. A more detailed list with the weights of each words can be found in Appendix C.

Topic 0	interview engin question comput experi career devop scienc salari epic
Topic 1	program engin degre tech salari job collabor intern without upcom
Topic 2	engin career question interview look learn devop scienc start creat
Topic 3	engin advic start program becom practic studi busi best recommend
Topic 4	comput science experi program engin differ learn germani team role
Topic 5	interview job experi engin thread comput applic tech work week
Topic 6	career engin devop question advic start posit agil manag use
Topic 7	engin comput salari language degre advic scienc differ process master
Topic 8	best devop engin major compani choos switch cours code consid
Topic 9	engin compani career look experi code applic use role salari
Topic 10	compani project base product engin idea life best posit dev
Topic 11	agil program start compani advic salari first amazon use project
Topic 12	devop engin internship role compani comput program career science degre
Topic 13	engin career work book advice hire look program field process
Topic 14	move make agil tool project start manage interview role intern
Topic 15	engin computer scienc differ project futur agil degre career advic
Topic 16	experi engin becom career team degre job senior tehcnic busi
Topic 17	career experi hire program manag skill design master possibl look
Topic 18	scienc comput engineer degre advic enough start career manag field
Topic 19	project devop compani first irish team becom engin design test

Discussion

Research Questions

These results do not appear to present coherent, distinctive topics. The contributing words to each topic do not not show a common theme, neither do they show much age-related aspects that can be discussed in the field of software development.
Therefore,

> **RQ1 Which age-related topics concerning software development are discussed on Reddit?**

is not able to be answered through the followed methodology in combination with the chosen data.
Since

> **RQ2 Which topics concerning age(ism) in software development are most frequently discussed on Reddit?**

relies upon the answer giving to RQ1, it also cannot be answered through the followed methodology in combination with the chosen data.
As there are no topics from the LDA to compare to the topics from the literature,

> **RQ3 How do the topics concerning age(ism) in software development discussed on Reddit compare to the topics concerning age(ism) in software development found in literature?**

cannot be answered as well using this methodology and dataset.

Implementations

Answering these research questions could be used to inform employers and managers in the software development industry about the issues their older developers may encounter. They could also contribute to understanding these challenges and how to potentially counter them.

Limitations

There are several limitations of the used methodology in combination with the collected data. First of all, they cannot be used to answer the research questions central to this study. The topics uncovered by the LDA analysis do not show much practical coherence and distinctiveness. It may be possible that using Latent Dirichlet Allocation, or even topic modelling is not the right method.

Second, a selection of keywords and subreddits was used to filter the data. Possibly this can lead to a bias in the results of the analysis. For instance, subreddits that do not have software development as theme were excluded. On these subreddits, Reddit users maybe discussed different age-related aspects in software development than on the more software development orientated subreddits.

Third, removing frequently used words and adding them to the stop word list might involuntarily push the analysis into a certain direction. The reason for this is that while some words may not appear to be useful or relevant for the topic analysis on their own, they could, together with other words, create a theme that 'sketches' the topic. Fourth, the reproduce-ability of this study is limited, since the topics change with every run, even when the parameters stay the same.

Future Research

Future research can be done to see the difference in how different subreddits discuss age in software development. Doing so could lead to insight of how different parts of the general public perceive age in software development. This also enables researchers to search for difference in perception between different countries, as some subreddits are country-related. A more advanced analysis could be made to look into the semantics of how age in software development is discussed on Reddit.

Conclusions

Age(ism) in software development is widely discussed in submissions on Reddit. Analysing the content of these submission through the Latent Dirichlet Allocation topic modelling method does not show any coherent, distinctive topics about age in software development. As a result, the in this study proposed research questions cannot be answered by performing the chosen methodology on the scraped data.

Bibliography

[1] S. Baltes and S. Diehl. Towards a theory of software development expertise. *In Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 187–200, 2018. 3

[2] S. Baltes, G. Park, and A. Serebrenik. Is 40 the New 60? How Popular Media Portrays the Employability of Older Software Developers. *IEEE Software*, 37(6):26–31, 2020. 1, 4

[3] D.M. Blei, A.Y. Ng, and M.I. Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003. 6

[4] L. Brooke. Prolonging the careers of older information technology workers: continuity, exit or retirement transitions? *Ageing and Society*, 29(2):237–256, 2009. 1, 4

[5] T.D. Comeau and C.L. Kemp. Intersections of age and masculinities in the information technology industry. *Ageing and Society*, 27(2):215–232, 2007. 4

