Samuel Gordon-Stewart: This is Samuel's Persiflage, Episode #5 for May 2006.

(Music)

Samuel Gordon-Stewart: Yes, hello and welcome to another episode of Samuel's Persiflage coming to you from the lovely city of Canberra. Well, today on the show we've got a fair bit coming up for you. We've got another Persiflage puzzle. I must say, there was an interesting reaction to the last Persiflage puzzle, and I'll go through that in the feedback, which comes up a little bit later on. There's also a bit of a change to how you can send feedback, just so that it makes just it little bit easier. Today though, we have, the major feature of the show is an interview with Margaret Phillips, the Director of Digital Archiving at the National Library of Australia. She'll be talking to us about number of things that the library does; including and this is something I am very excited about, the PANDORA Archive- which is their attempt to at Archiving the Australian Internet web space basically. So, we will go into that one; that's quite interesting. That will take up most of the show. We have also got some odd stories and as I mentioned, we do have the feedback and the Persiflage puzzle so, I think it's going to be a good show. I hope you can stick around for it. I don't know if saying that you should stay tuned is the right term to use in a podcast but we might use it anyway. I hope you can stay tuned to Samuel's Persiflage.

(Music)

SGS: But before we can do anything else, we better introduce the Persiflage puzzle and its an interesting one this month. There's few words in this one and I will give you a little clue- this one is probably one of the more well recognized music combinations of our history. I won't say much more than that except to say that, they're fairly influential, at least they were. I won't go any further than that but I'll let you think about that. Anyway, there are five words-- the first word has four letters, the second word has five letters, the third word has three letters, the fourth word also has three letters and the last word has six letters. And seeing as we're on our first go at the Persiflage puzzle for this month, I'll give

you five letters, lets start with the letter 'K' and there are none of those there, so we might try the letter 'W' and none of those either. Have a look, next one is the letter 'F' for Fred hmm… none of them so that's no Ks, no Ws, no Fs. Lets, pull out another letter and it's the letter 'N' for Neighbour and that's in the puzzle- it's the second letter of the third word, so the second letter of the third word is an 'N'. And the next letter out is the letter 'Z' for Zebra; none of them. So, that's a few letters and the letters that aren't in there, they might help you work it out as well because the letters that aren't in there help you eliminate some of the possibilities.

(Music)

SGS: The National Library of Australia is one of Australia's more important institutions but one that not a lot of people seem to understand properly. Joining me on the line to talk about the National Library is their Director of Digital Archiving, Margaret Phillips. Margaret, welcome to the program.

(00:05:04)

Margaret Phillips: Thank you, Samuel.

SGS: So, in a nutshell what is the role of the National Library?

MP: The role of the National Library is to collect and preserve Australia's documentary heritage, that's the primary role. We also collect material from overseas, particularly that material which gives context to the Australian collections. An important area of our overseas collecting is from South-East Asia and the Pacific, because of Australia's role and the geographic replacements in that area. That's our fundamental role and then of course, there are related activities such as providing researchers, readers, members of the public with access to those collections, so we have- of course information services. We have exhibitions and including some traveling exhibitions, which go around Australia so that Australians, wherever they are, can gain access to our collections and the information contained within them.

SGS: Right. I have met a few people who seem to get the impression that it's just a really big version of a suburban library but that's not true because you can't actually borrow the materials, can you?

MP: It's not true. Well, the materials can be borrowed but not in the same way as your local public library. The National Library is much more than a public library. We have other roles, which include setting up information infrastructure within Australia for all Australian Libraries to use; that includes the service that's recently been launched called, Libraries Australia.

SGS: Right.

MP: And that is a service, which enables people, wherever they are, at home, at schools, in their office- to search what is called, the National Bibliographic Database. The National Bibliographic Database is a collection of catalogue records or descriptions of library resources that, I think its somewhere like, 800 libraries around Australia have in their collections. So, all of the local public libraries, the University libraries, the special libraries such as libraries at hospitals, the state libraries- they all catalogue and describe the materials that they have in their collections and then they contribute it, they send those records to the National Bibliographic database and the National Library makes those available to everybody to search through this new service called Libraries Australia. Now, that assists people wherever they are, to locate a book or a serial or a picture, a manuscript, a map, that may help them with their research and it gives them information about how they can then obtain access to that information. And that includes services such as what librarians call 'document supply' so they can go to their local public library and request that library to borrow that book or obtain a copy of a journal article, for instance from another library, somewhere in the country that actually holds that item and that would include from the National Library. So, the National Library does lend its -- some of its collection materials to other libraries, upon request.

