

# Digital repositories: All hype and no substance?

Marinus Swanepoel  
University of Lethbridge

## Abstract:

The development of digital repositories has been a recent one, starting in late 2000 when the UK's University of Southampton released a software package called E-Prints. Since that time, the establishment of digital repositories has gained momentum. Factors such as the falling costs for online storage, the increase of broadband and gigabit networking technologies, as well as the development of metadata standards to describe repository content, have all contributed to their current popularity. Questions to be asked are: to what extent are digital repositories, as a method for communicating scientific and scholarly information, accepted or are they just hype?; how mature are institutional repositories as a technology?; and, to what extent are institutional repositories used by faculty and researchers?

The Gartner Hype Curve is a tool introduced by the Gartner Group in the 1990's to explain general phenomena of interest in new technologies. This framework plots the typical progression of a technology from its early introduction through its maturation to broad market acceptance. The first question is answered by an investigation into the status of digital repositories in the context of the Gartner Hype Curve. The second question is answered by plotting institutional repositories on the product life cycle (Sigmoid Curve) and the third question is answered by applying the Diffusion of Innovation Theory to institutional repositories.

## Keywords:

Digital repositories, scholarly communication, Gartner Hype Curve, Product Life Cycle, Diffusion of Innovation Theory

## 1 Introduction

### 1.1 Background

Education and research are dependant upon information, and the role of libraries in supplying the needed information is well established. Libraries have a specific responsibility to preserve information. In doing so, they need to ensure that information resources remain permanently accessible, in working condition and authentic.

However, because of developments in information and communication technology, the nature of information resources is changing fast. In fact at the moment, it is in flux with more information being migrated to digital format and more made available in digital format, because it is born digital. It is therefore not surprising that information in electronic format represents the prime and fastest growing collection in academic libraries today. These and other developments such as the open access movement have increased the interest in institutional digital repositories.

It should be remembered that many faculty and students have been posting their intellectual output on the World Wide Web for some time. This was usually done as part of personal pages or on departmental sites (11). With the establishment of digital repositories, this is no longer the only or preferred means of making scholarly output available. Creation of digital institutional repositories has meant organizing the random posting of these scholarly communications into well structured, secure and attractive virtual spaces. The concept of digital repositories has captured the imagination of leaders in the field who have a vision of the benefit to be derived from doing just that. The expected benefit of digital institutional repositories includes that the repositories will serve as meaningful indicators of the quantity and quality of the intellectual output of institutions and that they provide institutional visibility and prestige to the institutions. Institution repositories also serve as the basis for a new disaggregated model of scholarly publishing (11).

## 1.2 Research Question

The perceived benefits mentioned above, along with the financial advantage that is envisaged with an open access scholarly communication process, has created a hype, an excitement in not only the library and information services arena but also in post secondary institutions. Institutional repositories (IR's) are seen as the antithesis to the existing model of scholarly communication that tends to inhibit the availability of scholarly communications. This lack of affordable accessibility is exacerbated by continuous annual increases, in excess of the rate of inflation, of the price of scholarly journals. Also in this regard, the (inflated) expectation is that IR's will be a weapon to counter the lack of access for financial reasons.

The research question of this paper is to determine the level of maturity and acceptance of IR's; also whether IR's have fulfilled (some of) the expectations created around this concept or, if IR's have left those hoping for the advantages mentioned above disillusioned.

Specific questions to be asked are:

- a) to what extent are IR's, as a method for communicating scientific and scholarly information, accepted or are they just hype?
- b) how mature are IR's as a technology?
- c) to what extent are IR's used by faculty and researchers?

## 1.3 Methodology

The first question will be answered by an investigation into the status of digital repositories in the context of the Gartner Hype Curve. The second question will be

answered by plotting institutional repositories on the product life cycle (Sigmoid Curve) and the third question will be answered by applying the Diffusion of Innovation Theory to IR's. The Product Life Cycle is a tool used to show the various stages of a product or service from introduction, to growth, to maturity, to decline; while Diffusion of Innovation Theory shows the extent to which a technology has been adopted by users.

Using all three models as indicators, applied to IR's, the primary research question should be answered with an acceptable degree of accuracy.

