

# BioNT - Network for Training

Deliverable 2.1 | Training needs analysis report



**Co-funded by  
the European Union**

Version 1.0

31 October 2023

Grant Agreement number	101100604 - DIGITAL-2022-TRAINING-02
Action Acronym	BioNT
Action Title	Bio Network for Training
Deliverable number and title	2.1 Training needs analysis report
Work package number and title	2 Training needs in industry
Dissemination level	Public
Authors	R. Alves, S. Di Giorgio, I. Letunic, E. Ortega, L. Paladin, I. Paredes Cisneros, approved by all BioNT partners
Delivery date	2023-10-31

Project partners	1. EUROPEAN MOLECULAR BIOLOGY LABORATORY (EMBL)
	2. BIOBYTE SOLUTIONS GMBH (BIOBYTE)
	3. HPC NOW CONSULTING SL (HPCNOW)
	4. UNIVERSITETET I OSLO (UiO)
	5. UNIVERSITAT DE BARCELONA (UB)
	6. INFORMATION CENTRE FOR LIFE SCIENCE (ZBMED)
	7. SIMULA CONSULTING (SIMULA)
	8. ALBERT-LUDWIGS-UNIVERSITAET FREIBURG (ALU-FR)
	9. ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE (EPFL)

## Table of Contents

<b>Project overview</b>	<b>2</b>
<b>Training needs survey: Overview and introduction</b>	<b>3</b>
Design and content	3
Advertisement and dissemination	3
Results	5
Audience	5
Comparison between Academia and Industry	7
Training	7
Impact on BioNT's training offers	10
Workshop schedule	10
Workshop format	11
Training offers	11
<b>Next steps</b>	<b>12</b>
<b>Conclusion</b>	<b>12</b>
<b>Annex 1: Survey questionnaire</b>	<b>13</b>

## Project overview

The BioNT consortium is dedicated to providing a comprehensive training program and fostering a community for digital skills relevant to the biotechnology industry and biomedical sector. With a curriculum tailored for beginners and advanced professionals, BioNT aims to equip individuals with the necessary expertise in handling, processing, and visualising biological data, as well as utilising computational biology tools and computing resources at medium-large scale. Leveraging the consortium's strong background in digital literacy

training and extensive network of collaborations, BioNT is poised to professionalise life sciences data management, processing, and analysis skills.

## Training needs survey: Overview and introduction

The foundational pillar of the BioNT consortium project is Work Package number 1: Training design and development (WP1). As previously highlighted in the project overview, this component is fundamental given the consortium's primary goal of creating specialised training materials of superior quality tailored to the field. These training materials, while still free and openly accessible to the academic community, target trainees who are looking for a job in the biotech/biomed sector, as well as employees of the diverse ecosystem of small and medium enterprises (SMEs). Therefore, to ensure the successful realisation of this objective, the training design must be supported by a thorough analysis of the training needs of its target audience.

To address this specific BioNT task, the consortium collaborated with other entities sharing similar goals to create a survey aimed at gathering information about specific training requirements and preferences. This allowed for expanding the reach of BioNT's initiative, and nurturing and structuring future collaboration with these entities. In collaboration with [ELIXIR](#), and its German node, [de.NBI](#), BioNT designed, conducted, disseminated and analysed a training needs survey. The survey aimed to assess the interests of the training audience regarding the content, format and modality of the training.

### Design and content

The survey, titled "Bioinformatics training needs", was divided into two sections: "Audience" and "Training". The complete survey questionnaire is provided in [Annex 1](#), and the survey data was hosted and managed by EMBL. The survey included a total of 11 questions, 6 mandatory and 5 optional, collected anonymously.

The first part of the survey aimed to identify the participants by collecting their current employment situation, demographics and areas of expertise. This information allowed aggregating the collected data according to the audience. The second part of the survey assessed the participant's training preferences, including the topics of interest, their desired interaction with other participants and the time of the week/day during which they would prefer to be trained.

The survey was carefully designed with the contribution of the Industry Programme experts of ELIXIR and de.NBI, with the aim to find a balance between obtaining all the relevant information for BioNT and its collaborators, and minimising the effort required for survey participants to provide their input and submit their responses.

### Advertisement and dissemination

The survey was initially advertised on July 13th of 2023 through BioNT's social media accounts on [LinkedIn](#) and [X](#), posting approximately every two weeks. According to statistics

from the two platforms, the initial post reached 501 and 586 impressions, respectively. Examples of the social media posts are shown in Figure 1.