SGS: Okay.

MP: But no, you can't walk in to the National Library and just borrow a book over of the counter.

SGS: The National Library keeps on record, its one copy of each book published in Australia, isn't it?

MP: That's correct. That is enabled through the Legal Deposit Legislation which part of the Copyright Act of 1968, that Act stipulates that all publishers should send one copy of everything they publish to the National Library. And that enables the National Library to build a preservation collection of everything that's published in print in Australia.

(00:10:14)

Unfortunately, that legislation has not yet been extended to include other formats, such as CD ROMs, DVDs or web publications and that means that the Library does experience some difficulty in building up a comprehensive collection of those kinds of material but we are attempting to do that without the aid of legal deposit.

SGS: We'll talk a little bit more about your Internet archive project a little bit later on, but regular listeners would remember Gary Floyd- a School librarian from an earlier episode. I's chatting with him, probably a week or two ago, and he made an interesting point about Screensound Australia, collecting audio and video and the National Library doing that as well. He asked me if I could ask you, where you sort of draw the boundary between the two institutions?

MP: The National Library and Screensound Australia, which has now gone back its former name of, the National Film and Sound Archive.

SGS: Oh, has it? I can never keep up with that. They keep changing it.

MP: Yes, they have changed back to the former name; have complimentary roles. The collection at the National Film and Sound archive used to be part of the National Library until those two collections was separated out. But we're still in very close contact over our collection policies and we don't duplicate each other. We have defined collection policies.

The National Library's collection in this area is oral history recordings and also folklore recordings. Our recordings, by and large, are not commercial recordings, the recordings that we commission, that we engage either our staff or experts in the field to interview Australians of importance in some aspect of Australian life and we record those interviews and create transcripts of them and then add them to our collection. Like wise, we engage people with particular expertise to go out into various communities and to create folklore recordings.

The National Film and Sound Archive, on the other hand, is responsible for collecting the commercial audio and video recordings that are produced in Australia. So, there's a very distinct difference between the collecting tasks of the two agencies. Sometimes, we do have materials sent to us-- a video or a DVD, or a CD, something of that nature and if it contains primarily music or film, we will send it on to Screensound, but if it contains some text and documentary material and explanatory material sometimes, we do consider that it's more appropriate for our collection. So, sometimes there is a bit of a gray area, but mostly it's fairly clear, what the two organizations will collect.

SGS: All right. Going back to the books being added under the legal deposit part of the Copyright Act, what's the process for actually having a book added to your collection? The books obviously have to be sent, but do you get notified that a book is going to be published, before it's published or do they just turn up?

MP: Sometimes, we do have notifications. The Library also offers a service to publishers, which is called Cataloguing-in-Publication or CIP, as its known for short. And publishers can ring the CIP section of the library and ask for what is known as a -- it's a brief catalogue record based on the information that the publisher gives to the Library officer over the phone.

(00:15:03)

The Librarian then catalogues the item and supplies that record to the publisher, and the publisher then publishes that or inserts that catalogue record into the front of the book.

Now, that alerts us to the fact that that book is on the way but we don't always have that kind of advance notice and we often receive books, that, from publishers- particularly the

large amount of self publishing that takes place now, the desktop publishing, so that it's not an essential part of the process but it is a service that the library offers to publishers who want it.

The Library acquires books in a number of ways through legal depositors, I mentioned. But we also purchas copies of Australian books when the books come in on legal deposit. We look at them and evaluate whether they're the kinds of publications that the Library is likely to receive loan requests for, and if they are, then we purchase a second copy from the publisher. We also publish books from overseas; we do that most often through blanket orders. So, we have suppliers overseas to whom we give a selection profile, we tell them the sorts of books that we want to acquire, that the subjects that we give emphasis to, and the audience that we're interested in—so, mostly an adult research oriented audience, and then they send us books that meet that profile.

SGS: Right.