#### 1.4 Delimitations

IR's are part and parcel of the Open Access movement. However for the purposes of this study, this aspect of IR's will not be discussed; nor will the scholarly communication process be part of the investigation, even though IR's are part of the new emerging model of scholarly communication.

Disciplinary repositories are also outside the scope of this study. It will be limited to institutional repositories.

The focus of the paper is on determining the extent to which IR's have become mainstream technology.

## 2 Institutional Digital Repositories

### 2.1 Definitions

Clifford Lynch (12) defines IR's as: "a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members."

Barton and Walker (2) describes DSpace at MIT as a digital repository designed to manage, host, preserve and enable distribution of the scholarly output of MIT's faculty.

It seems that there is a difference of opinion concerning whether repositories should be restricted to scholarly output such as peer reviewed articles; or whether repositories should be expanded to include other material like grant proposals, radio/TV interviews, technical reports, photographs, etc. The supporters of the "pure" scholarly content of IR's seem also to be viewing IR's as a substitute for scholarly journals.

At the other end of the spectrum, there are proponents of the idea that IR's should be the "hold it all" of scholarly, academic and institutional output (3).

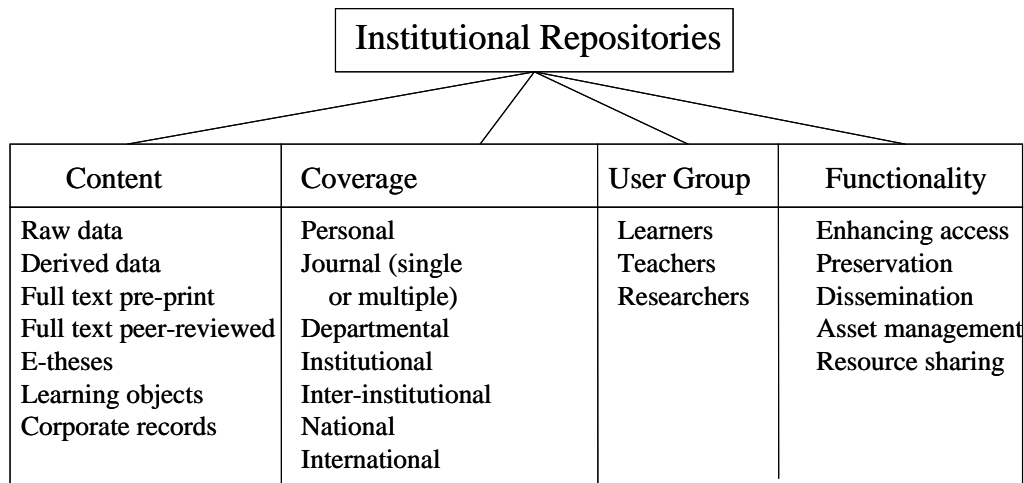
### 2.2 Characteristics

IR's are not just collections of data but offer functionality and features that distinguish them from other digital collections (16) (9):

- Repositories must provide permanent storage for scholarship. In this respect, concerns are expressed that post secondary institutions with IR's are paying too little attention to this aspect.
- IR's offer a set of basic services such as access control and search capability.
- IR software manages metadata and content.
- IR's are sustainable and trusted to continue to provide permanent storage.
- IR's are generally designed to grow in content through self-archiving by the creator of the item, though this is sometimes done on a departmental level (21).
- IR's provide open access to content
- IR's provide open access to metadata for harvesting.

### 2.3 Typology

IR's could be viewed as a recent development if the release of a software package called E-Prints in late 2000 by the University of Southampton, in the UK, is taken as the trigger that started it (19). They could also be viewed as being quite old if collections of electronic theses and dissertations are included in the definition of IR's. Probably the most well known repository of theses and dissertations is the Networked Digital Library of Theses and Dissertations started in 1996 at Virginia Tech (14).



**Figure 1: Typology of IR's**

The question of what should be included in IR's and what should not be included is as difficult to answer as it is to arrive at a generally acceptable definition of IR's, as shown in paragraph 2.1. However, Heery (9) developed a typology that provides a helpful framework for exploring IR's, as presented in Figure 1.

It is clear from this typology that many computations are possible. Figure 1 also illustrates that it will be difficult to achieve consensus on any single generally accepted definition.