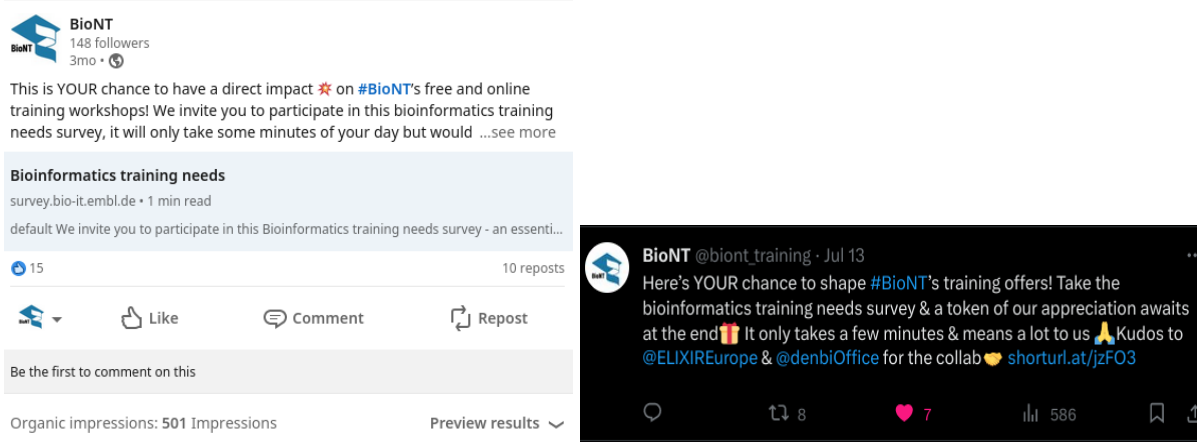


Figure 1: Examples of the posts in BioNT's social media channels, LinkedIn (left) and X (right) to advertise the training needs survey.

ELIXIR and de.NBI were tagged in these posts to facilitate and accelerate their dissemination. In addition, considering the survey's target audience, this was also distributed through the industry newsletters of the ELIXIR Hub and de.NBI. An example of ELIXIR's post on [LinkedIn](#) advertising the survey is shown in Figure 2.

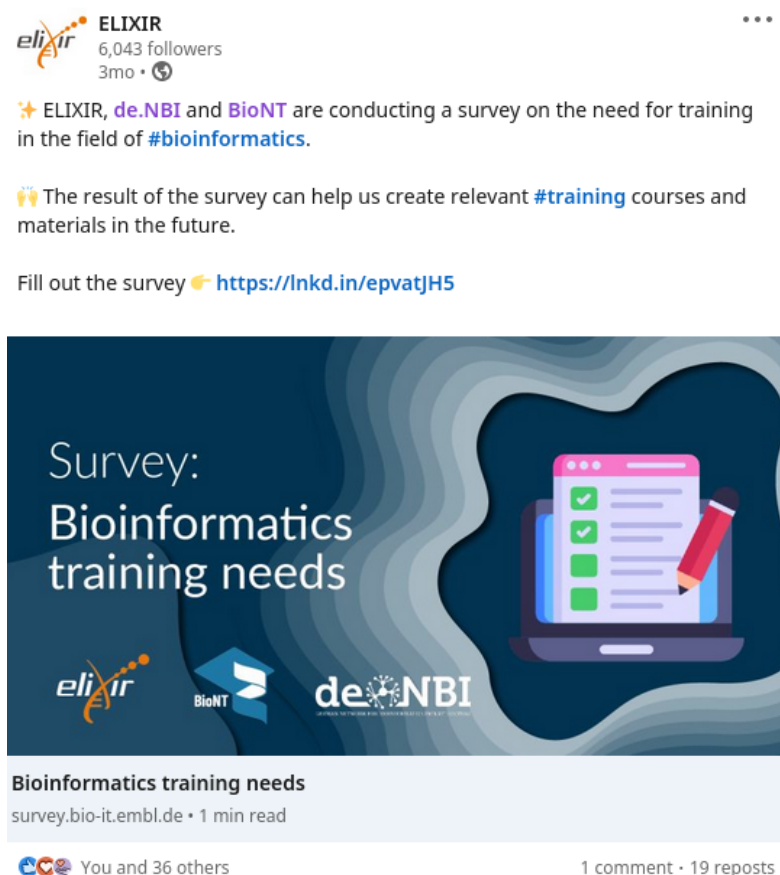


Figure 2: Example of post through ELIXIR's social channel in LinkedIn, advertising the survey.

## Results

The survey was successfully completed and submitted by 68 respondents. As not all questions were mandatory, the number of responses varied among the mandatory and optional questions. For the complete set of questions and answers included in the survey refer to [Annex 1](#).

### Audience

Job seekers and industry/SME-employees accounted for approximately 41.3% of the total survey participants: job seekers from various sectors (26.5%) and industry employees (SME and big enterprises, 14.7%). Most of the survey participants identified themselves as being in the academic sector, where the majority were constituted of academic employees (44.1%). The participant's responses for question Q1, with 68 respondents, are shown in Figure 3.

