Sustainability models for AgBioData member databases

January 2024

Prepared by Sabarinath Subramaniam, Director of Business Development, Phoenix Bioinformatics and Josh Young, Executive Director, Phoenix Bioinformatics

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

Phoenix BIOINFORMATICS
39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

1. Introduction

1.1 About this report

This report was prepared by Phoenix Bioinformatics as partial fulfillment of the deliverables within NSF Award # 2126334, "RCN: Reimagining a Sustainable Data Network to Accelerate Agricultural Research and Discovery". Information for this report was gathered by Phoenix Bioinformatics between January 2023 and November 2023, with the assistance of principal investigators of various AgBioData member databases. The final version of this report was submitted on Dec 31, 2023.

1.2 Disclaimer

The findings described in this report are highly dependent on the accuracy of the information provided to Phoenix Bioinformatics by AgBioData member databases and any errors in the underlying data will require correction before taking any recommended action.

1.3 About Phoenix Bioinformatics

Phoenix Bioinformatics was founded in 2013 as a nonprofit 501(c)3 organization. Our mission is to assist scientific data repositories and other research cyberinfrastructure components in developing innovative and sustainable funding support mechanisms to ensure long-term sustainability. We pioneered our novel approach with TAIR, the Arabidopsis Information Resource, a widely used plant genome database, and were successful in replacing grant funding with a $1.1M/yr revenue stream from users without significantly impacting usage of the resource. Since that early success, we have been working in partnership with a range of scientific resources to assist them in finding sustainable revenue streams. Current partners include Repbase, a database of genomic DNA repeats and transposable elements, CIPRES, a resource for phylogenetics and population biology, and, CyVerse, a national cyberinfrastructure for life science research as well as training scientists in using high-performance computing resources.

2. Background

2.1 Genomic, Genetic, and Breeding Databases GGB Databases

Genomic, Genetic, and Breeding (GGB) databases serve and respond to research and breeding stakeholder communities to provide value-added curated data and tools that meet stakeholder needs. To ensure that researchers continue to have access to reliable, high-quality, curated, and FAIR data in the future, GGB databases need to plan and develop infrastructure, strategies, and tools to ensure long-term sustainability of GGB data and GGB Databases. The AgBioData consortium (https://www.agbiodata.org) has agricultural biological databases with the mission of consolidating standards and best practices for acquiring, displaying, and reusing genomic, genetic, and breeding (GGB) data. Formed in 2015, the consortium involves 40 GGB databases

Phoenix BIOINFORMATICS
39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

and over 200 members, including database curators, researchers, librarians, and anybody that works with agricultural data.

2.2 AgBioData member Databases

Table 1 (Appendix 7.1) shows the list of AgBase member databases that were submitted with the RCN grant proposal. We excluded VectorBase because the resource is no longer part of the AgBioData consortium; TAIR and CyVerse, because they have implemented a sustainability plan based on subscriptions; and Araport because the database has become defunct with the tools and data now being hosted by other databases. Of the 40 AgBioData member databases, the survey was sent to 25 PIs representing 36 databases.

2.3 Rationale for sustainability efforts

The Sustainability of GGB Databases and Resources has emerged as an important issue. Most GGB Databases rely on short-term funding for a majority of their operating costs and are vulnerable to loss of personnel and knowledge if funding lapses, even while demand for their services by researchers continues to increase. Financial uncertainty and gaps in funding not only hamper efficient operation but also limit long-term planning, potentially resulting in higher costs for data access. Loss of database funding can also lead to permanent loss of valuable data and software gathered and built at taxpayer or industry expense, and slows the progress of research (56–60). Also, while there is a need to ensure that the data for new genomes is made accessible in a timely and cost-effective manner, it is simply not feasible nor desirable to create a GGB Database for every species. Consolidated and standardized database resources are needed.

To plan for the future of GGB Databases and data gather feedback from member databases and stakeholder communities via surveys and assess different sustainability options ranging from support for individual database projects to federated /cost-sharing models that can be applied across many databases.

3. RCN Aim 4.2: Conduct detailed analysis and modeling of sustainability solutions for representative AgBioData member databases.

This report focuses on achieving the goals set within Aim 4.2 of the RCN grant. We plan to delve into an in-depth study of sustainable practices and their potential application to each AgBioData Member Database and Resource. This will encompass a broad spectrum of strategies, categorized into three main groups:

3.1 Exploration of New Funding Avenues

We aim to explore diverse funding possibilities such as voluntary membership schemes, fees for data deposit, subscription models, 'freemium' options, crowdfunding, corporate backing, and philanthropic contributions.