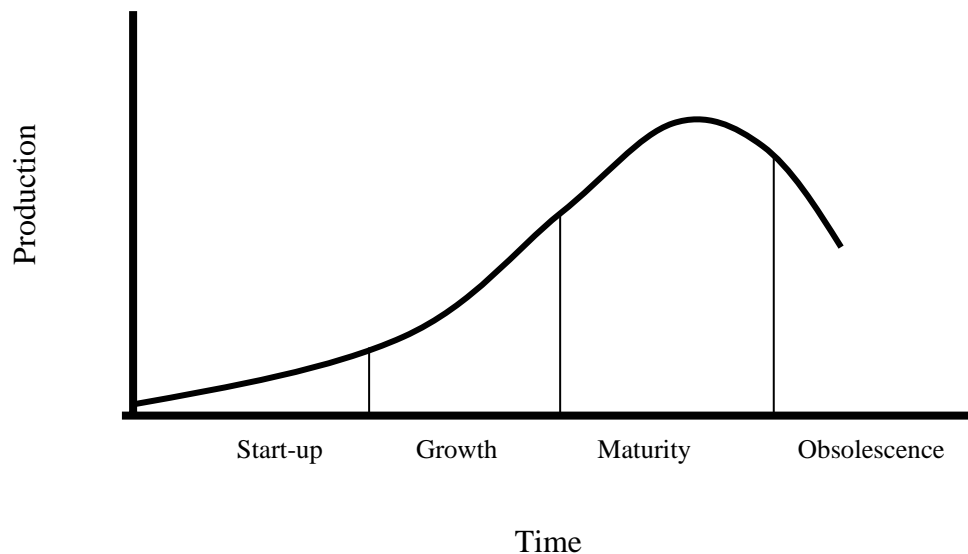
### 3 Technology models

#### 3.1 Product Life Cycle (Sigmoid Curve)

The Product Life Cycle or PLC provides a way to gauge the status of a product or service, from introduction to decline.

The four stages of a PLC can be described as follows (7):

- a) Introduction or start-up is the period from a new product's commercialization until takeoff.
- b) Growth is the period from a new product's takeoff until a slowdown in the growth of sales is experienced.
- c) Maturity is the period from a product's slowdown until sales begin a steady decline.
- d) Decline is the period of steadily decreasing sales until a product's demise or obsolescence.



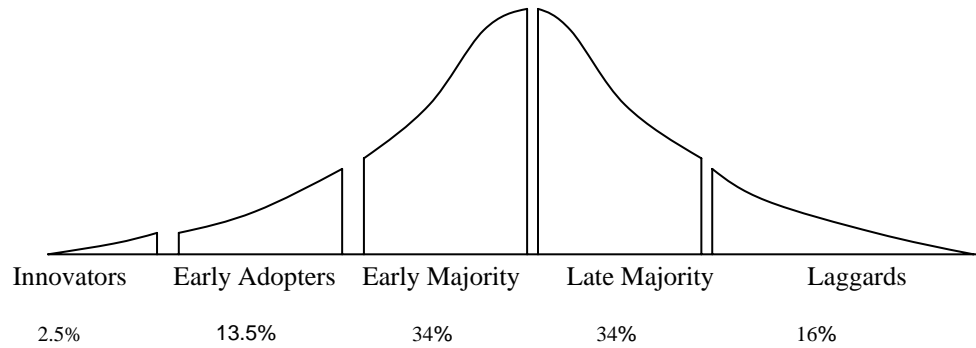
**Figure 2: Stages of the Product Life Cycle (15)**

#### 3.2 Diffusion of Innovation Theory (4)

Figure 3 shows the bell-shaped distribution of individual innovativeness and the percentage of potential adopters theorized to fall into each category. At one end of the distribution curve are the innovators. The risk takers and pioneers are innovators who adopt an innovation very early in the diffusion process. On the other extreme of the continuum are the laggards who oppose adopting an innovation until relatively late in the diffusion process, if they do so at all.

Earlier research on behavior in general, and technology adoption in particular, has indicated that there could be a common set of determinants of behavior among

different segments of the population, e.g., Davis, et al (5). This basic assumption is shared by the Innovation Diffusion Theory. It also assumes that everyone will ultimately adopt (17). These two assumptions are challenged. For example Moore (13) suggests that people in each of the different adopter categories are different from those in the category to their immediate left. These differences across categories are referred to as "cracks in the bell curve" (14).