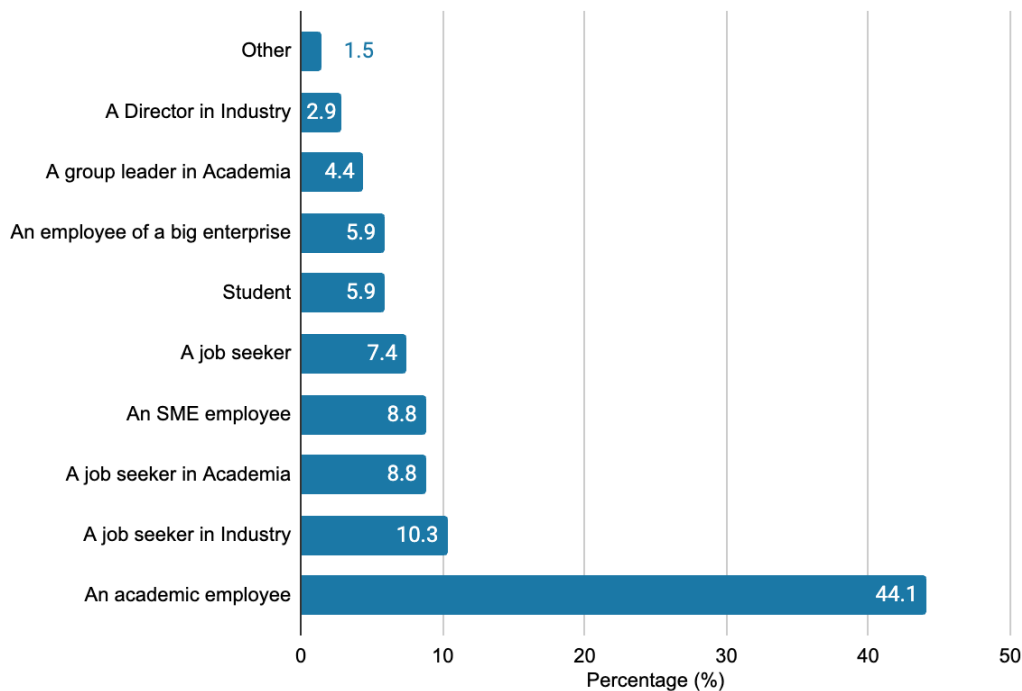


Figure 3: Answers to the question Q1 “Would you define yourself as?”. The category “Others” groups the answers “A group leader or Director in Academia” and “A director or entrepreneur in Industry”. The category “Student” was added separately as it was repeatedly found in the free-text field “Other”. This question was mandatory (N = 68).

Question Q2 covered 59 respondents and was relevant to better understand the target audience, based on their career stage. The majority of the respondents indicated to be employed as research staff (21.0%) and graduate students (19.7%), followed by analysts (13.1%) and postdoctoral researchers (10.5%), as shown in Figure 4.

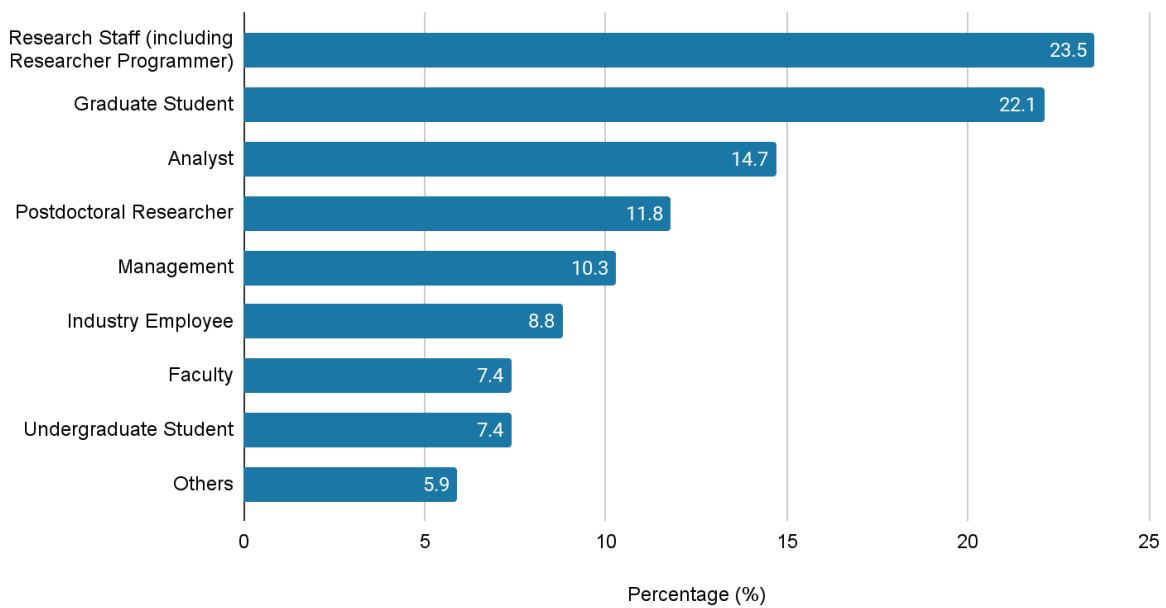


Figure 4: Answers to the question Q2 “If you are employed, what is your current occupation/career stage?”. The category “Others” groups the answers “administrative staff”, “commercial software developer”, “government employee”, “librarian or archivist”, “retired or not employed”, and “support staff (including technical support)”. This question was optional (N = 59).

Question Q3 covered 66 respondents, the majority (61.7%) of which were working or looking for a job within the European Union (EU), either in their country of origin (39.7% of EU nationals) or in a different country (13.2% of non-EU nationals and 8.8% of EU nationals). 30.9% of the respondents were non-EU nationals working or looking for a job outside the EU (22.1% in their country of origin and 8.8% in another non-EU country). A small portion of the respondents (7.4%) were EU nationals working outside of the EU.