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

3.2 Investigation of Cost Reduction Techniques

We will scrutinize potential cost-cutting strategies like collaboration in curation or software development, outsourcing curation to the community or data providers, and utilizing shared infrastructures, all to enhance sustainability and diminish operational expenses.

3.2 Assessment of Policy or Procedural Revisions About Grant Funding

Our focus will be on identifying methods to bolster databases' ability to secure a portion of research funds dedicated to data creation. Furthermore, we will assess the impact of a central coordination body in aiding sustainability efforts. These approaches for individual database sustainability will be designed to align with and contribute to the overarching long-term roadmap.

4. Exploration of New Funding Avenues

4.1 Voluntary membership schemes

Voluntary membership schemes represent a strategic approach to achieving financial sustainability for biological research data resources. This model relies on users' willingness to pay for access to enhanced services, creating a revenue stream that supplements traditional funding methods.

4.1.1 Structure and Benefits

In a voluntary membership scheme, users—ranging from individual researchers and academic institutions to corporations and nonprofits—can opt to pay a membership fee. In return, members receive exclusive benefits. These could include early access to new datasets, advanced analytical tools, priority customer support, or opportunities to influence future resource development. Tailoring these benefits to meet the specific needs and preferences of different user segments is crucial to maximizing membership appeal.

4.1.2 Community Engagement and Ownership

This model cultivates a sense of community and ownership among users. Members who contribute financially are likely to be more engaged and invested in the resource, providing valuable feedback and advocating for its use within their networks. This heightened engagement can lead to a more dynamic, user-driven resource, constantly evolving to meet the needs of its active user base.

4.1.3 Sustainability and Flexibility

Membership fees provide a consistent revenue stream, contributing to the financial stability of the resource. This stability allows for better long-term planning and reduces dependence on unpredictable grant funding. Furthermore, the model offers flexibility; membership tiers can be adjusted to accommodate different user capabilities and needs, making the resource accessible while still generating income.

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

4.1.4 Ethical Considerations and Accessibility

It's important to balance revenue generation with the ethical imperative of ensuring wide accessibility. The resource should remain freely accessible at a basic level, especially for individuals and institutions with limited funding. A voluntary membership scheme should enhance the resource's offerings without creating barriers to essential data and services.

In conclusion, voluntary membership schemes offer a viable path to financial sustainability for biological research data resources. By providing tangible benefits to paying members, these schemes can generate a steady income stream while fostering a more engaged user community. This approach ensures the resource's longevity and continuous evolution in response to the needs of its diverse user base.

4.2 Subscription models

Subscription models present a viable approach for ensuring the financial sustainability of biological research data resources. This strategy involves charging users a regular fee for access to the database, which can be structured in various tiers to accommodate different needs and budgets.

4.2.1 Tiered Access and Flexibility

A key feature of subscription models is the tiered access structure. Basic access might be free or low-cost, providing essential data and functionalities. Premium tiers could offer advanced features such as more comprehensive datasets, higher data processing capabilities, personalized support, and exclusive research tools. This tiered system ensures basic accessibility for all users while providing enhanced services for those who can afford to pay more.

4.2.2 Predictable Revenue Stream

Subscription fees create a predictable and steady revenue stream, a significant advantage over traditional grant funding, which is often subject to uncertainties and time constraints. This financial predictability allows for better planning, development, and maintenance of the database, ensuring its long-term viability and continuous improvement.

4.2.3 Alignment with User Needs

To be successful, subscription models must align closely with user needs. Regular feedback mechanisms and user engagement surveys can help tailor the service to the evolving requirements of the user base. Ensuring that the subscription fees are justified by the value provided is crucial for user retention and satisfaction.

4.2.4 Community and Collaboration

Implementing a subscription model also encourages a sense of community and collaboration among users. Paying subscribers can be given platforms to suggest features, participate in beta testing of new tools, and engage in community forums. This collaborative environment not only improves the resources but also fosters a loyal user base.

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

4.2.5 Balancing Accessibility and Revenue

While a subscription model can provide financial stability, it's essential to balance revenue generation with the need for wide accessibility, especially for students, researchers from low-income regions, and non-profit organizations. Offering discounted or sponsored subscriptions to these groups can help maintain inclusivity.

In summary, subscription models offer a practical solution for the financial sustainability of biological research data resources. By providing tiered access, aligning services with user needs, and ensuring broad accessibility, such models can secure a stable revenue stream while continually enhancing the value of the resource for its diverse user base.