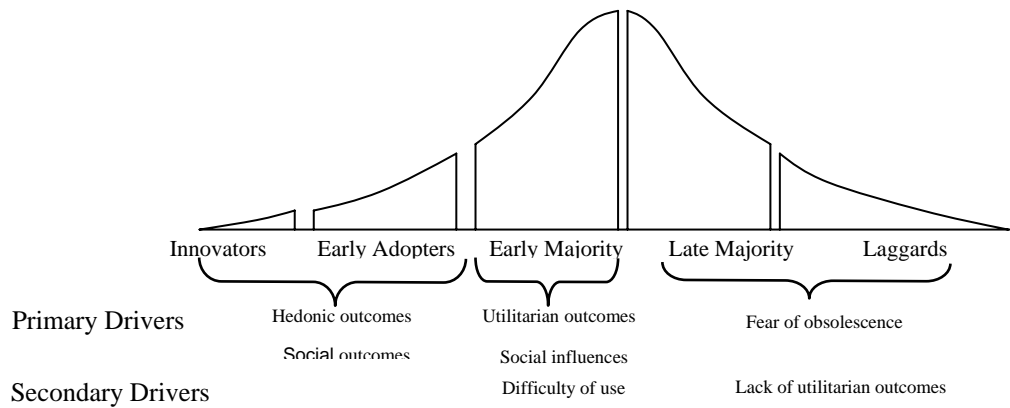
**Figure 3: Adopter Categories (24)**

This view suggests that innovations that succeed among innovators and/or early adopters may fail among the early majority or late majority since the innovation does not possess the characteristics that appeal to those in these later categories. Further, this implies that factors influencing different categories of adopters are fundamentally different. It also implies that since not everyone will adopt an innovation, it is quite likely that factors influencing non-adoption will be different from factors influencing adoption (17).

The motivating forces or drivers that result in adoption of the technology by the different adopter categories are known and presented in Figure 4.

Technology adoption decisions have been typically characterized by a strong productivity or utilitarian orientation. Across the different categories, drivers related to the use-productivity possibility (e.g., perceived usefulness, relative benefit, job fit, etcetera.) have emerged as the strongest predictors of adoption (17).

The role of utilitarian and hedonic outcomes is also supported by motivation theory. Motivation research suggests that there are two main classes of motivation: extrinsic and intrinsic. Extrinsic motivation pertains to achievement of a particular goal whereas intrinsic motivation is the enjoyment and pleasure resulting from a particular behavior (23).



**Figure 4: Mapping Adopter Determinants to Adopter Categories of Rogers (17)**

Social outcomes can be thought of as the social rewards (e.g., "public" acknowledgment) that would be achieved as an outcome of adopting an innovation. Adoption may lead to an elevation in power, knowledge, and/or status if the decision is viewed by others to be a good one (23). Prior research has emphasized the importance of social outcomes as a determinant of behavior. Similarly, innovation literature suggests that the desire to gain status is an important reason for the adoption of an innovation (17).