MP: We also receive books on donation. Sometimes, members of the public having their personal collections- books of value that they think the library might not hold and would be interested in and we also purchase retrospective Australiana, and so, in the past, for some reason, we may missed out on collecting published material and we strive to fill those gaps.

SGS: Right. I suppose newspapers come under similar legislation. I'm not aware of the exact legislation there but you keep copies of all of the newspapers or just some?

MP: You're right. Newspapers, being printed library material, are subject to the legal deposit legislation.

SGS: Right.

MP: And we do collect a comprehensive collection of Australian newspapers. The kinds of newspapers that we would not include would be very local newspapers; the kind of free newspapers that would be put on into your letterbox, for instance, for local areas. But the State Libraries are responsible for collecting those kinds of local interest newspapers so they are definitely collected within Australia.

The National Library administers what is called the 'Newspaper Plan' -- the 'Australian Newspaper Plan' or 'N Plan' for short, and the National and State Libraries, are on the Planning Committee to ensure that all Australian newspapers are collected and preserved in at least one library. But the National Library acquires a comprehensive collection of Australian newspapers and for many of them; we also collect them in a microfilm version as well.

SGS: Right. So, how did you decide which newspapers to put on to microfilm and which ones not to? I would imagine, based on how people come in to access the newspapers,

that perhaps the ones that are more likely to be accessed, would be put on microfilm to res -- not restore, to save the original version.

MP: That's exactly right. I mean, as you would be aware, newspaper is very easily torn, it's a low-grade quality and it degrades fairly quickly with use.

(00:20:11)

So, it's the high-use newspapers that we acquire the microfilm for. Usually, we don't create that microfilm ourselves; these are publishers, supply a microfilm version although we have to buy that. That doesn't come with legal deposits.

SGS: That's interesting.

MP: We do have a microfilming project with some of the members of this 'Newspaper Preservation Plan' and that is, microfilming the older newspapers, for which the commercial microfilm isn't available. We microfilm old the newspapers as a preservation measure, microfilm and also digitize so that they can be viewed online.

SGS: Oh, okay. Just going back to the local newspapers, just for a moment, would it be fair that as well as being the National Library you also fulfill the role of ACT State Library, because you collect 'The Chronicle' which is a very Canberra centric publication

MP: Yes.

SGS: And I suppose, seeing as the ACT is such a small area, you probably fulfill that role as well?

MP: Yes, I think, it's fair to say that we do. I mean, its unofficial, I think. The ACT has its own heritage library and it is collecting important ACT materials of heritage value. Certainly, when we select material here for the National Library's collection, we do take into account that the Library building is situated in the ACT and that the ACT Government Library System doesn't have quite the resources that we do. And so, we do give more preference to ACT publications including government publications and newspapers than they -- than we would, say, for the other states.

SGS: Okay, well, what we'll do now, we've just got a few other quick things to do, so we will come back to this in just a moment. But right now on Samuel's Persiflage.
(static sound)

Voiceover: Samuel's Persiflage Listener Feedback, send your e-mails to podcast@samuelgordonstewart.com….

(Music)

SGS: Yes, indeed its time for the Samuel's Persiflage Listener Feedback, and as you'd be well aware, you can send your feedback to podcast@samuelgordonstewart.com.

Now, you can send written feedback, after all that's what e-mail was intended for. Or if you want, you can send some audio feedback and I appreciate the audio feedback because, after all, this is an audio product, really. So, having that audio feedback just breaks a bit of the monotony of me talking a lot. Now, you can attach that to an e-mail quite easily, in mp3 format or wave format or Ogg Vorbis format. I'm happy to receive any of those, or there is an even easier way to do it now. I've added this for your convenience-- on the podcast website which is samuelgordonstewart.com/samuels–persiflage or just go to samuelgordonstewart.com and click on the Samuel's Persiflage link, you'll find there is a button to send audio feedback; if you click on that, that would take to another page, thanks to the people at odeo.com where you can send through some spoken feedback. It'll record the feedback for you; right there on the web page and it will send it for you. You can attach a short written message if you want, I will receive that. I can get that from the site in mp3 format and quite easily include it here on the podcast.

So, you've got all of these options open to you and, as much as, I love getting the written feedback it would just be great if I could also get some spoken and even musical feedback if you like because, as I said, it just helps to break the monotony of me speaking a lot.

