European Genomic Data Infrastructure

Deliverable D2.2

Infrastructure Cost report

Project Title Grant agreement no	**Genomic Data Infrastructure** Grant agreement 101081813		
Project Acronym (EC Call)	GDI		
WP No & Title	WP2: Long term sustainability		
WP Leaders	Regina Becker (51. PNED) Troels Rasmussen 13. NGC)		
Deliverable Lead Beneficiary	13. Danish National Genome Center		
Contractual delivery date	31/10/2023	**Actual delivery date**	31/05/2024
Delayed	Yes		
Partner(s) contributing to deliverable	NGC		
Authors	Troels Rasmussen, NGC, DK		
Contributors	Regina Becker, PNED, LU		
Acknowledgements	N/A		
Reviewers	Rob Hooft, Health-RI, NL Tommi Nyrönen, CSC, FI		

Log of changes

Date	Mvm	Who	Description
13/04/2024	V01	Troels Rasmussen, NGC	First draft
30/04/2024	V02	Troels Rasmussen, NGC	Uploaded for review
21/05/2024	V03	Mercedes Rothschild Steiner (ELIXIR Hub)	Deliverable sent to GDI MB for review

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

29/05/2024	V04	Troels Rasmussen, NGC	Addressed comments
06/06/2024	1V0	Mercedes Rothschild Steiner (ELIXIR Hub)	Final version submitted to EC portal
10/06/2024	2V0	Mercedes Rothschild Steiner (ELIXIR Hub)	Amended final version submitted to EC portal

Table of contents

1. Executive Summary 3
2. Contribution towards project outcomes 5
3. Methods 7
4. Results 9
    4.1. Central hub - Cost calculation and distribution 9
    4.2. Central hub – 10-year Budget 10
    4.3. Central hub – Membership distribution key 11
    4.4. National costs and obligations 17
        4.4.1. Discussion on cost parameters 17
        4.4.2. Estimation of staff level – nationally 22
5. Subjects for further discussion and analysis 24
6. Next steps 27
7. References 28

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

1. Executive Summary

In the context of Work package 2, this report follows the work in the D2.1 report *First analysis of cost elements for the setup of 1MG infrastructure*, which was focused on identifying the different cost items that go into building and running the 1+MG infrastructure, both on a national node level and at a central hub level.

In this deliverable, it is assumed that the infrastructure will be based on a European Digital Infrastructure Consortium[1] (EDIC). The identified cost items are compiled into a budget that includes implementation, operations and maintenance costs over a ten-year period – for the central activities. However, it should be noted from a ten year budget perspective, given the level of complexity in an area that changes all the time – the budget shall only be considered a baseline and indicative – particularly for years 6 through 10.

For the local activities, only budgeting guidelines are given, since the actual cost on local level is dependent on the amount of data managed, used and shared by the infrastructure and since the level of connection and synergy with existing national infrastructure components will vary a lot across different countries. The focus has been on assessing how many FTEs of personnel will be needed to get a basic service layer into production, and which known costs that will incur – particularly on the storage and risk management part.

This report will give the reader the first overview of the financial obligations for a country in becoming a member of the EDIC. In this report, the assumption is that the costs of the central operation will be distributed across member countries based on the individual countries GDP level. While this is only the first iteration of the budget and cost distribution, it should provide an indication of the expected size of the general membership fee.

Below are some key messages, derived from the report. The reader should consider that the findings and the report are the first iteration, and will likely be corrected in the course of this project.

Key messages:

Central costs

- The indicative central budget increases from 700.000 EUR in the first year of operation to 4.5 MEUR in the 10th year, with a relatively quick buildup phase in the first five years.
- A budget of 325.000 EUR has been allocated for the two years preceding the formation of the EDIC, in order to have dedicated staff in place to drive the progress towards implementation of the EDIC .

[1] https://digital-strategy.ec.europa.eu/en/policies/edic

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

- The Cost distribution key between countries is based on National Net Income + a flat rate fee of 25.000 EUR, as suggested by the EDIC Task Force.
- The indicative contribution per country is based on an assumed number of member states joining as founding members.
- Currently there are no financial provisions that would incentivise countries to join early - or vice versa financial provisions that would allow some of the development costs to be recovered from countries joining later.

National costs

- Storage: The cost for storing the data comes to 14-16 EUR per whole genome sequence.
- Indicatively if in a country 0,4 % of the population is sequenced every year, total storage costs could be 60.000 EUR per million citizen accumulated each year
- In order to operate a national node each country can expect to need to allocate 5-8 FTE's (Full-Time Equivalent) to provide minimal services to a limited number of users. The number of people needed will grow further when the infrastructure starts onboarding significant data and running many research projects, which will also require increased compute capacity - please consult the D2.1 report for further discussion on these subjects.
- To perform a Data Protection Impact Assessment (DPIA) and set up a local Risk management system that can be audited and/or certified according to international standards and subject to NIS2, each country will likely need to make a one-time investment of 100.000-500.000 EUR for external consultants, and allocate 12-24 PMs (Person-Months) internally to succeed. 1-2 FTE's and around 30.000 EUR should be allocated on a permanent basis to support risk management (these figures are subject to national cost level, the complexity of operations and existing structures)
- The EDIC depends on national structures for data subject information portal and consent registration.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

2. Contribution towards project outcomes

With this deliverable, the project has reached or the deliverable has contributed to the following project outcomes:

	Contributed
Outcome 1 Secure federated infrastructure and data governance needed to enable sustainable and secure cross border linkage of genomic data sets in compliance with the relevant and agreed legal, ethical, quality and interoperability requirements and standards based on the progress achieved by the 1+MG initiative.	Yes
Outcome 2 Platform performing distributed analysis of genetic/genomic data and any linked clinical/phenotypic information; it should be based on the principle of federated access to data sources, include a federated/multi party authorisation and authentication system, and enable application of appropriate secure multi-party and/or high-end computing, AI and simulation techniques and resources.	No
Outcome 3 Clear description of the roles and responsibilities related to personal data and privacy protection, for humans and computers, applicable during project lifetime and after its finalisation.	Yes
Outcome 4 Business model including an uptake strategy explaining the motivation, patient incentives and conditions for all stakeholders at the different levels (national, European, global) to support the GDI towards its sustainability, including data controllers, patients, citizens, data users, service providers (e.g., IT and biotech companies), healthcare systems and public authorities at large.	Yes
Outcome 5 Sustained coordination mechanism for the GDI and for the GoE multi-country project launched in the context of the 1+MG initiative.	No