[6] H.P. van Dalen, K. Henkens, and J. Schippers. Productivity of Older Workers: Perceptions of Employers and Employees. *Population and Development Review*, 36(2), 2010. 3

[7] J.L. Davidson, U.A. Mannan, R. Naik, I. Dua, and Jensen C. Older adults and free/open source software: A diary study of first-time contributors. *OpenSym '14: Proceedings of The International Symposium on Open Collaboration*, pages 1–10, 2014. 3, 4

[8] J.L. Davidson, R. Naik, U.A. Mannan, A. Azarbakht, and Jensen C. On older adults in free/open source software: reflections of contributors and community leaders. *IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 93–100, 2014. 4

[9] C. Klein, P. Clutton, and V. Polito. Topic Modeling Reveals Distinct Interests within an Online Conspiracy Forum. *Frontiers in Psychology*, 9:189, 2018. 6

[10] W. Kopec, B. Balcerzak, R. Nielek, G. Kowalik, A. Wierzbicki, and F. Casati. Older adults and hackathons: a qualitative study. *Empirical Software Engineering*, 23(4):1895–1930, 2018. 4

[11] R. Krampe and N. Charness. Aging and Expertise. *Cambridge University Press*, (2):835–856, 2018. 1, 3

[12] Bloomberg Law. https://news.bloomberglaw.com/daily-labor-report/hewlett-packard-hit-with-new-class-age-suit-over-retrenching?context=article-related. 1

[13] P. Morrison, R. Pandita, E. Murphy-Hill, and A. McLaughlin. Veteran developers' contributions and motivations: An open source perspective. *IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 171–179, 2016. 4

[14] Y. Murakami, M. Tsunoda, and H. Uwano. WAP: Does Reviewer Age Affect Code Review Performance? *IEEE 28th International symposium on software reliability engineering (ISSRE)*, pages 164–169, 2017. 3

[15] United Nations. https://www.un.org/en/global-issues/ageing. 1

[16] nltk. https://www.nltk.org/. 5

[17] U.S. Bureau of Labor Statistics. https://www.bls.gov/cps/cpsaat11b.htm. 1

[18] R.A. Posthuma and M.A Campion. Age stereotypes in the workplace: Common stereotypes, moderators, and future research directions. *Journal of Management*, 35(1):158–188, 2009. 3

[19] PRAW. https://praw.readthedocs.io/en/stable/. 5

[20] ProPublica. https://features.propublica.org/ibm/ibm-age-discrimination-american-workers. 1

[21] reddit. https://www.reddit.com/. 1

[22] G. Rodríguez-Pérez, R. Reza, and M. Nagappan. Perceived diversity in software engineering: a systematic literature review. *Empirical Software Engineering*, 26(2):102, 2021. 3

[23] U. Schloegel, A. Stegmann, and R. van Dick. Age stereotypes in agile software development: an empirical study of performance expectations. *Information Technology People*, 31(1):41–62, 2018. 3

Appendix A

List of subreddits that were used in the analysis:

- wgu_devs
- development
- gamedev
- agile
- askcomputerscience
- developersindia
- devops
- mobiledevelopment
- softwaredevelopment
- cscareerquestions
- webdev
- softwareengineering
- learnprogramming
- develeire
- cscareerquestionseu
- askprogramming
- computerscience
- experienceddevs
- epicsystems
- csmajors

These subreddits appeared as result of the following queries:
"software development"
(these appeared as results in both queries)
"software developer"

Appendix B

Additional stop words:

- would
- could
- should
- also
- even
- lot
- get
- time
- like
- feel
- good
- want
- wanting
- nan
- really
- anyone
- please
- find
- know
- knowing
- way
- one
- need
- go
- many
- much
- https
- www
- com
- data
- cases
- week
- day
- vs
- last
- web
- ca
- etc
- cs
- take
- year
- years
- game
- level
- png
- new
- back
- got
- well
- since
- going
- software
- developer
- development
- developing
- develop
- developers
- people
- think
- help
- getting

Appendix C

The results of the 20 topic LDA with weighted words:

Topic 0
0.030*"interview" + 0.019*"engin" + 0.017*"question" + 0.014*"comput" + 0.013*"experi" + 0.013*"career" + 0.012*"devop" + 0.012*"scienc" + 0.008*"salari" + 0.008*"epic"

Topic 1
0.033*"program" + 0.018*"engin" + 0.013*"degre" + 0.013*"tech" + 0.008*"salari" + 0.008*"job" + 0.007*"collabor" + 0.007*"intern" + 0.007*"without" + 0.007*"upcom"'