4.3 Implementing fees for data deposit

Implementing fees for data deposit is a strategic approach to developing a financial sustainability model for biological research data resources. This model involves charging researchers or institutions a fee to deposit their data into the database. This strategy can be particularly effective if the database offers value-added services that enhance the visibility, accessibility, and usability of the deposited data.

4.3.1 Rationale and Implementation

The rationale behind data deposit fees lies in the recognition that storing, curating, and maintaining data incurs costs. By charging for data deposits, the resource can cover these operational expenses. The fee structure can vary depending on factors like the size of the dataset, the complexity of the data, or the level of curation and maintenance required. It’s crucial to establish a pricing model that is transparent and justifiable, reflecting the value provided to the depositors.

4.3.2 Value-Added Services

To justify data deposit fees, the resource should offer services that add significant value. These could include high-quality data curation, ensuring data integrity and consistency, metadata generation for easier data discovery, and robust data preservation and archiving services. Additional services could include data analytics tools, visualization capabilities, and integration with other databases or platforms.

4.3.3 Balancing Costs and Accessibility

While fees for data deposit can generate essential revenue, it’s important to balance this with the need for open access to data. Excessive fees might discourage researchers, particularly those from smaller institutions or developing countries, from using the database. Offering tiered pricing, waivers, or discounts for certain groups can help maintain accessibility and inclusivity.

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

4.3.4 Incentives for Depositors

To encourage data deposits despite the fees, the resource could offer incentives such as enhanced data visibility, citation tracking, and acknowledgment in publications. These incentives not only benefit the depositors but also enrich the database, making it a more comprehensive and valuable resource for the entire research community.

In summary, fees for data deposit represent a potential revenue stream for sustaining a biological research data resource. By offering value-added services and balancing the cost implications with incentives and accessibility measures, this approach can contribute to the financial health of the resource while ensuring its ongoing relevance and utility to the research community.

4.4 Corporate backing

Corporate backing can play a pivotal role in developing a financial sustainability model for a biological research data resource. The collaboration between scientific data resources and corporate entities offers a unique opportunity to achieve long-term viability and growth.

Firstly, corporate backing can provide much-needed financial resources for the development and maintenance of the data resource. Biological research data resources require substantial investments in infrastructure, data collection, curation, and dissemination. These expenses can be daunting for non-profit or academic organizations to bear alone. Corporate partnerships can inject capital into the project, allowing it to scale and improve its services.

Furthermore, corporate sponsors can bring valuable expertise to the table. Many corporations have deep experience in data management, technology, and business strategy, which can be leveraged to enhance the efficiency and effectiveness of the research data resource. Their insights can help streamline operations, develop innovative tools, and optimize data management processes.

In addition to financial support and expertise, corporate partnerships can also broaden the reach and impact of the biological research data resource. Corporations often have extensive networks and customer bases, which can help disseminate the resource's data and findings to a wider audience. This increased visibility can attract more users and collaborators, further enhancing the resource's sustainability.

However, it is crucial to approach corporate backing with caution and a clear set of principles. Maintaining data integrity, ensuring open access to research findings, and protecting against conflicts of interest are paramount concerns when engaging with corporate partners. A well-defined governance structure and transparent guidelines should be established to safeguard the resource's integrity and mission.

Phoenix BIOINFORMATICS
39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

In conclusion, corporate backing can be a valuable component of a financial sustainability model for a biological research data resource. By providing financial resources, expertise, and expanded reach, corporate partnerships can help ensure the resource's long-term viability and its ability to support scientific advancements in the field of biology. Careful planning and ethical considerations are essential to forge mutually beneficial collaborations that prioritize the interests of both the research community and corporate partners.

4.5 Philanthropic contributions

Philanthropic contributions are a crucial component of developing a financial sustainability model for a biological research data resource. These charitable donations, often provided by foundations, individuals, or nonprofit organizations, can have a profound impact on the resource's ability to thrive and fulfill its mission.

4.5.1 Seed Funding and Start-Up Support

Philanthropic contributions can provide the initial funding needed to establish the biological research data resource. This seed funding can cover essential costs such as infrastructure setup, data collection, and hiring key personnel. It serves as a catalyst for the resource's development, allowing it to get off the ground and start making a difference in the scientific community.

4.5.2 Long-Term Sustainability

While grants and donations from philanthropic sources are typically time-limited, they can offer a bridge to long-term sustainability. By demonstrating the resource's value and impact during the initial funding period, it becomes more attractive to other potential stakeholders, including government agencies, corporate partners, and academic institutions, who may provide ongoing support.