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Outcome 6 Communication strategy – to be designed and implemented at the European and national levels.	No
Outcome 7 Capacity building measures necessary to ensure the establishment, sustainable operation, and successful uptake of the infrastructure.	Yes
Outcome 8 Financial support to the relevant stakeholders to enable extension, upgrade, creation and/or physical connection of further data sources beyond the project consortium or to implement the communication strategy and for capacity-building.	Yes

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

3. Methods

In this deliverable, we detail the different costs associated with the 1+MG infrastructure, including set up, scale up, operational, maintenance, and decommissioning costs. In this report the reader will find an indicative budget for the central EDIC activities for a ten-year period and the distribution of those costs over the member states, bearing in mind that we are still fairly early in the project to understand the full scope of activities and costs.

In the previous deliverable D2.1 we identified the different cost items involved in running the 1+MG infrastructure, including central costs and local node costs. The costs were divided into individual tasks/responsibilities in order to have a more granular approach to understanding the complex interface between different parts of the infrastructure. The D2.1 Cost item report also included further development for the central and national functionalities.

Fixed costs are costs that are incurred independently of the level of activity. Variable costs are costs that are directly associated with activity level. In the context of the EDIC this translates into: The more user requests, the more data and the more advanced compute services are used - the more costs are generated. It is extremely important to acknowledge that the Genome-EDIC will have relatively high fixed costs – it will be a huge coordination effort, and will have to operate core functionalities within a tight security framework, even before the first users start using the services – it will take time and a continued investment to run the platform and build the needed capacity. The federated infrastructure will need to successively add more advanced functionalities in collaboration with the national nodes.

Another aspect is how to manage the build-up of service capacity as the users start using the platform. As pointed out in the D2.1 report, scaling the infrastructure with the amount of data coming in, the potential user base, and having the right compute capacity at any given time, will be a challenge. Careful planning, and continued support from the national funders, will be key for the EDIC to fulfil its mission. In terms of hardware investment, the risk goes both ways: the infrastructure must avoid the ***too-little-too-late trap*** – that service providers are unable to honour user requests due to not being able to secure the funding in time, that will allow their infrastructure to grow with data and users, and also the ***too-much-too-soon trap*** – investing in compute capacity that far exceeds the capability of the service platform and the early number of users, since it is impossible to recover the investments done in idle computer hardware. For that reason, it should be considered to establish models that allow to reserve a certain growth budget for a timeframe rather than spend immediately or apply for increased funds only when the need is already apparent.

For the EDIC, ideally, the national investment into compute capacity should be tied into a larger drive towards building data- and platform capacity for Precision Medicine and life science research linked to EHDS and the 1+ MG initiative, allowing the users of the EDIC to make use of resources that are also used outside genomics. Practically, the national investments on supercomputing services must be able to support research on access controlled sensitive data, such as human genomic data.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Since the countries committed to 1+MG have very different approaches and existing setups for national life science research platforms (if one is present), as such it makes little sense to try and calculate specifically which infrastructure needs the EDIC requires at national level. On the supply side - for some countries completely new investments are needed, for others, running EDIC services will be incorporated into already existing platforms and investments. For others again, there might be considerations on not running storage and/or SPE services at all. On the demand side – at this point in time, it is very difficult to calculate how many users the infrastructure will have, and for which services. Particularly running advanced services is unlikely to happen until the service platform has reached an advanced level of maturity and a more comprehensive analysis of business models can be performed.

While this document gives a budgetary overview of 10 years, it should be considered a first iteration and only indicative. The budget loses precision the longer into the future costs are forecasted. This is certainly true for the period after the first three years of operation – the ramp-up phase - setting up the central hub and functional nodes based on a Minimal Viable Product. Without any past reference budgeting operations 5 to 10 years out in the future is extremely difficult. Also, most of the funding for the first 5 years of operation will likely come from research funders and health care funders in the form of direct financial contribution, whereas hopefully as the EDIC grows and creates momentum with users, it will be in a better position to generate income from projects, users or their sponsors.

A number of the development tracks identified in the D2.1 report will not be addressed further now, since they will anyway depend on future funding opportunities as well as the direction of the EDIC going from the ramp-up phase towards steady operations. The development tracks that have been included are the ones that are critical in terms of starting operations – defined as actual users requesting real data from the platform.

What this deliverable offers at national level is an indicative contribution level expected from for each country's to the central costs, and an overview of the minimal requirements in terms of manpower (FTEs) to run the EDIC service layer nationally as a functional element towards the federated European infrastructure, as well as providing examples and baselines for calculating the needs in terms of storage, risk and service management. High Performance Compute (HPC) infrastructure in itself will not be addressed:. there is an assumption that each node has some level of compute capacity to run services initially, or that this capacity will be acquired outside the budget for running services.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

4. Results

4.1. Central hub - Cost calculation and distribution

In this section the cost calculation and budget for the central hub is presented, and the budget is split into three sections:

Organisational costs: The budget for running the central hub, including staff costs, rental of office space, travel, computers etc. These are the indispensable costs that are part of establishing and running an infrastructure. The budget priorities are fairly quick build-up of staff and thus competences that will be critical for the central hub to manage its responsibilities and engagement with user communities, local nodes and other stakeholders.

Operational costs: Since the central hub is also a direct part of the delivery system, this second line of cost items has the costs that are related directly to the operation of the services; these costs are mostly proportional to the usage of the infrastructure. This includes operating a single-point-of-entry help desk, that can issue tickets to product-, service-, and data owners across the infrastructure, and an application request management system, including reviewing the applications and resource management towards the local nodes, and advanced user services. Part of the operational costs for central functionalities may be operated by a local node or distributed across several nodes - particularly for advanced user services.