(00:25:00)

Anyway, its on to the feedback and Pete is first this month. He says, "Samuel, I listened to this on the way to Sydney on Friday." He is talking about Episode #4, "It certainly made the trip shorter and I could only smile when you read my comments on the air. Thank you, Pete." Well, thank you very much, Pete. That's very nice of you and I am glad that I made your trip to Sydney much shorter.

Chuck sends us an e-mail, "Hi Samuel, how are you? I really enjoyed Samuel's Persiflage episode #4. I think, you are really getting the hang of it now. My favorite part was the interview with Mrs. Gordon Stewart. I enjoy listening to people's stories who I don't know. I found it fascinating. You seemed very relaxed interviewing your mum. Have you been to the Netherlands?" Well, Chuck, no. I haven't, although it would be nice to go sometime. Chuck goes on to say, "I sometimes listen to Radio Netherlands via News Radio on the ABC." That's interesting. I can't say, I've actually listened to it. I would assume that if its on News Radio, its in English—I'd hope it is because it would be a bit hard to understand for most people in Australia if it wasn't. Chuck then goes on to say, "Your mum's story was very interesting. Please pass on my regards to her and thank her for her interview," I've already done that, Chuck. So, thank you for that. She appreciated that and Chuck says, "Maybe Mr. Gordon Stewart could appear on the next Persiflage." Well, maybe. We'll see; he might appear some time in the future, maybe, we will see what happens. Chuck goes on- this a lengthy e-mail, "Echuca is on the Victoria side of the border." Chuck is referring to an incident in the last episode during the interview with Mrs. Gordon Stewart. We'll just play that now.

Mrs Gordon-Stewart: I helped with looking after her children because she had…

SGS: Yeah, sorry to interrupt. I just found Echuca, its on the Murray river; it's a bit north of the New south Wales- Victoria border but carry on.

Mrs. Gordon-Stewart: And apart from that.

SGS: Chuck goes on to write, "If you cross the bridge at Echuca over the Murray River you arrive at Moama. I used to go on holiday there." Well, Chuck, you are right. Echuca is in Victoria, see what happened is-- the map behind me in this studio- it shows railway lines and state lines and unfortunately, its not terribly clear and the lines look very-very similar. And as I, in the middle of the recording it, I sort of looked it the wrong line, Echuca was north of the railway line, which I realise now, but actually south of the state line which I didn't see. So, I mistakenly thought that Echuca was in New South Wales, which obviously it isn't. But not to worry, you're not the only one who told me that, Chuck. Plenty of people sent through e-mails about that, so thank you to everyone who highlighted that for me. And yes, these accidents will occur; it has to be said.

Now, people who were listening to Episode #3, would remember Danny O'Brien from the Electronic Frontier Foundation. Well, he sent through this e-mail, "Hey Samuel. Thanks so much for sending me the stats on the Persiflage appearance but don't worry, there is no need to air mail this to me every month. If you want to let me know feel free to just e-mail it to me. Best wishes, Danny." Well, thank you very much for that, Danny. I've replied to that. See, my basic policy with all of the people who appear on the show, is that I will send them letters or in Danny's case, now e-mails, outlining how many downloads there have been of their episode of Samuel's Persiflage each month. And I send these through, just after the beginning of the month once all the stats have been collated and calculated-- I print them out and I send them through with the total downloads so that they know how much it has been heard. Much like I do with the blog view stats on the samuelgordonstewart.com website. There the Persiflage stats are, indeed, included there and that's pretty much what I sent through to the people who appear on the show; the stats for their episode. And overall, the response has been very good. I actually received a few letters from people who've been on the show previously, thanking me for these letters and Danny was really no exception there. But yes, certainly printing them out and sending it by air mail is not only taking up some time but it also means I am using a lot of paper.

(00:30:01)

So, I am starting to send through some of them by E-mail mail now, which is certainly easier, but I think its important that I do send through these stats, because ultimately with any broadcast medium such as radio or television-- if you are on an interview there you can quite easily find out roughly how many people are listening. But with podcasts, there is no easy way to do that. So, what I do, I have decided that the best thing to do is to make sure that the people, who appear on the show, know roughly how many people

have listened to their interview. Obviously, download statistics aren't exact listener statistics due to the very nature of downloads, but it's a rough estimate that seems to work most of the time. So, that's basically the policy there.