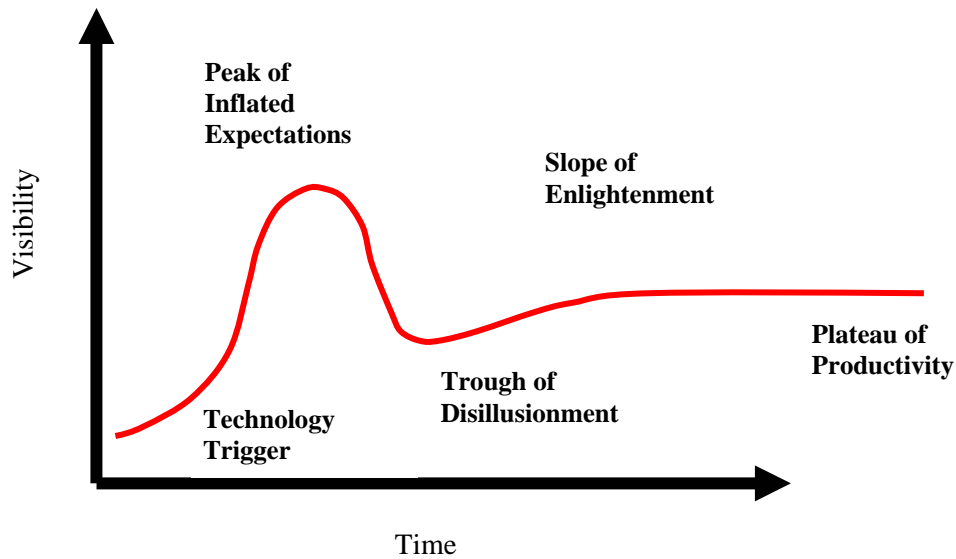
### 3.3 Gartner Hype Cycle

The Gartner Hype Curve is a tool introduced by the Gartner Group in the 1990's to explain general phenomena of interest and excitement (also called hype) in new technologies. This framework plots the typical evolution of a technology from its early beginning through its maturation to general market acceptance.

The different stages are (6) (22):

- **Technology Trigger:** The first phase of a Hype Cycle is the "technology trigger" or breakthrough, product launch or other event that generates significant press and interest.
- **Peak of Inflated Expectations:** In the next phase, a whirl of publicity normally generates over-enthusiasm and idealistic expectations. There may be some successful applications of a technology, but there are typically more failures.
- **Trough of Disillusionment:** Technologies enter the "trough of disillusionment" because they fail to meet expectations and rapidly become unfashionable. Consequently, the press usually abandons the topic and the technology.
- **Slope of Enlightenment:** Although the press may have stopped covering the technology, some businesses persist past the "trough of Disillusionment" and climb the "slope of enlightenment". Gradually, through continued experimentation they recognize the benefits and practical application of the technology.

- **Plateau of Productivity:** During this phase of a Hype Cycle, the real benefits of the technology are established and accepted. Tools and methodologies are more and more stable as they enter their subsequent generations. The final height of the plateau varies according to whether the technology is generally applicable or benefits only a niche market. Approximately 30 percent of the technology's target audience have or are adopting the technology as it enters the Plateau.



**Figure 5: Gartner Hype Cycle (6)**

#### 4 The Models Applied

In deciding where IR's should be positioned on the curves presented by the three models, the following represent some of evidence to be considered:

On November 18-19, 2004, SPARC and SPARC Europe presented a workshop titled: "Institutional Repositories: The Next Stage" (18). The 276 registrants from 184 institutions discussed issues such as:

- how to populate IR's
- how to manage policy issues of IR's
- legal issues: copyright, authors' agreements, etc.
- IR business models
- understanding digital preservation
- IR technical solutions
- what other institutions are doing with their IR's.

Looking towards the future, Ober (from a 2004 perspective), as reported by Cantara (3), sees the need to build enough repositories with value-added services that would continuously realign goals to suit requirements of the environment and also make



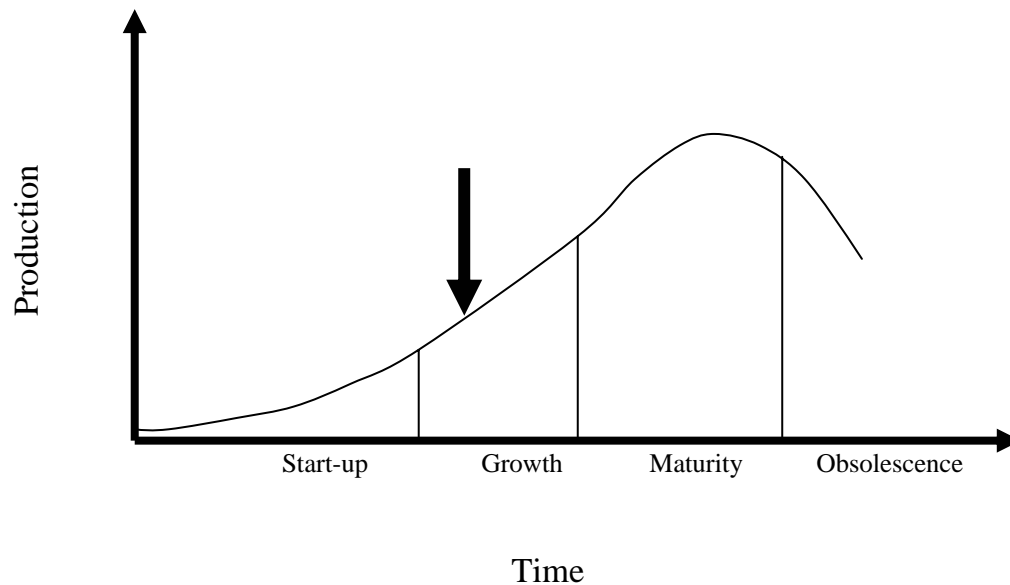
available core services such as metadata enhancement and information lifecycle managers.