Central development track: As stated previously, several development tracks are expected to continue during the life cycle of the EDIC – some of which refer to technology and methodology – and will likely require external funding. Others, like the central development tracks, are the top priorities at a central level to move the infrastructure into production, which will require funding even before the EDIC starts operating. This includes implementing a risk management system, supporting the implementation of an eID based authentication mechanism. Developing a clinical user interface and deploying a documentation system for the EDIC. These development tracks must be included into the short-term budget. Some, but not all of these developments are started in the GDI project.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

NOTE: It should also be taken into account that the current plans for the European eWallet by 2026 do not have plans to support fine-grained data access authorisation already foreseen in the GDI project. The Data governance framework for genomic data requires further development to be usable for health dataset access controlled at the desired level by human data access authorities,

4.2. Central hub – 10-year Budget

Table 1 below shows the calculated Budget for Genome-EDIC over 10 years with an addition of two years of pre-operation implementation phase. The budget total is shown both including and excluding development costs.

Table 1: Central hub indicative budget 10 years

YEAR	-2	-1	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10
Organizational costs												
Staffing cost	250	250	500	800	1300	1700	2100	2500	2800	3200	3600	4000
Running costs	50	50	150	175	200	225	250	275	300	325	350	368
Capital expenses and equipment	25	25	50	50	55	60	65	68	70	73	76	79
Organizational costs in cash (1000 EUR)	325	325	700	1025	1555	1985	2415	2843	3170	3598	4026	4447
Operational costs Central Hub												
Total in cash				**286**	**650**	**971**	**1.283**	**1.364**	**1.792**	**1.986**	**2.588**	**2.971**
Central development tracks - projects												
Total in cash		300	650	1.350	1.400	1.000	500	650	600	600	750	700
Total Cash need Central hub (1000 EUR)	**325**	**625**	**1.350**	**2.661**	**3.605**	**3.956**	**4.198**	**4.856**	**5.562**	**6.184**	**7.364**	**8.118**
Cash need Central hub minus dev. (1000 EUR)	**325**	**325**	**700**	**1.311**	**2.205**	**2.956**	**3.698**	**4.206**	**4.962**	**5.584**	**6.614**	**7.418**

Note: Numbers in 1000 EUR, at present value (not indexed over time) : 2.2 Cost report budgets and forecasts.xlsx - Google Sheets

The budget starts with the adoption of the GENOME-EDIC (year 1). Year 1 and 2 are implementation years, focussing on getting the EDIC ready for operation, setting up operations and testing. In year 3 it is expected that the EDIC can start operating services for external users. For this reason, the budget increases fairly quickly going from year 2 to year 3 forward, which also applies for the operational budget.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Central development tracks are prioritised in the implementation phase, since they are key to be able to start operating the infrastructure year 3. The budget also contains an option to have the EDIC's leadership in place and start the risk management development track even before the legal entity is established (-2 and -1). This will require funding outside the EDIC contribution scheme, but will significantly speed up the establishment of the legal entity and for the EDIC to become operational as soon as possible.

Salary level is based on salary levels in Luxembourg. The staff number is approximated according to the number of staff of existing mature research infrastructures, including ELIXIR, BBMRI, EGA and Eurobio-imaging. The current budget is indicative, and may particularly from year 5 forward) look very different in practice. Unlike other research infrastructures, the Genome EDIC is operating an infrastructure with a much broader user base than any life science research infrastructure, catering not only to public researchers but also to clinical actors, policy makers, and industry that may or may not have differing needs and requirements..

4.3. Central hub – Membership distribution key

In this section we present the distribution of the central hub budget between the member states. According to the Statutes for Genome-EDIC currently being drafted, the contribution towards the central budget shall be based as a proportionate share of the members distribution of National Net Income (NNI).

In order to explain NNI, it is important to understand the two more familiar terms GDP and GNI. GDP or Gross Domestic Product is the collective value of a country's goods and services within a certain time frame – usually a year. It measures a country's output as a performance indicator. However, it does not consider that this value may be benefiting people and entities outside that country itself. GNI or Gross National Income is a more precise measurement of a country's wealth, since it adds income from foreign sources, and subtracts income going to foreign sources. Many countries will have little difference between GDP and GNI – however some countries do – particularly if they either channel large shares of the profits to other countries or if they pay salaries to people not living and paying taxes inside the country. NNI considers depreciation/inflation – to reflect the real value of a country's economic activities, acknowledging that some countries have higher debts and thus interest rates and/or high inflation – that reduces the real value of the GNI. So, put short: *NNI reflects the real value of a country's economic activity that goes back into that country directly.*

In table 2, the budget for the first four years of operation is shown, derived from the overall budget from table1. The budget is then distributed among EDIC member states according to NNI distribution.

For each year there is a fixed and a scaled budget part. The fixed part is the 25.000 EUR minimum membership fee, that is applied to all 16 participating countries – which equals 400.000 EUR. The rest of the budget is then distributed according to NNI. The sum for each

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

country per year is the total sum including the fixed and the NNI scaled part. Table 2 also shows the total financial commitment from each country over the four-year period. (Note that the NNI distribution model nor the fixed membership fee has been formally approved yet – only suggested by the EDIC Statutes Task force March 2024).

Deciding on the number of countries to distribute costs towards

In order to have an outline for a cost distribution it was necessary to make a calculated guess as to which countries would be in a position to join the EDIC as founding members or at least sooner rather than later. For that reason, a selection of 16 countries out of the 27 countries involved in the 1+MG initiative, have been included into the distribution key. There is an obvious risk that it may take considerably longer for countries to join than what is planned for. However, it is still important to do this for two reasons:

1) A realistic budget and cost distribution cannot assume that ALL 1+MG partner countries will be able to join as founding partners. National decision-making processes are inherently slow, and subject to changing priorities. So there needs to be a calculated assumption about who could potentially join quickly, while acknowledging that in real terms this might take longer than what would be ideal (1 to 3 years).
2) If on the other hand, only the countries who can say with some certainty that they will become founding members are included, the costs will have to be distributed among very few countries, which in itself creates a clear disincentive for getting countries to join early on.

The inclusion of a specific country into the cost calculation is to some extent based on an estimation as to which countries are active in the GDI project. However, this is only indicative – no country has, at this point, committed to joining the EDIC as a founding member. Also, Europe consists of a relatively homogenous group of smaller countries and a few very big countries: (Spain, France, Italy, and Germany (Currently the UK is not qualified to become a member) . How many of the big countries will join early will have a significant impact on the cost distribution. While all of the big four have joined the 1+MG vision, only ONE of the big countries is included into the cost distribution calculation – in this case France - to ensure that the indicative cost distribution will not be too dependent on several of the big countries joining early.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Table 2: Cost distribution EDIC operational phase year 1-4 – not including development costs.