Topic 2
0.038*"engin" + 0.027*"career" + 0.021*"question" + 0.020*"interview" + 0.014*"look" + 0.013*"learn" + 0.011*"devop" + 0.010*"scienc" + 0.010*"start" + 0.007*"creat"

Topic 3
0.026*"engin" + 0.025*"advic" + 0.023*"start" + 0.014*"program" + 0.012*"becom" + 0.007*"practic" + 0.007*"studi" + 0.007*"busi" + 0.007*"best" + 0.007*"recommend"

Topic 4
0.019*"comput" + 0.016*"scienc" + 0.014*"experi" + 0.014*"program" + 0.011*"engin" + 0.010*"differ" + 0.010*"learn" + 0.010*"germani" + 0.010*"team" + 0.010*"role"

Topic 5
0.017*"interview" + 0.016*"job" + 0.014*"experi" + 0.013*"engin" + 0.011*"thread" + 0.010*"comput" + 0.010*"applic" + 0.010*"tech" + 0.009*"work" + 0.008*"week"

Topic 6
0.036*"career" + 0.018*"engin" + 0.013*"devop" + 0.013*"question" + 0.011*"advic" + 0.011*"start" + 0.008*"posit" + 0.008*"agil" + 0.008*"manag" + 0.008*"use"

Topic 7
0.082*"engin" + 0.020*"comput" + 0.017*"salari" + 0.016*"languag" + 0.015*"degre" + 0.014*"advic" + 0.011*"scienc" + 0.011*"differ" + 0.009*"process" + 0.008*"master"

Topic 8
0.036*"best" + 0.023*"devop" + 0.014*"engin" + 0.014*"major" + 0.013*"compani" + 0.013*"choos" + 0.012*"switch" + 0.010*"cours" + 0.010*"code" + 0.009*"consid"

Topic 9
0.020*"engin" + 0.018*"compani" + 0.017*"career" + 0.016*"look" + 0.012*"experi" + 0.010*"code" + 0.010*"applic" + 0.009*"use" + 0.009*"role" + 0.009*"salari"

Topic 10
0.028*"compani" + 0.016*"project" + 0.014*"base" + 0.013*"product" + 0.013*"engin" + 0.011*"idea" + 0.009*"life" + 0.008*"best" + 0.008*"posit" + 0.008*"dev"

Topic 11
0.023*"agil" + 0.021*"program" + 0.017*"start" + 0.017*"compani" + 0.012*"advic" + 0.011*"salari" + 0.010*"first" + 0.009*"amazon" + 0.009*"use" + 0.009*"project"

Topic 12
0.028*"devop" + 0.020*"engin" + 0.017*"internship" + 0.016*"role" + 0.015*"compani" + 0.015*"comput" + 0.014*"program" + 0.013*"career" + 0.012*"scienc" + 0.011*"degre"

Topic 13
0.034*"engin" + 0.022*"career" + 0.018*"work" + 0.015*"book" + 0.015*"advic" + 0.014*"hire" + 0.014*"look" + 0.012*"program" + 0.009*"field" + 0.009*"process"

Topic 14
0.012*"move" + 0.012*"make" + 0.012*"agil" + 0.012*"tool" + 0.010*"project" + 0.010*"start" + 0.009*"manag" + 0.009*"interview" + 0.009*"role" + 0.009*"intern"

Topic 15
0.068*"engin" + 0.027*"comput" + 0.026*"scienc" + 0.014*"differ" + 0.011*"project" + 0.011*"futur" + 0.011*"agil" + 0.011*"degre" + 0.009*"career" + 0.009*"advic"

Topic 16

0.032*"experi" + 0.021*"engin" + 0.014*"becom" + 0.012*"career" + 0.011*"team" + 0.011*"degre" + 0.009*"job" + 0.009*"senior" + 0.009*"technic" + 0.009*"busi"

Topic 17

0.016*"career" + 0.013*"experi" + 0.012*"hire" + 0.010*"program" + 0.010*"manag" + 0.010*"skill" + 0.010*"design" + 0.009*"master" + 0.008*"possibl" + 0.008*"look"

Topic 18

0.021*"scienc" + 0.017*"comput" + 0.015*"engin" + 0.011*"degre" + 0.011*"advic" + 0.009*"enough" + 0.009*"start" + 0.009*"career" + 0.008*"manag" + 0.008*"field"

Topic 19

0.031*"project" + 0.014*"devop" + 0.011*"compani" + 0.011*"first" + 0.011*"irish" + 0.011*"team" + 0.010*"becom" + 0.010*"engin" + 0.010*"design" + 0.009*"test"