In the November 07, 2003 issue of the *Chronicle of Higher Education* Atkinson encourages universities to establish IR's: "...giving faculty members the necessary tools to make their publications more accessible. Universities should shoulder the costs of developing, managing, and publicizing research -- including peer review of scholarly papers -- and build the online capacity to distribute those works worldwide. The costs, though not insignificant, pale in comparison to those that libraries must bear to buy access to our faculty members' publications." (1)

The *Timeline of the Open Access Movement* compiled by Suber (20) also provides useful information on the development of repositories, providing a sense of the current status of repositories as a technology.

#### 4.1 Product Life Cycle (Sigmoid Curve)

Taking into account the situation around IR's as described in the first part of par. 4 it is clear that IR's are no longer a start-up technology. It is equally clear that they do not represent a mature technology. Therefore IR's must be in the growth phase of the Product Life Cycle. The estimate is that they have made modest but significant progress in this phase but that the degree of progress and adoption differs from country to country and even from institution to institution.



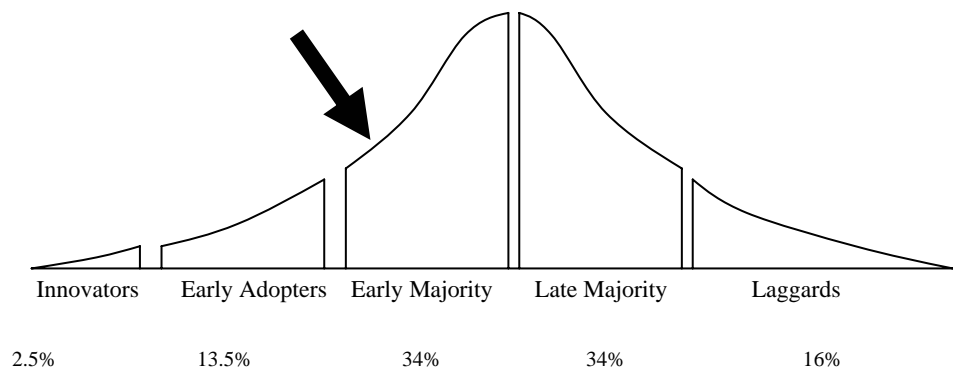
**Figure 6: Estimated position of IR's on the Product Life Cycle**

#### 4.2 Diffusion of Innovation Theory

Even though it cannot be assumed that everybody will display a similar attitude towards a new technology, varying only in the degree or speed of acceptance, the

Theory of Innovation Diffusion does provide an indication of the status of IR's in general terms.

The diffusion of IR's among faculty can be gauged to some extent by statistics reported at the Berlin 3 meeting on Open Access held February 28 – March 01, 2005 in Southampton (8). At this meeting it was reported that although faculty are still ill informed about Open Access, 25% were already providing open access to their scholarly output. It is acknowledged that the use of open access does not necessarily equate to archiving in IR's. However it is argued that faculty making use of open access journals are also more likely to use IR's. Furthermore, 79% of faculty indicated their willingness to self-archive, provided their institutions require them to do so (8).



**Figure 7: IR's have moved into the early majority**

This could serve as evidence to argue that IR's are no longer receiving support only from the Innovators and Early Adopters but also from members of the Early Majority.

The drivers that motivate the various adopter categories (mentioned in paragraph 3.2) are seen to be applicable to faculty in adopting or not adopting the use of IR's.