	Year 1 - EDIC Budget	Year 2 - EDIC Budget	Year 3- EDIC Budge	Year 4- EDIC Budget	Total 4 years	Pr. year distributed
Total budget withou	700.000 EUR	1.311.000 EUR	2.205.000 EUR	2.956.000 EUR		
Fixed budget	400.000 EUR	400.000 EUR	400.000 EUR	400.000 EUR		
Scaled budget	300.000 EUR	911.000 EUR	1.805.000 EUR	2.556.000 EUR		
France	136.150 EUR	362.526 EUR	693.753 EUR	971.999 EUR	2.164.428 EUR	541.107 EUR
Netherlands	62.867 EUR	139.990 EUR	252.833 EUR	347.627 EUR	803.317 EUR	200.829 EUR
Belgium	47.740 EUR	94.053 EUR	161.818 EUR	218.743 EUR	522.354 EUR	130.588 EUR
Sweden	46.388 EUR	89.948 EUR	153.685 EUR	207.226 EUR	497.247 EUR	124.312 EUR
Romania	43.929 EUR	82.480 EUR	138.888 EUR	186.273 EUR	451.570 EUR	112.893 EUR
Czechia	39.242 EUR	68.247 EUR	110.687 EUR	146.339 EUR	364.514 EUR	91.129 EUR
Denmark	38.924 EUR	67.284 EUR	108.779 EUR	143.636 EUR	358.623 EUR	89.656 EUR
Norway	38.828 EUR	66.991 EUR	108.199 EUR	142.815 EUR	356.834 EUR	89.208 EUR
Portugal	37.752 EUR	63.725 EUR	101.727 EUR	133.651 EUR	336.856 EUR	84.214 EUR
Ireland	36.398 EUR	59.613 EUR	93.580 EUR	122.114 EUR	311.705 EUR	77.926 EUR
Finland	34.920 EUR	55.124 EUR	84.687 EUR	109.520 EUR	284.252 EUR	71.063 EUR
Croatia	29.807 EUR	39.597 EUR	53.921 EUR	65.954 EUR	189.278 EUR	47.320 EUR
Slovenia	27.955 EUR	33.972 EUR	42.777 EUR	50.173 EUR	154.876 EUR	38.719 EUR
Estonia	26.832 EUR	30.562 EUR	36.020 EUR	40.605 EUR	134.018 EUR	33.504 EUR
Luxembourg	26.820 EUR	30.528 EUR	35.952 EUR	40.509 EUR	133.810 EUR	33.452 EUR
Malta	25.448 EUR	26.360 EUR	27.695 EUR	28.816 EUR	108.319 EUR	27.080 EUR
TOTAL	**700.000 EUR**	**1.311.000 EUR**	**2.205.000 EUR**	**2.956.000 EUR**		

(Source: Eurostat 2022; see section 7): [2.2 Cost report budgets and forecasts.xlsx - Google Sheets]

For budgeting purposes and for countries to be able to evaluate and decide on whether or not to join the EDIC, the contribution of each country towards the central costs have been calculated based on a suggested membership roster, and are shown in Table 2. Although only indicative it should give a fair and realistic number in the foreseeable future- although in practice the timing of each country's joining may be different. This is why the cost distribution considers an "ideal" number of countries joining, and locks the contribution fee over the first four years of operation. This will allow countries to make financial planning for a four-year period based on a reasonable figure that does not change depending on which other countries join in that period. The other side of that coin is that financial planning for the EDIC will be challenging, since it is likely that the income from member states will be less than what is budgeted for. The only way to navigate this, is to downscale activity levels accordingly.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

To balance the need for budgeting clarity for the EDIC and the need to provide a transparent and fixed membership fee for the countries, it is vital that the Statutes include provisions that do not allow countries to leave the infrastructure within the first five years of joining. Otherwise it will be completely impossible for the EDIC to do any sort of financial planning.

The following observations can be made from the budget and cost distribution in Table 2:

- The impact of the fixed membership fee is most significant in the early years of operation, and levels out the cost distribution among the partners.
- The cost distribution assumes that all the countries included are in a position to join as founding members. This is an obvious financial risk, since that may not be feasible. However, the addition of the fixed contribution fee, to some degree, alleviates that risk – particularly if bigger countries are not able to join as founding members.
- The budget reflects a fairly rapid build-up phase, which has been prioritised in order to keep up momentum.
- At this point there are no financial provisions that would incentivise countries to join early. This requires further investigation and needs to include a better understanding of the membership structure, including, duties, rights and obligations included into the Statutes.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Cost distribution including development costs

In the following table, the cost distribution is applied again. This time it includes the development costs identified and included into the overall budget in Table 1.

Table 3: Cost distribution EDIC operational phase year 1-4 - including development costs.

	Year 1 - EDIC Budget	Year 2 - EDIC Budge	Year 3- EDIC Budg	Year 4- EDIC Budge	Total 4 years	Pr. year distributed
	1.350.000 EUR	2.661.000 EUR	3.605.000 EUR	3.956.000 EUR		
Fixed part	400.000 EUR	400.000 EUR	400.000 EUR	400.000 EUR		
Distributed part	950.000 EUR	2.261.000 EUR	3.205.000 EUR	3.556.000 EUR		
France	376.975 EUR	862.701 EUR	1.212.453 EUR	1.342.499 EUR	3.794.629 EUR	948.657 EUR
Netherlands	144.912 EUR	310.391 EUR	429.546 EUR	473.851 EUR	1.358.701 EUR	339.675 EUR
Belgium	97.009 EUR	196.382 EUR	267.937 EUR	294.542 EUR	855.871 EUR	213.968 EUR
Sweden	92.729 EUR	186.194 EUR	253.495 EUR	278.519 EUR	810.938 EUR	202.734 EUR
Romania	84.941 EUR	167.660 EUR	227.222 EUR	249.369 EUR	729.192 EUR	182.298 EUR
Czechia	70.098 EUR	132.334 EUR	177.148 EUR	193.811 EUR	573.391 EUR	143.348 EUR
Denmark	69.094 EUR	129.944 EUR	173.759 EUR	190.051 EUR	562.848 EUR	140.712 EUR
Norway	68.789 EUR	129.218 EUR	172.730 EUR	188.909 EUR	559.646 EUR	139.912 EUR
Portugal	65.383 EUR	121.111 EUR	161.239 EUR	176.159 EUR	523.891 EUR	130.973 EUR
Ireland	61.095 EUR	110.906 EUR	146.772 EUR	160.108 EUR	478.881 EUR	119.720 EUR
Finland	56.414 EUR	99.765 EUR	130.981 EUR	142.588 EUR	429.748 EUR	107.437 EUR
Croatia	40.221 EUR	61.227 EUR	76.353 EUR	81.976 EUR	259.778 EUR	64.944 EUR
Slovenia	34.356 EUR	47.267 EUR	56.564 EUR	60.021 EUR	198.209 EUR	49.552 EUR
Estonia	30.800 EUR	38.804 EUR	44.567 EUR	46.710 EUR	160.880 EUR	40.220 EUR
Luxembourg	30.764 EUR	38.719 EUR	44.447 EUR	46.577 EUR	160.508 EUR	40.127 EUR
Malta	26.418 EUR	28.376 EUR	29.785 EUR	30.309 EUR	114.889 EUR	28.722 EUR
TOTAL	**1.350.000 EUR**	**2.661.000 EUR**	**3.605.000 EUR**	**3.956.000 EUR**		