Moving on and Jerry poses a very interesting question, "Samuel, why do you use podcast@samuelgordonstewart.com and not Persiflage@samuelgordonstewart.com. The second one would make more sense." Well, Jerry, before I started the show in December last year, I was thinking the exact same thing, I might as well just use Persiflage@samuelgordonstewart.com, but if you have a look through some of the sites that have actually linked to Samuel's Persiflage, it will become quite apparent that regardless of how often people see the word Persiflage, they have no idea how to spell it. Now, I really don't know how many people send e-mails to the samuelgordonstewart.com domain-- to addresses that don't actually exist there. But the only way I could really capture all of the Persiflage possible spellings is to either setup a few dozen e-mail accounts and have them all forward to the one, which would be a pain in the neck. And in fact, mean that I'd end up with more multiple bits of spam in the e-mails, or I could just have all of the false e-mails just forward into there and that would just be painful as well--as I would end up spending hours trawling through all of these pointless E-mails which would ultimately be caught by the spam filter for the most part. But I generally look through my spam, just a quick skim to see if anything legitimate has been picked up by the spam filter, usually isn't, but it would just mean that process would take much longer.

So, podcast is a word-- its two syllables, its very easy to spell and most people can spell it, and if they can't, there's a link on the Samuel's Persiflage website; not only to podcast@samuelgordonstewart.com, but also to the spoken feedback link, which is very helpful as well. So, if you'd like to send me some feedback, those two options are open to you-- podcast@samuelgordonstewart.com or you can send me some written-- spoken feedback rather by clicking on that link on the site.

(Music)

Well, I think, we might do just a quick update on the Persiflage puzzle before we go back to Margaret Phillips from the National Library of Australia. And we will draw out a few more letters for you. We'll start with the letter "J" for Jerry and there are none of them. We seemed to be having a run of letters that aren't in the puzzle. Okay, well, lets have a look at another letter--the letter "O." Yes, that's in there. It's the second letter in the fifth word, so the second letter in the last word indeed. All right, well, we'll go on and pull out another letter. How about the letter "C" for Cat? Well, that one's in there as well. That's the first letter of the last word. We'll have another letter now-- the letter "V" for Victor and it's not in the puzzle at all. So, keeping in mind that clue I gave you earlier you might be able to work this out, if you know your music.

(00:35:00)

But, I don't know this one; looks like it could be a bit tricky. We'll have another update, a little bit later on.

(Music)

Voiceover: You're listening to Samuel's Persiflage.

(Music)

SGS: Well, now its back to Margaret Phillips who is the Director of Digital Archiving at the National Library of Australia. And we might talk about one of your other projects, which is the Internet Archive Project or PANDORA, which is an acronym, but I seem to misplaced the meaning of it, at the moment. But, as you were saying earlier, the newer technology, such as CD ROMs and in fact web pages, don't fall under the Legal Deposit legislation. So, you've started the PANDORA project to help archive some of the Australian websites.

MP: That's correct, yes. This is an enterprise that's very dear to my heart. I have been working on it since 1996. We have our 10th anniversary this year.

SGS: Happy Birthday!

MP: Thank you.

SGS: I am, my website just celebrated one year so I've got some way to catch up yet.

MP: Its quite an achievement, we feel, PANDORA stands for 'Preserving and Accessing Network Documentary Resources Of Australia' and that's really our mission, I suppose, that gives rise to that acronym.

In well, late 1995, it was really we, the Collection Managers at the National Library, began to discuss the fact that there was a lot of material now appearing on the Internet that wasn't also being published in print. So, we weren't collecting it in the usual way. And we decided that we needed to do something about this and we set up a small group to examine what kind of publications were appearing within the Australian web domain. And from that, we developed some selection guidelines and then April of 1996, we setup a small unit of staff to start to select more of those publications, to catalogue them into the National Bibliographic Database and then towards of the end of that year, we also made some initial steps in actually archiving them.