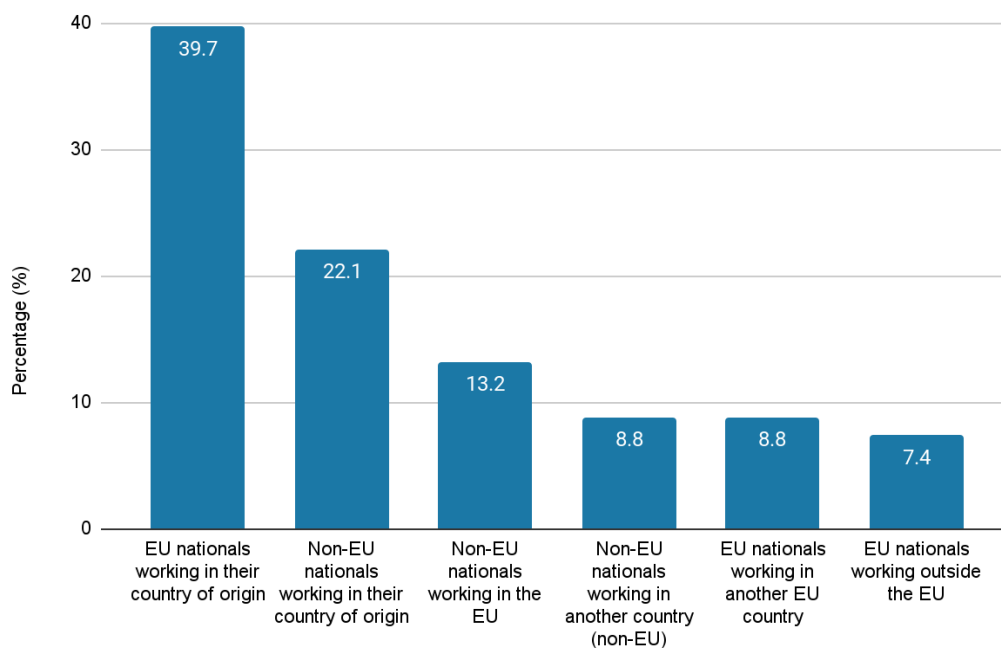


Figure 5: Combined answers of the question Q3 “What is your country of employment or the country you are looking for a job in?” and Q4 “What is your nationality?”. The answers have been grouped according to the categories shown. The term “working” in the labels covers employees as well as job seekers. The question Q3 was optional, whereas Q4 was mandatory (N = 66).

## Comparison between Academia and Industry

From this point onwards, the survey answers have been categorised into two groups, “Academia” and “Industry”, defined as follows:

Based on question Q1, respondents who selected an answer containing the word “industry” or “enterprise” were grouped into the “Industry” category (N = 19). Similarly, respondents were grouped into the “Academia” category (N = 44) if their answer included the terms “academia”, “academic” or “student”. Individuals identifying solely as “job seekers” were excluded from the results presented in the rest of the report (N = 5). Percentages in the subsequent bar charts were calculated based on the total number of responses in each category.

In the last question of the Audience section (Q5), participants were asked about the area(s) of expertise within the Life Sciences relevant to them. The answers, displayed in Figure 6, revealed that the most common areas of expertise were Genomics (72.7% in Academia and 68.4% in Industry) and Sequencing (59.1% in Academia and 73.7% in Industry).

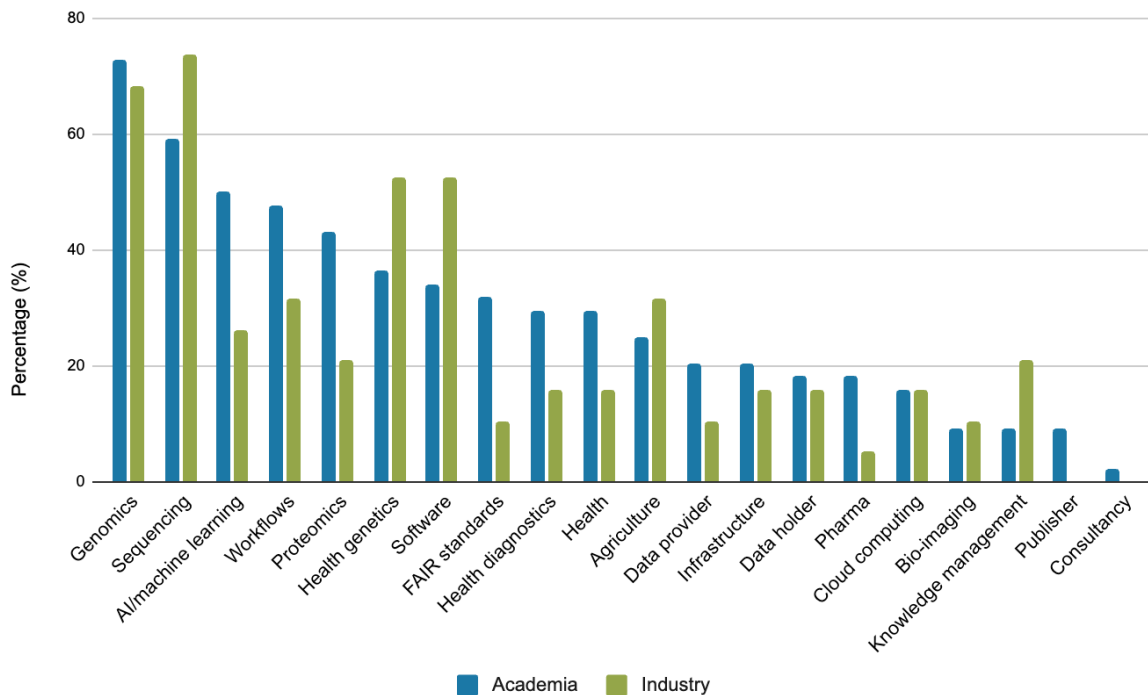


Figure 6: Answers to the question Q5 “Within the Life Sciences, is any of these areas of expertise relevant to you?”. The categories “Publisher” and “Consultancy” had no responses within the “Industry” category. This question was optional. (N-Industry = 18, N-Academia = 44).