(Source: Eurostat 2022, see also section 7): 2.2 Cost report budgets and forecasts.xlsx - Google Sheets

As can be seen from the table, the addition of development costs significantly increases the overall cost and the contribution from the countries – in most cases it doubles the contribution fee. The inclusion of development costs is critical for the EDIC to go into an operational phase. Installing an ISO certified risk management system and performing a DPIA is no small feat, and how to do this exercise requires significant consideration.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

The same goes for a documentation system and a clinical user interface framework. Currently no eID based solution that the EDIC (or EHDS) can depend upon has been implemented, but that must be included into the overall framework. To bridge those gaps there must be funding in place to develop and implement these critical systems moving from a demonstration level into production. But this also raises some questions.

- Should continued development costs be put on the founding members, and if not how else to cover those costs?
- Should member states joining later, donate into the development tracks, covered by others, as they join?
- To what extent can the EDIC plan for external funding for these bridging development tracks?
- To what extent can development funds be allocated to non-member entities?

- eID provides a solution for a trusted electronic identification of the users. However, data access governance needs additional user information to be present. In GDI and EHDS this is mainly data/resources access permissions in a standard format to be carried to the data holders with the eID information, such that this access can be verified at the time of processing by the party issuing the access.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

4.4. National costs and obligations

In this section the overall costing framework for the national participation in the Genome-EDIC is presented and discussed. This will however NOT be as detailed as the national contribution towards the Central Hub. The reason for this is that no two countries are alike. They have different starting points and they have different structures and frameworks in place already. For that reason, what already exists and what needs to be built is very different across the countries. Particularly, the ways to structure data inclusion and compute services will be very different across the nodes. Obviously also, the actual costs vary considerably across countries. For these reasons this section will highlight and provide guidance on what the requirements are and what frameworks need to be in place, depending on the ambition at national level.. This will be presented in FTE's, and if possible, there will be some indicated costing baselines for storage and risk management in EUR.

Essential item per country is providing European interfaces from the country to plug national operations into the European federation.

4.4.1. Discussion on cost parameters

EDIC coordination and management: On a national level, it is expected that there is a dedicated manager that is legally and financially liable for the running of the national node, including financial, legal and contractual management. This includes being responsible for service provisioning towards users, and towards the EDIC for adhering to both agreed service levels, adaptation of technical requirements and the EDIC's risk management requirements. Management is also responsible for reporting to the EDIC, and responsible in general for adherence to data protection regulation as a data processor for the EDIC. At the coordination level there will need to be resources to handle communication with stakeholders and towards the rest of the EDIC ecosystem, and quality assurance services, to ensure that the performance of the national node is monitored.

National node service delivery: The national node, assuming the node will be part of the EDIC service delivery system, must provide the standard functionalities defined for the nodes as services within the EDIC ecosystem. In addition, any node must perform its own data management and provide a help desk to engage with the users.

Part of the national service delivery system is to integrate services with existing compute and storage resources OR to run high performance compute services within the node's own system independently. It is likely that it will take some time after deployment of the Minimal Viable Product (MVP) before there is a need for a full-service package including advanced compute services; that will in large part also depend on the availability of data resources at national level and the attractiveness of those resources for users.

Since the countries will have VERY different starting points in terms of existing infrastructure and capacity, how a compute system should be set up and run, and at what capacity will not be

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

addressed further in this deliverable, but should be explored further, both within the EDIC framework and at national level, as the data resources and the needs from users are better understood, both in scale and capability. There will be ample opportunity for national discussions on how to build hardware capacity and national compute capability and capacity for the EDIC moving from an MVP and address how these needs fit in with other life science compute needs, implementation of genomics into health care and provide synergy towards EHDS. Moving from a short time frame of 1-4 years into the next phase, the infrastructure will likely see a massive increase in users and use cases, and at that point the EDIC will be in a much better position to build capacity based on a business and investment model that mixes revenue from several sources.

National data subject interface: As agreed with the GDI governing partners, the direct interaction between data subjects and the EDIC infrastructure must be handled at a national level, as part of the overall national system for consent management and data subject engagement under GDPR. As such this means that these functionalities must be in place, for the EDIC to operate at national level, even if the EDIC node itself may have little impact on these services. The required functionality includes running a data subject information portal – an online portal in which data subjects can find information regarding the use of their genomic data in the context of the Genome-EDIC, and possibly their (health) data in general also for other data reuse-infrastructure. There must be a dynamic and updated consent management framework in place that allows the EDIC to offer these data to users based on the scope of the nationally collected consents. That means having in place coherent and documented consent form and procedures as well as a registration system that will allow subjects to be removed from the data pool, either permanently or for specific purposes. There must also be in place a system for reporting and handling incidental findings, that information from the EDIC data processing can be fed into.