It was late 1997, I think, before there was actually an archive online, which members of the public could have a look at, but since then we have gone from strength to strength. And we now have almost 11 and a half thousand Australian websites and web publications in the Archive. Many of those titles have been archived multiple times, some, many websites change over time and we want to capture the significant changes. There are also quite a number of serial titles, which add new issues every month or every

quarter so, we go back and capture that new content on a regular basis. And there are over 22,000 of those recaptured, what we called 'instances'-- every time we recapture a website, we call that a new instance and there are 22,000 of those in the archive. So, its now quite a sizable body of information.

SGS: Hmm. So, how do you actually go about choosing websites to record in this way? Because obviously, you can't just go and take the .au top-level domain, because there are plenty of websites that are from Australia- mine included for that matter, which don't use the .au domain. I suppose, that could very well be for cost reasons. Hosting in the US, for example, is extremely cheap and .com domain names are also extremely cheap. So, a lot of the Australian web content is being hosted overseas. So, how do you go about choosing the websites to archive?

(00:40:05)

MP: The answer to that question is quite a complex one, and I will attempt to answer it. Firstly, we have the Selective Archive, which we have been working on over the last ten years and I'll talk a bit more that in a minute, but last year we did, in fact, do a harvest of the .au domain. That was a major development in our work of archiving the Australian web domain. We went to the Internet Archive, which is a non-profit organization in San Francisco in the US. They create, I don't know whether all your listeners would have heard of the Internet Archive, they actually go out and tend to archive the whole Web every two months and so, they have been doing a top slice if you like, of the Australian domain, but they haven't been concentrating on -- they haven't allowed enough time for the Australian domain to do it to any depth. So, we decided that we would do just that, we would contract the Internet Archive to run their crawlers over the .au domain for a period of six weeks. And so, they brought us back a lot of material-- 185 million unique files.

So, we have, in fact, attempted that-- as part of that crawl which took place in June and July of last year, the Internet Archive tried out some new technology, which would attempt to pick up at least some of the Australian material without the .au. It did this by, when it was looking at an Australian site, like a .au site, if there were links out to a site that was not .au it went to a Look-up table to see if that IP address was registered as being Australian. And if it was, it would also collect that. So, using that technique we have picked up a proportion, we don't know what proportion, but at least some websites without the .au suffix.

SGS: Right.

MP: So, getting back to the Selective Archive, we have a team of five staff whose job is to identify and select publications from the Australian web domain-- including those, interpreted broadly, including those without the .au at the end. As you know, your own site is included in the Archive.

SGS: Yes.

MP: And they then go through the process of contacting the publisher for permission to archive. Scheduling the harvest using the PANDORA Digital Archiving System that has been developed in house by the National Library and then once its been harvested, they do the quality checking work to make sure that the item is complete and functional. They catalogue the item into the National Bibliographic Database and then they make it available in the public archives and then the public can gain access to it.

SGS: Hmm. I remember, it would have been, I'm just trying to remember now, it was early this year from memory that my site was archived and I remember once I gave permission for it to be archived and I had a look at the result of the archiving work-- I was amazed that it is really a carbon copy of what you see on the web. It has, all the image is there, has all the original formatting so-- not only are you preserving the textual content of it, but you are also preserving the style of the time as well.

(00:45:10)

MP: That's correct. That's what we're attempting to do. Its what we call 'collecting the content' and also what we call a 'look and feel' and we give quite a lot of emphasis to that aspect of it, with the Selected Archive. The big advantage with the Selective Archive, even though we can do only a very small proportion of Australian Web publications in this labour-intensive way, its that we do do the quality checking to make sure that it's all there. And because we have the publishers' permission, we can also provide access to it immediately. Becausethat Legal Deposit legislation does not yet cover online publications; we can't give access to it without the publisher's permission.

SGS: Right.

MP: So, this, the whole domain harvest that we have done, we can't guarantee that the items that we've collected there--the Web resources, will be complete or functional and for the time being, we also can't provide access to it. We felt that it was important for the future, for research purposes into the future that we at least collect it and keep it.

SGS: Yes. The PANDAS system that you use to archive all these web sites is in itself, I am led to believe, quite an amazing piece of technology. Without going into the nitty-gritty of how it works, would you like to give an overview of it?