## Training

The actual assessment of the training needs started with question Q6, focusing on evaluating topics relevant to the respondents. Moving forward, a comparative analysis of responses is presented, with the data aggregated as mentioned in the [previous section](#). Question Q6 covered 68 respondents, where the majority expressed interest in the “Data analysis” field (94.7% in Industry and 84.1% in Academia), followed by “Software usage” and “Data management” (52.6% in Industry and 54.5% in Academia, for both fields). Less

appealing fields within the Industry category included “Develop data-based solutions”, “Text mining” and “Develop software solutions”.

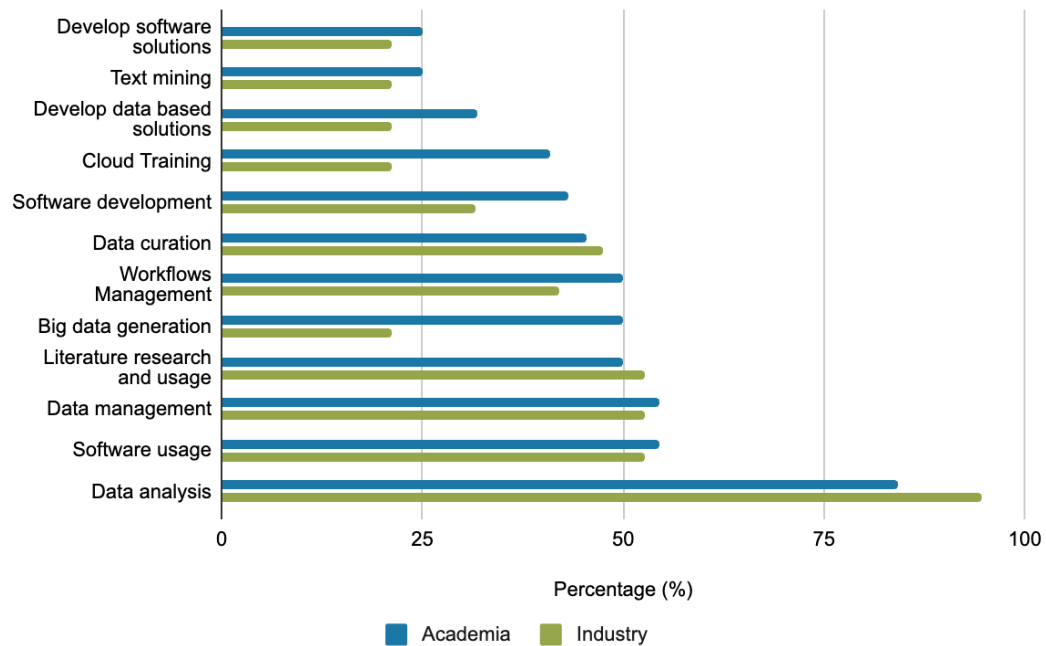


Figure 7: Answers to the question Q6 “Within Bioinformatics / Computational Biology, is any of these topics relevant to you?”. The category “Others” had no entry. This question was mandatory (N-industry = 19, N-academia = 44).

The question Q7 enquired about which of BioNT’s training offers were most relevant to the target audience. This question covered 67 respondents, and among the 8 different training workshops provided by BioNT, nearly all of them were considered relevant for the Academia and Industry categories. The only exceptions were “Open and Fair principles and data management”, “Instructor Training” and “System administration”, where the interest from the both categories was low.

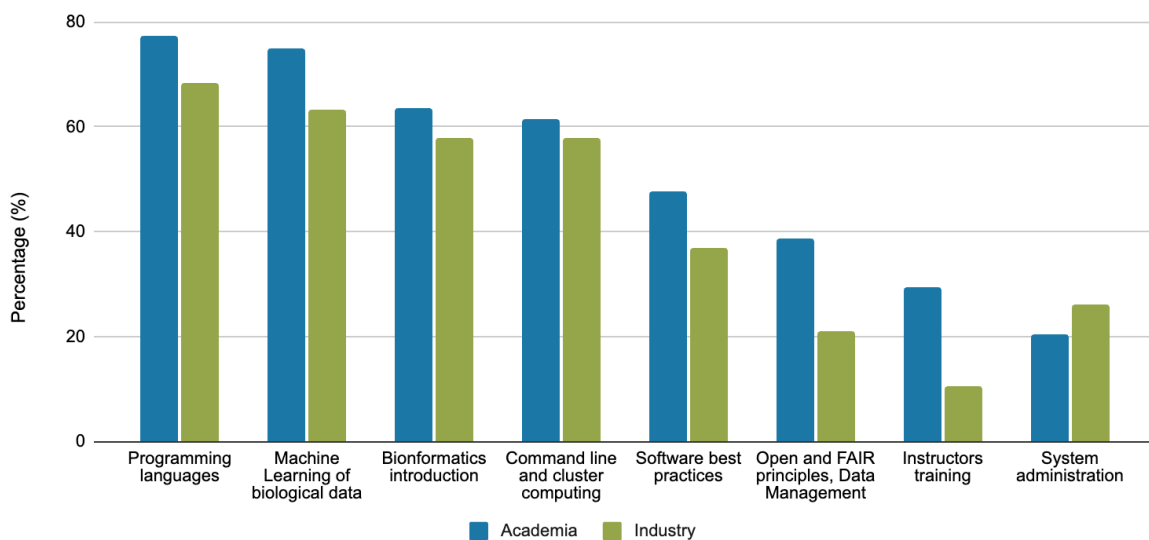


Figure 8: Answers to the question Q7: “These macro-area(s) of training will be specifically addressed in the BioNT training programme. Which one(s) would mostly interest you?”. This question was optional. (N-academia = 44, N-industry = 18).