Risk management framework:The following section is based on the authors' past experience and estimation as to the risk management requirements of the EDIC - which may or may not represent the views of others. A lot of details on scope and cost still remains:

It is expected that setting up a proper risk management framework will cost a national node between 100.00 to 500.000 EUR for consultancy fees and between 12-24 Person-Months to implement - depending on the complexity of the national organisation and already implemented risk management systems. Running a risk management system requires the involvement of a CISO (Chief Information Security Officer) and a DPO (Data Protection Officer) and 30 kEUR annually in recertification costs. Risk management is critical for a distributed data infrastructure on sensitive data. For the EDIC to run, it will require rigorous protocols, controls and documentation in order to function and maintain a zero-trust risk framework based on data-protection-by-design-and-default.

The EDIC will likely need to have an ISM (information security management) system in place that is ***certified*** according to ISO 27001 (Information security) and ISO 27701 (on Privacy) – on ALL if its activities – that also includes service provider management. For the EDIC HUB to function as a

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

data controller, it will depend on a Service Level Agreement (SLA) with the local nodes, establishing clear guidelines that the nodes must adhere to and accept to be controlled on.

What all the nodes being part of the service delivery system will need to do, in compliance with GDPR, is to do a Data Processing Impact Assessment (DPIA) on all the systems and software, interfaces, Standard Operating Procedures (SPOs) and staff before any real data can be accessed by external users. Since this is a high-risk area for the individual, it is likely that in most countries data protection authorities will need to approve/be consulted on the DPIA before actual processing of real data can take place.

The countries are strongly encouraged to support the nodes in building the data processing infrastructure based on the principles outlined in the NIS2 directive on Cyber security and implemented through an Information Security and Privacy Management System (ISPMS) as this is a vital exercise, in order to ensure compliance and adherence to both the EDIC requirements, GDPR and national regulation in a very dynamic digital environment.

The nodes' ISPMS should ideally be ISO 27001 and ISO 27701 certified[2] on ALL of its operations - at the very least on ALL the operations involved in the Genome-EDIC activities (*the notion of being ISO 27001 compliant is not desired – since that is very much subject to an entity's own interpretation).* However, it should be recognized that becoming ISO 27+ certified is a costly and time-consuming endeavour, requiring extensive analysis and documentation.

National node will need access to CISO and DPO resources, who have the capacity to engage with the EDIC and guide the node, including performing the DPIA and build the ISPMS including performing the initial risk assessment and put in place controls and rules for the operation in compliance with the risk management regimes employed. Risk managers typically make use of a documentation system like XWiki to handle versioning of the substantial amounts of documents related to risk management – risk assessments, rules, controls, reports etc. These people must be trained in risk management and legal requirements with regard to data protection, and the level of engagement of these resources needs to match the level of activity at the node. It will likely not be enough just to depend on a CISO and/or a DPO that is managing risks for an entire university, as these resources will likely not be able to allocate large portions of their work hours towards the EDIC's activities.

The nodes must have an ISEA3000[3] Type 2 audit conducted each year. Whereas an ISO27+ audit will measure the performance of an entity's risk management system, an ISEAT 3000 Type 2 will audit the actual risk management performance of an entity (not its risk management system) and give an audit statement on that. An ISEA 3000 Type 2 audit is completely independent of the status of the risk management system and whether that system is certified or not – it gives a

[2] ISO/IEC 27001:2022 - Information security, cybersecurity and privacy protection — Information security management systems — Requirements

[3] What Is The ISAE 3000 And How It Can Help Your GDPR Work (cyberpilot.io)

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

performance evaluation. It is a national responsibility to fund the audit, provide the documentation for the SOP's to the auditors, participate in the audit process, and address the findings that are being reported.

The risk management requirements stated above, are deemed vital in order to ensure that the individual's data is well protected and that the integrity of the EDIC is also protected, since the EDIC itself has reputational risks, as data controller and for the processing activities at local level.. There can be no doubt, that while cumbersome and costly, proper risk management is critical for this type of operation, and that any data breach incidents that may happen, will be a massive liability in terms of loss of reputation.

Baseline for national cost calculation on risk management framework: In order to have a functioning risk management system (besides staff profiles within that area) the node needs to do a Data Protection Impact Assessment (DPIA), set up a Risk management system and define rules and controls. This is a considerable workload that is likely to involve most specialists working within the node, and one that is likely to require outside expertise to be done properly.

To make an estimate of the costs involved, we can look at an existing organisation that has gone through these efforts. The Danish National Genome Centre (DNGC), who is ISO certified on information security and privacy, conducted a DPIA and put in place a risk management system in a year – after some initial effort had already been conducted. DNGC spent at least 1.5 MEUR on getting the system in place – most of which was spent on external consultancies, with a core staff of four people working with the external consultants. Arguably, the National nodes in the Genome EDIC will not have the same risks to manage, as do DNGC which is operating not only a platform for secondary use, but also a complex clinical pipeline for primary use - direct patient care. Also, the extensive use of external consultants in DNGC was a choice made in order to speed up the process.

The nodes should be in a position to do a DPIA and install a risk management system for considerably less and also in a shorter timeframe – But only if there is a commitment to do it, at management level, as it will require a considerable allocation of time towards it. It is important that the DPIA is conducted on all activities related to the EDIC.

At least 12-24 person-months over the course of 6-18 months from own staff should be allocated to the task of conducting a DPIA and implementing a Risk management system, with an additional one-time budget of 100,000-500,000 EUR for external consultants, depending on national cost level and complexity of operations. Certainly, the complexity of the interaction between the EDIC node and its host institution organizationally and technically, will directly impact the complexity of the DPIA and thus the implementation efforts for an ISPMS.

An ISO certification will likely be a one-time cost of 20,000 EUR with biennial recertification costing around 15,000 EUR. An ISAE3000 Type 2 audit will cost around 20,000-25,000 EUR and will need to be conducted each year. These are external fees that need to be budgeted for at national level.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Data inclusion: For the EDIC to maintain its credibility as a high-quality data space, based on selectively curated and standardised data, there must be mechanisms and resources available to ensure that data accessible through the platform is quality and sanity controlled. Ideally the data is captured at the source according to standards and protocols, which will ensure that data coming from multiple sources can easily be integrated into the database. However, this will likely not be the case across multiple actors across multiple countries, which means that there will likely need to be resources available in or around the national node, that do metadating and documentation work on the data sets coming into the platform. Mechanisms and resources must be available in or in close collaboration with the node, in order to engage with data holders on standardisation, on guidelines, and how to improve data quality and support the adaptation of GA4GH standards into research and clinical practice. Continued maintenance of the country's data within the metadata catalogue will be required.