4.5.3 Independence and Integrity

Philanthropic contributions often come with fewer strings attached compared to other funding sources. This financial independence allows the resource to maintain its integrity and focus on its mission without compromising its objectives due to commercial interests or profit-driven motives.

4.5.4 Support for Research and Innovation

Philanthropic donors may prioritize specific areas of research or technology development that align with their interests or values. This targeted support can enable the resource to invest in cutting-edge technologies, data analysis tools, or research projects that advance the field of biology, making it even more valuable to the scientific community.

4.5.5 Community Engagement

Philanthropic contributions can be used to engage with the broader community of scientists, educators, and stakeholders. This may include organizing conferences, workshops, or

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

educational programs to foster collaboration and knowledge sharing, further enhancing the resource's impact.

4.5.6 Diversification of Funding Streams

Relying solely on philanthropy may not be sustainable in the long term. However, these contributions can help diversify the resource's funding streams. This diversity reduces the risk associated with a single funding source and makes the resource more resilient to financial challenges.

In conclusion, philanthropic contributions are instrumental in kick starting and sustaining a biological research data resource. They provide the necessary initial capital, maintain the resource's independence and integrity, and pave the way for attracting a broader spectrum of stakeholders. When combined with other funding sources and a well-planned financial sustainability model, philanthropy can significantly contribute to the resource's long-term success and its ability to serve the scientific community effectively.

4.6 'Freemium' options

Implementing a 'freemium' model is a strategic approach to achieving financial sustainability for a biological research data resource. This model combines free access with premium, paid services, catering to a broader user base while generating revenue to support the resource's operations and growth.

4.6.1 Free Access

Offering free access to essential data and basic features is a key aspect of the freemium model. This approach ensures inclusivity by allowing researchers with limited budgets or those exploring the resource for the first time to benefit from its offerings. Free access can attract a large user base, raising the resource's profile within the scientific community.

4.6.2 Premium Features

To monetize the resource, premium features and services can be introduced. These could include advanced data analysis tools, more extensive datasets, faster access, or personalized support. Researchers or organizations willing to pay for these premium features can do so, providing a source of revenue to sustain the resource.

4.6.3 Tiered Pricing

Implementing tiered pricing for premium services allows flexibility in catering to diverse user needs. Different subscription levels can be offered, ranging from individual researchers to larger institutions or corporations. This tiered approach ensures that users pay according to their specific requirements and budget, maximizing revenue potential.

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

4.6.4 Data Licensing and Commercial Partnerships

Under the freemium model, commercial partnerships can also be established. Corporations or organizations interested in accessing and utilizing the resource's data for commercial purposes can enter into licensing agreements. This can be a significant source of revenue, especially if the data resource contains valuable proprietary information.

4.6.5 Community Building

Beyond revenue generation, the freemium model encourages community building. Users who benefit from free access often become advocates for the resource, helping to expand its user base. This organic growth can lead to more premium subscribers, further strengthening the resource's financial position.

4.6.6 Monitoring and Adaptation

It's essential to continuously monitor user behavior and feedback to refine the freemium model. Adjust pricing, features, and offerings based on user preferences and evolving research needs. This agility ensures the resource remains relevant and competitive in the dynamic field of biological research.

4.6.7 Transparency and Trust

Maintaining transparency about data usage, pricing, and the value proposition of premium features is crucial. Building trust with users by demonstrating the value they receive from both free and premium services is essential for long-term sustainability.

In conclusion, the freemium model is a versatile approach to developing a financial sustainability strategy for a biological research data resource. By balancing free access with premium services, the resource can attract a wide user base, generate revenue, and promote community engagement. When executed effectively, this model can secure the resource's financial stability while advancing its mission of supporting scientific research and discovery.

4.7 Crowdfunding

Crowdfunding is an innovative approach to fostering financial sustainability for a biological research data resource. This model leverages the collective support of individuals, organizations, and the public to raise funds for specific projects, initiatives, or ongoing operations of the resource. Here's how crowdfunding can be integrated into the sustainability plan:

4.7.1 Diverse Funding Sources

Crowdfunding allows the biological research data resource to tap into a wide range of potential donors, including scientists, educators, interested citizens, and philanthropic organizations. This diversity of funding sources reduces dependence on a single revenue stream, enhancing financial stability.