The preferred training format was explored with question Q8, which covered 68 respondents, the majority expressing a preference for “Online workshop with few attendees, very interactive, closed assistance” (78.9% in Industry and 86.4% in Academia). Additionally, a significant portion indicated a desire for “E-learning materials to consult anytime” (63.2% in Industry and 84.1% in Academia).

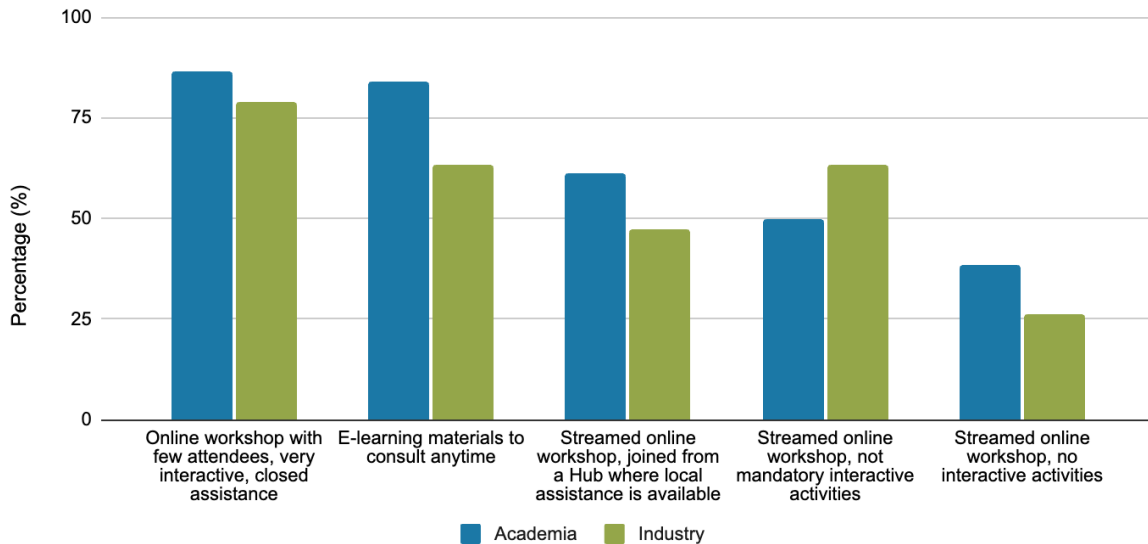


Figure 9: Answers to the question Q8: “Would you attend a workshop in the following formats?”. This question was mandatory (N-industry = 19, N-academia = 44).

Question Q9 delved into participants' schedule preferences, considering their needs. This question was answered by 68 respondents, showing considerable variability. Among the Industry category, 52.6% expressed a preference for "Half days of training in the afternoon," while 57.9% indicated interest in "Full days of training (few in total)". On the other hand, most of the respondents in the Academia category preferred "Half days of training in the evenings" (61.4%).

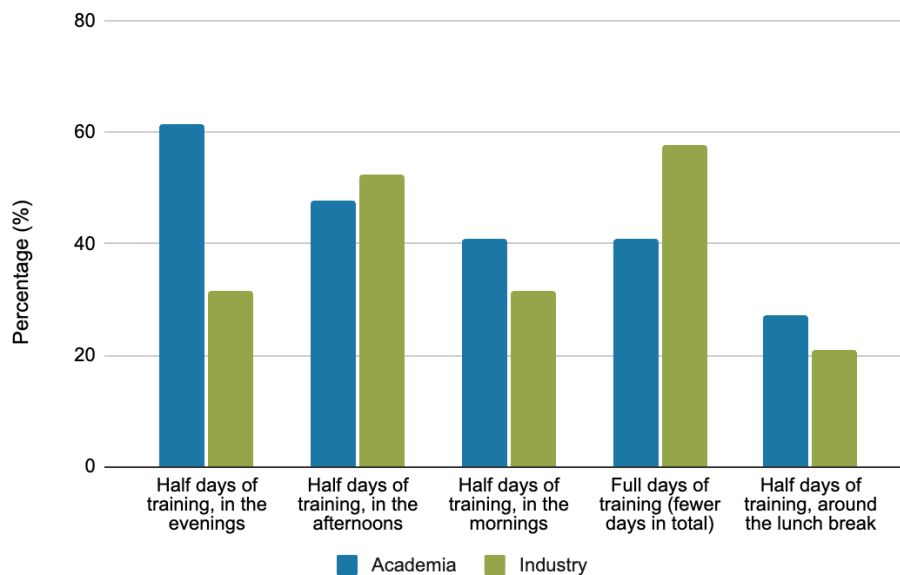


Figure 10: Answers to the question Q9: “Which time of the day would you be able to join:?”. This question was mandatory (N-industry = 19, N-academia = 44).