It is important to note, that on the data side, the ability of the national node to engage with clinical and research actors may differ significantly. Ensuring that the data can be adopted into the EDIC data portal cannot be the responsibility of the EDIC. Only at national level will there be an opportunity to have a dialogue with data holders on ways to standardise, and thus minimise the work that must be performed ex post to include data.

Storage: Estimates for the storage cost for the data in the EDIC with current technology come to 14-16 EUR per whole genome sequence. If in a country 0,4 % of the population is sequenced every year, total storage costs could be 60,000 EUR per million citizens for the first year, and growing with 60,000 EUR per year as more data is added over time (i.e. after 10 years, the costs are 600,000 EUR per year (estimate, based on 30x CRAM).

The EDIC will expect the member states to cover storage for the data that forms part of the EDIC data resources and financially support the incremental build up and operation of data storage. Each country can decide to build new storage resources directly linked to the EDIC or charge a subcontractor to host data on behalf of the EDIC, assuming that the data can be fully queried and processed under the EDIC framework.

Any storage system within the EDIC ecosystem, must include an off-site storage facility for the genomic data included into the metadata-catalogue, ensuring that the core data sources can be recovered. Having a maximum loss time of 24 hours 24/7 is significantly more challenging than a requirement of 48 hours during work days.Data ingested into the data storage system must be synced with the back-up system directly. Only backup storage can be a tape solution. There must be sufficient bandwidth on the system to handle peak loads moving data from storage to data processing.

Referring to the Storage discussion in the D2.1 Cost report, the actual cost of storage can be fairly precisely calculated based on current storage technology and current short read WGS methodology.

Baseline for national cost calculation on storage (estimate as of April 2024):

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

- **WGS data size: 100-120GB per. WGS:** 1 WGS (30X coverage with current sequencing techniques) as short read lossless CRAM plus gVCF and VCF files constitutes 80 GB of data. With back-up and some operational doubling, this may constitute up to 120 GB of data (Short read 30X Lossless CRAM), depending on the availability of phenotypic and clinical data and pipeline regime (This number has been adjusted after the publication of the 2.1 report.
- **Cost of storage: 1PB of data - 400,000 EUR over three years** - 1 Petabyte server storage and off-site backup requires a hardware investment of 400,000 EUR, including overhead over three years (2024 prices). 1 Petabyte of server storage will be able to accommodate up to 8-10.000 WGS's in a single copy.
- **Unit cost for storing one WGS for a year should be between 14- 16 EUR maximum.**
- **Clinical production: 3-5.000 WGS's produced each year per 1 million citizens**. Based on a forecast from the Netherlands and experience from Denmark clinical production of WGS data seems to level off, at between 3- 5.000 WGS's pr. one million citizens.

The calculation above does not take the following aspects into account:

A) If at any point long-reads WGS becomes part of the data pool, storage requirements will increase dramatically – up to 1TB pr. WGS.

B) It is likely that data storage costs will decrease over time, as it has done historically.

C) The precision of the sequencing tools MIGHT reduce the WGS data size, however past experience suggests that this may not be the case.

4.4.2. Estimation of staff level – nationally

Disclaimer – Since this is still fairly early in the GDI project, the following presentation is only indicative. It is extremely difficult to budget an activity in terms of FTE's and actual cost, without having a full picture of the operational requirements, which is also in large part, depending on the availability of data and data processing resources.

The following estimation of resources needed within a relatively short timeframe, to be able to have a functioning federated infrastructure in place, that would allow users to access actual data. The estimation does not factor in any specific national pre-condition or resource-availability. For what it is worth, the following is a starting point for a more in-depth discussion within the EDIC and with the national partners with respect to resources needed to operate the infrastructure.

FTE's to operate the national node.

For basic operations, and to some extent independent of the actual number of users or scale of activity it is estimated that the national node requires the following full-time equivalents (FTE's.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Role	FTE Minimum	Remark
Coordinator / Node lead	0,5	Liable for data processing at national level
Financial and Legal Support	0,5	Contracts, agreements, cash flow, budget, accounts
Compliance	1-2	Chief Information Security Officer, Data Protection Officer
Help Desk Management	2-3	Quality assurance, software integration and update. User ticket management, compute system support, User management – **large scaling requirements depending on activity level**.
Data Inclusion Management	1-2	Documentation and standardisation. **Must scale with incoming data, and quality thereof.**
Total	**5 - 8**	

2.2 Cost report budgets and forecasts.xlsx - Google Sheets

Obviously, this is a very early estimation on basic operations – and does not consider operating an advanced platform for many different users across health care and research. It should be considered the minimum capability to build, service and initiate running the basic infrastructure. For a longer discussion on the scaling issue please consult the 2.1 report on cost items. At national level it is critical to plan that the level of activity will increase dramatically going from a MVP for a few users to operating an effective infrastructure catering to thousands of individual users and hundreds of different projects. If scaling is not incorporated into the national funding model for the node, both on people and hardware it is likely that the node runs into the "success trap" – not being able to accommodate all the potential users because of a funding or investment ceiling. Some costs are independent of usage and data volume, while others directly scale with usage and data volume.

Continuation of efforts from GDI towards EDIC

It is to some degree expected/assumed that the same people/entities who are currently building the infrastructure (The current GDI Pillar II community) will also be the ones operating the infrastructure in the context of the EDIC. While the national node structure is completely a national decision, this requires some careful consideration in terms of ensuring a smooth transition from the GDI project towards an operational infrastructure. For some countries, there is

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

likely to be a time gap between the end of the GDI project and a national decision to join the EDIC. How to ensure continuation of coverage and keep the momentum in the interim period needs to be considered. Countries should not assume that developers working in academia have the desire to work within a strictly managed operational service environment. Rather, it would be important to consider how to move the infrastructure from a development and conceptual framework into operation, which may need plans for how to transfer tasks, know-how and capability from developers to a devops and sysadmin level, also to ensure that the service operation is matched by job profiles at various levels of education and skill, thus expanding the recruitment base.