#### 4.3 Gartner Hype Cycle

Gartner indicates that IR's (called E-Learning Repositories in Figure 8) are moving down into the Trough of Disillusionment. Looking at the issues under discussion, as indicated in the first part of par. 4, it is clear that these are issues of a developing technology. However the number of successful repositories to be seen in the *Institutional Archives Registry* (Figure 9) indicates that this technology may have moved further along the curve than it seems at first glance (10). If the number of faculty participating at some individual institutions is taken into account, it suggests that this technology of IR's has moved further down the slope closer to the plateau. Some institutions are reporting a 25% participation rate (8) while 30% penetration (22) is seen as a technology already entering the plateau.

On the other hand, if the flood of literature on IR's that is still being published in professional literature is taken into account, it argues that the rightful position of IR's is still not far from the Peak of Inflated Expectation.



Figure 8: Hype Cycle for Higher Education Technology (6)

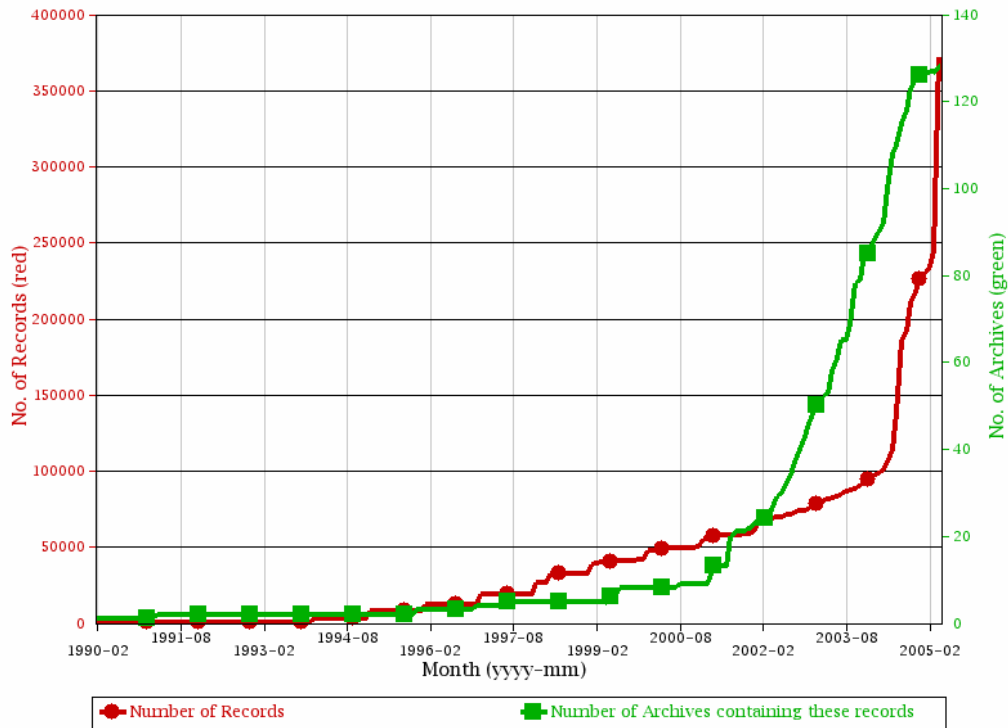


Figure 9: Growth in number of records and number of institutional archives (10)

## 5 Conclusion

All three models seem to indicate that IR's are well on their way to becoming mainstream technology. In terms of the Product Life Cycle the technology of IR's is gathering critical mass and should move to maturity within the next 5 years. This is confirmed by the Theory of Innovation Diffusion where IR's have clearly achieved significant acceptance by the early majority. The Gartner Hype Cycle indicate that the IR technology is on its way to the Trough of Disillusionment but will pass the Slope of Enlightenment reasonably soon to reach the Plateau of Productivity within five to ten years. Taking cognizance of the other two indicators it should be closer to five years rather than ten. This is supported by Suber's view of the momentum building in the Open Access movement, to which IR's are inextricably linked (20).