Question Q10 aimed to identify the quality of engagement the participants preferred during the workshop. This question was answered by 68 respondents and showed that most of them favour the interaction with other workshop participants (90.9% in Academia and 84.2 in Industry).

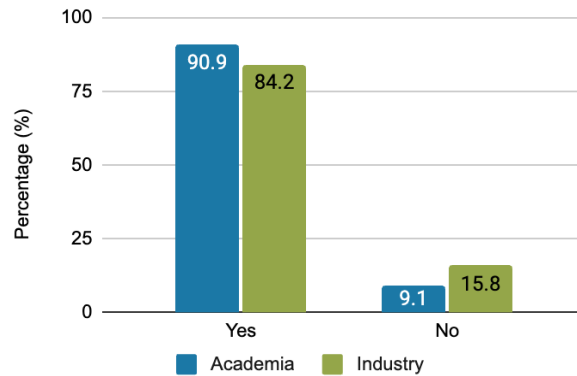


Figure 11: Answers to the question Q10: “Do you think it would be an added value to favour interaction among the workshop participants?”. This question was mandatory (N-industry = 19, N-academia = 44).

## Impact on BioNT’s training offers

Based on the results from the survey, the consortium aims to identify how the training workshops can be made more appealing, tailoring the delivery format and content to the target audience.

### Workshop schedule

The first training workshop titled “[A practical introduction to bioinformatics and RNA-seq using Galaxy](#)” was successfully conducted from September 4th to 8th in the mornings (9:00 - 13:00 CEST). After a preliminary analysis of the survey results, specifically focusing on question Q9 (Figure 10) from the training needs survey, the BioNT consortium decided to test delivering the training workshops in the evening. As a result, the second workshop, “[From Hero to Zero with Python](#),” is scheduled for November 21st to 24th from 18:00 to 21:00 CET. This change in timing will allow for an assessment of audience engagement during these evening sessions.

According to the responses gathered in question Q9, the participants in the Academia category expressed a preference for evening training sessions, organised as half-day sessions. In contrast, participants from the Industry category preferred full-day training sessions, condensed into fewer total days. To cater to the specific needs of SMEs and industry professionals, the BioNT consortium decided to continue testing with delivering the training workshops at various times of the day, guided by the presented survey results. The remaining training workshops in the basic curriculum would be delivered with different schedules to further assess how the training schedule impacts the participation of the different audiences interested in BioNT’s training offers.

## Workshop format

The first BioNT workshop was delivered in a Zoom Webinar format, with participants' visibility disabled to preserve their privacy. With the aim to protect the participant's anonymity, they were able to watch and learn from the trainers with no direct interaction (e.g. chat or voice) possible during the sessions. Instead, a collaborative document, previously set up by the trainers, was shared with all the participants before the session. Trainees were able to interact and engage anonymously with other trainees, as well as with the workshop helpers and trainers, directly in the collaborative document. Using this format, the trainees could remain anonymous, if desired, which facilitated the interactions. This decision was taken to accommodate the needs of all participants, specifically the ones who expressed their concern about directly interacting with other participants. In fact, according to the survey, specifically to question Q10, 9.1% of respondents from Academia and 15.8% from the industry sector expressed a negative preference regarding a possible interaction with other participants. The overall survey results, however, show that generally trainees are interested in having direct interactions with the trainers, as well as with other trainees. Consequently, the consortium will assess the possibility of offering more interactive workshops for the delivery of BioNT's advanced curriculum.

Following BioNT's project proposal, the training of the basic curriculum will be delivered twice. This provides an opportunity to explore a completely different delivery format. The first round of training takes place using the Zoom Webinar format with minimal trainer-trainee interaction, as specified above, where the sessions are video-recorded. These records could be used to create comprehensive electronic materials for self-paced learning, which has generated significant interest among the survey participants (Q8, Figure 9).

## Training offers

The survey participants were questioned about their interest in specific training offerings from BioNT (question Q7). The most attractive options for both categories, Academia and Industry, were "Programming languages" and "Machine Learning of biological data". However, certain offers, such as "Instructors training" and "System administration", gathered very little interest. Based on this data, the consortium will discuss the possible course of action. An important clarification is that the information provided in question Q7 about the BioNT workshops was limited to the titles as presented in the proposal. In practice, once the workshop is developed and advertised, a more precise and engaging title, accompanied by further information about the training content, is provided. Therefore, the data might indicate that a precise and appealing title should be given in each case, and the workshop content itself should consider the most relevant and highly demanded areas within the Life Sciences. This was also validated on question Q6 (Figure 7), where the relevance of topics within the Life Sciences was assessed. In this particular question, the "data management" workshop was among the most popular answers for the Academia and Industry categories. However, when participants were asked about BioNT-specific offers in question Q7 (Figure 8), the workshop titled "Open and FAIR principles, Data Management" did not rank among the most popular, particularly within the Industry category. This suggests that the term "Open and FAIR principles" may not be interesting or well-known to the audience and that the provisional title was not very appealing. This situation highlights the need for further clarification or adaptation of the workshop content.

## Next steps

The training needs in digital Life Sciences assessed in the survey will continue to be relevant for BioNT, as well as for ELIXIR and de.NBI. Through the continuous collaboration with them, and with the survey analysis, different improvements to the survey were identified. Once implemented, an improved version is planned to be set up and disseminated, to provide further information about the training needs of the different audiences targeted by BioNT.