Phoenix BIOINFORMATICS
39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

4.7.2 Project-Specific Campaigns

Crowdfunding can be particularly effective for financing specific projects or initiatives within the resource, such as the creation of new datasets, the development of user-friendly tools, or the organization of scientific conferences. Each project can have its crowdfunding campaign, attracting donors passionate about that particular aspect of the resource's mission.

4.7.3 Engagement and Awareness

Crowdfunding campaigns provide an opportunity to engage the scientific community and the public in the resource's activities. These campaigns can raise awareness about the importance of biological research, data sharing, and the resource itself, potentially fostering a sense of ownership and support among donors.

4.7.4 Feedback and Collaboration

Crowdfunding campaigns offer a platform for direct interaction with donors. Through updates, surveys, and feedback mechanisms, the resource can gain valuable insights into donor preferences, needs, and expectations. This information can inform future development and sustainability efforts.

4.7.5 Seed Funding and Validation

Successful crowdfunding campaigns can serve as a form of seed funding. They can validate the resource's value and attract additional funding from other sources, such as government grants, corporate partnerships, or philanthropic foundations. Crowdfunding can kickstart larger initiatives by demonstrating community support.

4.7.6 Tiered Rewards

To incentivize contributions, crowdfunding campaigns often offer rewards or incentives to donors. These can range from access to premium features, exclusive content, or recognition in research publications. These rewards can be tailored to align with the resource's mission and the interests of donors.

4.7.7 Transparency and Accountability

Transparency is key to building trust with donors. Crowdfunding campaigns must clearly communicate how funds will be used and provide regular updates on progress. This accountability reinforces donor confidence in the resource's financial stewardship.

In conclusion, crowdfunding is a dynamic and engaging method to support the financial sustainability of a biological research data resource. By harnessing the collective power of supporters, crowdfunding can provide vital funding, raise awareness, and foster a sense of community involvement. When integrated into a comprehensive financial sustainability model, crowdfunding can help ensure the resource's continued growth and contribution to the scientific community.

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

5. Potential Revenue Streams

Based on the data available to us from a subset of AgBioData resources, we have come to these conclusions.

5.1 Pay for access model(academic)

This model, also known as an academic subscription model, requires researchers to pay a fee for access to the database. Some examples where this model has succeeded include ICPSR (social sciences data repository) and TAIR (plant genomic data). If the data to be accessed is valuable and unique researchers have a strong incentive to pay the fee. Disadvantages include the loss of access to those who can't afford to pay and the effort and expense of developing linked payment and data access mechanisms for individual researchers and/or negotiating licenses with research institutions for access for their researchers. One way to assess the feasibility of this support model is based on the amount of website traffic a database receives, as a proxy for how many researchers value the content.

5.2 Pay for access model (companies)

This model brings in revenue by charging for access to curated datasets with commercial value. The datasets of value would need to be gleaned out of public data either through curation or analysis. The specialized datasets would be made public after a moratorium period of one or more years. The advantage of this model is that the unanalyzed data would remain freely open to all. Disadvantages include potentially considerable effort and cost to develop the datasets of value, likely in advance of knowing whether the resulting revenue would cover those costs. Given that companies are capable of carrying out their own curation and analysis if the cost of such a dataset is too high, the excess revenue/profit margin that could be brought in would be limited, and for this reason would most likely require several customers for the same dataset in order to be worth the effort of creating it. Additional costs would be incurred in negotiating agreements for use of the data with companies, and possibly also support for the corporate clients.

5.3 Additional recommendations

The scope of our work did not allow us to reach out to additional groups of stakeholders but this would be a valuable exercise for AgBioData resource staff. A suggestion we received in an interview was that it may be possible to identify datasets of commercial value within the AgBioData consortium and require commercial entities to pay for access to these datasets. We recommend that the AgBioData consortium reach out to researchers who are currently working with commercial groups to pay for access to specialized datasets to understand if this would be an effective model for other databases within AgBioData. As suggested earlier, the usage numbers are not sufficient enough to provide a comfortable cushion of funding without additional funding sources. Increasing curated datasets, developing easy-to-use curation tools and generating specialized curated datasets will help drive more traffic to the databases. Another option is to increase visibility for the AgBioData consortium as a whole, as opposed to the individual databases. Even though these databases serve different research communities, a

Phoenix BIOINFORMATICS
39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

bundled approach to pay-for-access might be worth looking into. In this way, the pay-for-access will include all AgBioData databases, serving a larger researcher pool.