5. Subjects for further discussion and analysis

Development cost funding streams: As mentioned in the report, key development tracks need funding within a short timeframe. In the indicative Budget in Table 3, these costs are put on the member states. There might be other vital elements that still need to be developed going from the GDI project into production. This depends to a large degree on how the GDI project develops from now until the end of the project and how well the operational requirements can be defined in advance. Moving further, it should be investigated to what extent the EDIC short term development costs can be covered or supported through external funds, ea. through the Digital Europe Programme, The European Rare Diseases Research Alliance (ERDERA), IC PerMed, EH4Health or other sources. These funding programmes will no doubt become important in supporting development tracks within the infrastructure long term. But in order to be successful the EDIC and the local nodes MAY need to work together in producing the relevant grant applications, potentially prepared by small pilot projects - all funded through EDIC sources.

Also, as mentioned in the section on development costs, there is a need to better understand the interface of development teams across the countries, and also to understand how to put development projects into production – ensuring that the software is manageable in a production environment: Risk assessed, maintainability, based on open licence source code, etc.

Licencing costs: At this point there has been no discussion on licensing costs, but dependent on how the EDIC develops, this is likely to be something that needs to be managed. The EDIC must be able to function, independently of specific service providers, or at least be in a position to change service provider, should the need arise, with enough redundancy to ensure continuation of services. This requires that the functionalities are understood, specified and separated from any one tool used.

For commercial tools, it must be recognized that there will be licensing costs associated with using commercial software and also a risk of vendor-lock-in mechanisms. While vendor-lock in should be avoided, it must be recognized that commercial software is often used in (genomic)

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

research and that they are cost-effective for what they do. Also, genomic pipelines[4] are often based on commercial tools – certainly in health care. Since users will likely need access to commercial tools, this must be addressed in the funding and business model – which commercial tools should be offered (as standard), at what cost and to whom is the cost charged? Using commercial software must also form part of a risk assessment. Since source code is not available and since many commercial tools are cloud based there are inherent compliance challenges in using commercial software.

EuroHPC collaboration: For long term sustainability, it should be explored, to which extent a future collaboration can be set up between the EDIC and the EuroHPC programme with a goal of making EuroHPC compute resources suitable for certain types of operations on 1+MG data, subject to feasibility and acceptable levels of interdependency and liability in terms of performance and risk management.

Costs for FAIR : On a more generic level, further discussions are needed on data inclusion and data storage. Particularly it should be made clear how roles and thus distribution of tasks and obligations on data inclusion and long-term storage could be distributed across health care and research. Particularly for research data, it should be explored how to stimulate inclusion of FAIR principles into research funding – including allowing long term storage (within the EDIC) to be eligible costs within EC and national research funding schemes.

Costs for storage: For health care it should be assumed that the infrastructure itself and the storage of data should not operate in isolation. The same data, that will be used under the EDIC framework, will be used for health care purposes and EHDS – at a national level consideration must be made, on how to build cost-effective storage solutions that will accommodate all needs. (Although in some countries there is currently a very clear division between health care and research - particularly on storing data - even if data should be stored in a clinical setting, which makes finding synergies more difficult).

Risk management framework synergy: Since all the operational nodes must do a DPIA and install a risk management system, there should be a possibility to create synergies – The fundamental data processing will be similar – since the EDIC is operating a federated system using standardised tools. Having a common framework, and allowing countries to piggyback on the work already conducted by others, will significantly reduce the workload in performing the risk assessment, installing rules and controls and documentation -which could also include operating similar documentation systems. That could be achieved among other things by investing in setting up a support structure from the Hub towards the nodes – who will anyway need to comply with requirements set by the EDIC. A proper federated risk management system

[4] *A genomic pipeline is a set of complex algorithms (tools), which is used to process sequence data, in order to generate a list of variants or assemble a genome(s)*

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

should allow nodes to notify the community about potential risks and ensure that those risks are mitigated across the ecosystem.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

6. Next steps

This report forms part of the baseline for the countries to make informed decisions on joining the EDIC, by giving realistic estimations on the funding level required to support the central operations, and addressing the key factors to consider at national level, for the overall budgeting for the membership, working out what funding requirements are needed, what is already established and what needs to be in place to operate the node on a basic level at first and support the scaling up of the node over time.

Referring back to the D2.1 report, that showcased just how big of an effect the actual number of users and use cases will have on the needs for processing capacity, it is important that the countries understand that operating the EDIC on a basic level, will not be able to accommodate advanced use cases and a very large user base without additional investments in compute capacity and support structure. It is likely that, as the business model develops, the infrastructure will be able to charge a reasonable fee for its usage, which can include writeoff for the investments. However, the financial obligations (and hence the risks) for the investments into the scaling of the infrastructure will be with each of the nodes. Nevertheless, it will be important to create the necessary capacity, and hence that the local nodes have the possibility to get access to resources that will match needs in the longer perspective.

Going forward it is essential to develop the business model for the EDIC, which will be the end goal of the WP2 Long-term Sustainability framework. Understanding and documenting the financial impact of precision medicine in general and specifically the impact of the EDIC in supporting better health care in a health economy perspective, will be essential, in ensuring the continued support for new investments in infrastructure and services.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

7. References

NNI distribution - EU27+NO, UK, CH 2022

	NNI total (MEUR)	NNI share %
Total	**22.855.406**	**100,00**
Germany	4.290.478	18,77
United Kingdom	3.027.012	13,24
France	2.938.597	12,86
Italy	2.492.401	10,91
Spain	1.790.790	7,84
Poland	1.345.667	5,89
Netherlands	1.001.133	4,38
Belgium	601.197	2,63
Sweden	565.459	2,47
Switzerland	547.798	2,47
Romania	500.440	2,19
Austria	473.703	2,07
Czechia	376.521	1,65
Denmark	368.135	1,61
Norway	365.589	1,60
Portugal	337.151	1,48
Greece	318.389	1,39
Hungary	304.494	1,33
Ireland	301.351	1,32

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

Finland	262.272	1,15
Slovak Republic	164.980	0,72
Croatia	127.082	0,56
Lithuania	113.012	0,49
Slovenia	78.113	0,34
Latvia	55.252	0,24
Estonia	48.422	0,21
Luxembourg	48.126	0,21
Malta	11.842	0,05
	22.855.406	

Source: EUROSTAT 2022: Data table for: Net national income, Total, US dollars/capita, 2022 or latest available,Aggregate National Accounts, SNA 2008 (or SNA 1993): Disposable income and net lending/borrowing: 2.2 Cost report budgets and forecasts.xlsx - Google Sheets

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.