One of the aspects to be improved is related to the clarification of the different terms used in the survey. A glossary will be implemented to define terms used in various places throughout the questionnaire such as “FAIR”, “SME” and “job seeker”. Specific clarifications will also be provided in certain questions when inquiring about participants' current career stage. For instance, respondents will be informed that if they are in the final stages of their PhD and contemplating their future career paths, they should preferably be classified as “job seekers” rather than “students”. Additionally, further clarification will be given to PhD students emphasising that they would preferably identify themselves as either academic or industry employees and not “students”.

Where possible, the number of answers would be minimised, providing categories that encompass the relevant answer groups, for example for questions Q2 and Q5. In contrast, question Q1 will be split into two different questions: The first one enquires about the respondent's current situation, and the second about the next steps they would like to take. In this context, a clearer set of categories would be defined.

## Conclusion

The training needs survey conducted by the BioNT consortium provided valuable insights into the preferences of the target audience. The survey successfully gathered responses from 68 participants, including job seekers, and people in the academic and industry sectors. The survey highlighted the most common areas of expertise as Genomics and Sequencing, with participants expressing interest in topics such as Data Analysis, Software Usage, and Data Management. The preferred training format was defined as an interactive online workshop, as well as e-learning materials for self-paced learning. These findings from the survey influenced changes in the workshop schedule and format, with the consortium testing evening sessions and implementing a collaborative document approach. The survey also prompted discussions on refining workshop titles and content based on participant interest. Moving forward, the consortium plans to continue the collaboration with ELIXIR and de.NBI to further address training needs and make improvements to the survey design. Overall, the results have guided the consortium in tailoring the training program to meet the needs of individuals in the biotechnology and biomedical sectors.

## Annex 1: Survey questionnaire

Question number and text	Type	Answers
<b>Section 1: Audience</b>		
Q1 - Would you define yourself as*:	Single choice with open text alternative	An academic employee An employee of a big Enterprise An employee of a Small or Medium Enterprise (SME) A group leader or Director in Academia A Director or entrepreneur in Industry A job seeker in Academia A job seeker in Industry A job seeker
Q2 - If you are employed, what is your current occupation/career stage?	Multiple answers with free text alternative	Administrative Staff Analyst Commercial Software Developer Faculty Government Employee Graduate Student Industry Employee Librarian or Archivist Management Postdoctoral Researcher Research Staff (including Researcher Programmer) Retired or Not Employed Support Staff (including Technical Support) Undergraduate Student
Q3 - What is your country of employment or the country you are looking for a job in?	Single choice	List of countries
Q4 - What is your nationality?*	Single choice	List of countries

<p>Q5 - Within the Life Sciences, is any of these areas of expertise relevant to you:</p>	<p>Multiple answers with free text alternative</p>	<p>FAIR standards          Software          Workflows          Health          Health genetics          Sequencing          Data provider          Data holder          Publisher          AI/machine learning          Knowledge management          Bioimaging          Health diagnostics          Infrastructure          Pharma          Agriculture          Cloud computing          Consultancy          Proteomics          Genomics</p>
<b>Section 2: Training</b>		
<p>Q6 - Within Bioinformatics / Computational Biology, is any of these topics relevant to you*:</p>	<p>Multiple answers with free text alternative</p>	<p>Data analysis          Data curation          Data management          Literature research and usage          Text mining          Software usage          Software development          Big data generation          Develop software solutions          Develop data based solutions          Workflows Management          Basic or advanced Cloud Training (virtual machines, container orchestration etc.)</p>
<p>Q7 - These macro-area(s) of training will be specifically addressed in the BioNT training programme. Which one(s) would mostly interest you?</p>	<p>Multiple answers</p>	<p>Bioinformatics introduction          Programming languages          Command line and cluster computing          Open and FAIR principles, Data Management          Instructors training          Software best practices          Machine Learning of biological data          System administration</p>

<p>Q8 - Would you attend a workshop in the following formats*:</p>	<p>Multiple answers</p>	<p>Online workshop with few attendees, very interactive, closed assistance          Streamed online workshop, joined from a Hub where local assistance is available          Streamed online workshop, not mandatory interactive activities          Streamed online workshop, no interactive activities          E-learning materials to consult anytime</p>
<p>Q9 - Which time of the day would you be able to join*:</p>	<p>Multiple answers</p>	<p>Full days of training (fewer days in total)          Half days of training, in the mornings          Half days of training, in the afternoons          Half days of training, around the lunch break          Half days of training, in the evenings</p>
<p>Q10 - Do you think it would be an added value to favour interaction among the workshop participants?*</p>	<p>Single choice with open text alternative</p>	<p>Yes          No, I have concerns regarding the interaction with other participants</p>
<p>Q11 - Please consider the Bioinformatics needs of your (potential) working environment (e.g. access to training, need to hire bioinformaticians, usage of cloud services, alignment with standards, consultancy on building ad-hoc data and software solutions, etc.) and the specific training-related requirements (e.g. video-conferencing platform, certification requirements, tools licences, time commitment, etc.). Is there anything else you would like us to know about your needs, in terms of training content, format, duration, technical requirements?</p>	<p>Free text answer</p>	
<p>*The questions marked with a star (*) were mandatory</p>		