## References

1. Atkinson, R. C. (2003). A New World of Scholarly Communication. In *Chronicle of Higher Education*. November 07, 2003. <http://chronicle.com/free/v50/i11/11b01601.htm> (Accessed April 28, 2005).
2. Barton, M. R. & Walker, J. H. (2003). Building a Business Plan for DSpace, MIT Libraries' Digital Institutional Repository. In *Journal of Digital Information*. 4(2) Article No. 173, 2003-05-07. <http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Barton/barton-final.pdf> (Accessed April 22, 2005).
3. Cantara, L. (2004). Conference report: Institutional Repositories: The Next Stage. SPARC/SPARC Europe Workshop, Washington, DC, 18-19 November 2004 [http://library.case.edu/ksl/ecoll/digitalcase/sparc\\_ir\\_workshop.pdf](http://library.case.edu/ksl/ecoll/digitalcase/sparc_ir_workshop.pdf) (Accessed April 29, 2005).
4. Clark, R. (1999). A Primer in Diffusion of Innovations Theory. <http://www.anu.edu.au/people/Roger.Clarke/SOS/InnDiff.html> (Accessed April 26, 2005).
5. Davis, F. D. et al (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. In *Management Science*. 35, 982-1003.
6. Gartner Group (200?). Understanding Hype Cycles. <http://www.gartner.com/hc/asset50595.jsp> (Accessed April 22, 2005).

7. Golder, P. N. & Tellis, G. J. (2004). Growing, Growing, Gone: Cascades, Diffusion, and Turning Points in the Product Life Cycle. In *Marketing Science*. 23 (2).
8. Harnard, S. (2005). The implementation of the Berlin Declaration on Open Access: Report on the Berlin 3 Meeting Held 28 February – 1 March 2005, Southampton, UK. In *D-Lib Magazine*. 11(3). <http://www.dlib.org/dlib/march05/harnad/03harnad.html> (Accessed April 22, 2005).
9. Heery, R. & Anderson, S. (2005). Digital Repositories Review. <http://www.jisc.ac.uk/uploadeddocuments/rep-review-final-20050220.pdf> (Accessed April 20, 2005).
10. *Institutional Archives Registry*. (2005). <http://archives.eprints.org/> (Accessed April 25, 2005).
11. Johnson, R.K. (2002). Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication. In *D-Lib Magazine*. 8(11) <http://www.dlib.org/dlib/november02/johnson/11johnson.html> (Accessed April 29, 2005).
12. Lynch, Clifford. A. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. <http://www.arl.org/newsltr/226/ir.html> (Accessed April 22, 2005).
13. Moore, G. A. (1991). *Crossing the Chasm: Marketing and Selling Technology Products to Mainstream Customers*. Harper Collins, New York.
14. Peek, R. (2004). Are dissertations next? In *Information Today*. 21 (10) p.17-18.
15. Product Life Cycle. (2004). <http://www.witiger.com/marketing/PLC.htm> (Accessed April 22, 2005).
16. Quint, B. (2004). Librarians as Santa's Elve's. In *Information Today*. 21 (11) p.7.
17. Rogers, E. M. (1995). *Diffusion of Innovations*, 4th ed., The Free Press, New York.
18. SPARC Presents a Workshop on Institutional Repositories. (2004). ARL no. 237D2004. [http://library.case.edu/ksl/ecoll/digitalcase/sparc\\_ir\\_workshop.pdf](http://library.case.edu/ksl/ecoll/digitalcase/sparc_ir_workshop.pdf) (Accessed April 22, 2005).
19. Suber, P. (2004). Open Access Builds Momentum. *ARL Bimonthly Report* 232. <http://www.arl.org/newsltr/232/openaccess.html> (Accessed April 22, 2005).

20. Suber, P. (2005). Timeline of the Open Access Movement. <http://www.earlham.edu/~peters/fos/timeline.htm> (Accessed April 22, 2005).
21. Tennant, R. (2002). Institutional Repositories. In *Library Journal*. September 15, 2002. <http://www.libraryjournal.com/article/CA242297?display=search&Results&stt=001&text=institutional%2Brepositories> (Accessed April 22, 2005).
22. Twist, J. (2004). Peering beyond the technology Hype. <http://news.bbc.co.uk/1/hi/technology/3577746.stm> (Accessed April 25, 2005).
23. Vallerand, R. J. (1997). Towards a Hierarchical Model of Intrinsic and Extrinsic Motivation. In *Advances in Experimental Social Psychology*. 29, 271-360.
24. Venkatesh, V. and Brown, S. (1998). *A longitudinal investigation of personal computers in homes: Adoption determinants and emerging challenges*. CSI Working Paper # WP-98-01. <http://www.slis.indiana.edu/CSI/WP/wp98-01B.html> (Accessed April 25, 2005).