6. Investigation of Cost reduction techniques

In the context of the AgBioData consortium, achieving financial sustainability for its biological research databases is a multifaceted challenge. Key strategies for cost reduction, while ensuring high-quality data and service delivery, revolve around collaboration, community engagement, and leveraging shared infrastructures. These approaches not only aim to reduce operational costs but also to enhance the overall sustainability and effectiveness of the consortium.

6.1 Collaboration in Curation and Software Development

Collaboration plays a pivotal role in cost reduction. By pooling resources and expertise, consortium members can collectively develop and maintain software tools and curation processes. This shared approach reduces duplication of efforts and allows members to benefit from economies of scale. Collaborative software development, where multiple institutions contribute to a common codebase, can lead to the creation of more robust and versatile tools, reducing the need for each member to develop individual solutions. Similarly, shared curation efforts ensure high-quality data while distributing the workload and associated costs.

6.2 Outsourcing Curation to the Community or Data Providers

Engaging the wider community in the curation process can be an effective strategy to manage costs. Encouraging data providers and users to contribute to data curation and annotation efforts not only distributes the workload but also fosters a sense of ownership and community. This crowdsourcing approach can enhance the accuracy and richness of the data, as contributors bring diverse expertise and insights. However, it's crucial to maintain quality control mechanisms to ensure the reliability of the curated data.

6.3 Utilizing Shared Infrastructures

Leveraging shared infrastructures is another key strategy for cost reduction. By using common platforms for data storage, processing, and distribution, the consortium can significantly lower its IT expenses. Shared infrastructures can offer economies of scale, reducing the cost per unit of data stored or processed. Moreover, such infrastructures can provide robustness and scalability, accommodating the growing needs of the consortium without necessitating proportional increases in investment.

In summary, the AgBioData consortium can achieve financial sustainability through strategic collaboration in software development and data curation, outsourcing curation to the community, and utilizing shared infrastructures. These strategies not only aim to reduce operational costs but also enhance the consortium's capacity to provide high-quality, sustainable data services to its users. By implementing these approaches, the consortium can ensure its long-term viability and continue to support the vital research needs of the biological community.

39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

7. Assessment of Policy or Procedural Revisions Pertaining to Grant Funding

Assessing the sustainability of biological research data resources requires a critical examination of existing policies and procedures, particularly in relation to grant funding. This assessment is imperative for ensuring the continued relevance, accessibility, and utility of these databases, which are crucial for the global scientific community.

7.1 Mechanisms to Capture User Interest

One of the essential aspects of maintaining a biological research database is understanding and capturing user interest. This can be achieved through various means such as user feedback surveys, analytics on database usage, and engagement metrics. By identifying the most frequently accessed data, types of queries, and user demographics, administrators can tailor the database to better serve the needs of its users. This user-centric approach not only enhances the database's usability but also strengthens its relevance in the scientific community.

7.2 Importance of the Database to Its Users and Stakeholders

Biological research databases are invaluable to users such as researchers, educators, and policymakers. They provide a repository of data that can be used for a variety of purposes including academic research, policy development, and education. For stakeholders, the database is a critical resource for advancing scientific understanding and facilitating collaboration across different fields. It's a tool for innovation and discovery, enabling new insights and breakthroughs in biological research.

7.3 Mechanism to Capture Citations

Implementing a mechanism to capture citations is a pivotal step in quantifying the impact of a database. Such a system not only provides tangible evidence of the database's utility in scientific research but also aids in securing future funding. Granting agencies and donors are more likely to invest in a resource that demonstrably contributes to scientific advancement. Citations serve as a metric of success and relevance, showcasing the database's role in fostering research and discovery.

7.4 Global Core Biodata Coalition

Being part of the Global Core Biodata Coalition (GCBC) can significantly enhance the sustainability of biological research data resources. The GCBC brings together various stakeholders, including funders, policymakers, and database providers, to establish best practices and standards for data resource sustainability. By aligning with GCBC, a database can benefit from shared expertise, increased visibility, and potentially, more stable funding. This coalition promotes the integration of databases into a global network, enhancing their utility and ensuring they remain a vital resource for the international scientific community.

Phoenix BIOINFORMATICS
39899 Balentine Drive, Suite 200
Newark, CA 94560, USA
Phone: 650-995-7502
Fax: 8778205814
www.phoenixbioinformatics.org
info@phoenixbioinformatics.org

In conclusion, revising policies and procedures for grant funding with a focus on user engagement, citation tracking, and participation in global coalitions like the GCBC is crucial for the sustainable operation of biological research databases. This approach ensures these vital resources continue to serve and advance the global scientific community.