MP: Yes, I can do that. The Library developed this software because when we, when the Archive began to be of a size that was hard to manage, we needed some software to manage the various work processes that are involved. So, we needed to register our publications, so enter the details about it- the title, its web address. We needed to record information that supported the work processes such as the publisher's name, whether or not they'd given us permission, any restrictions on access that needed to be applied that the publisher had asked for and then we needed that to interface with the harvester, the gatherer. What we preferred to do was to go out into the marketplace and buy some software. But this was such a new activity at the time that it hadn't, the commercial

software providers hadn't seen a ready market for it, so it just wasn't there; we had to develop it ourselves.

SGS: Well to the same extent, a lot of the search engines have developed their systems themselves as well. And they don't do the same job but they do a similar job in that they gather a lot of the information from the web sites and organize it so that when people type something into the search engine, it can return relevant results. And Google, for example, have their own, very unique systems, which they developed. So, to that extent, you're not too dissimilar to the search engines. You have developed your own system that works the way that you need it to work.

MP: That's exactly correct, yes. The other need that we had was to be able to support the other libraries that participate in the PANDORA Archive. The National Library is the Coordinator and we provide the technical infrastructure but all of the Mainland State libraries and the Northern Territory library as well the National Film and Sound Archive, the Australian War Memorial and the Australian Institute of Aboriginal and Torres Strait Islander Studies- all select web publications according to their own selection criteria and they use the PANDAS software to put those into the PANDORA Archive.

(00:50:01)

So, we needed software that would enable these other participating libraries to contribute remotely- its web-based software. And we have had some expressions of interest from other libraries around the world, for instance, the United Kingdom Web Archiving Consortium, which includes the British Library, the National Archive of the UK and some of the main University libraries—formed a consortium to do exactly this same kind of work that we have been doing in Australia and they asked us if they could use software which they have been doing. The National Library Board at Singapore is also interested in using our software. So, it has filled a gap that other libraries have experienced, as well.

SGS: Yes. I could go on talking about this all day because its absolutely fascinating but I will wrap it up here just with this question—How do you foresee the PANDORA project evolving over the next, say, 10 to 20 years? Its bit of a long-range question, I know, but I suppose I'll narrow it down, how do you see this going forward? What do you think is in the future for PANDORA?

MP: In the area of the Web and information technology, it's very hard to see even ten years ahead. Perhaps I'll talk in terms of the next five years…

SGS: That's reasonable.

MP: I would see us developing our understanding and expertise of whole Domain Harvesting. We're planning Whole Domain Harvest later this year so that we can hone our skills and our expertise in that area.

I would hope within the next five years, that we will get legal deposit for online publications; the Library has certainly been lobbying parent department, the Department of Communications, Information Technology and the Arts and the Attorney General to help us to obtain the extension to legal deposit that we need. And if that is the case, then the Whole Domain Harvest, I would foresee, will become a more important aspect of our work because we will then be able to give access to that material. I would expect that we will continue to do the Selective Archiving that we are currently doing but we will do less of that and we will concentrate more just on the key, nationally, significant Australian publications and rely on the Whole Domain Harvest for the wider range of websites, organizational websites, community websites; that sort of thing. So, in the future I would anticipate that we'll see a blending of these two archiving techniques that we engaged in and a development of the Whole Domain Harvesting.

SGS: Right. Just on the legal deposit thing—if the legal deposits were extended to the websites, would that mean that web publishers would need to notify you that they've, that they do have an Australia based web site or would you work that out yourselves? I just ask because if web publishers did have to notify you of it, I can just foresee an awful lot of web publishers not notifying you and probably not realising they are breaking the law by doing such.

MP: That's right. We foresee the same thing and what we have asked for in the legal deposit legislation for online materials, is not that the publishers need to notify us but simply that we have permission to go out and harvest the material that we want from their site and this would only be for the freely available material. We wouldn't be able to harvest the commercial material in that way because there are usually passwords or other mechanisms, which prevent people gaining access without paying, and those mechanisms will prevent our harvester harvesting the material anyway. So, with the commercial material, we would still need to approach the publishers and ask for their permission and their cooperation. So, this would relate only to the freely available material.

(00:55:18)

SGS: All righty, then. Its hard to believe this but we have been talking for nearly 40 minutes here so it's probably a good time to wrap up. How can people find out more about PANDORA and the National Library?

MP: People can find out more about both by going to the National Library's web site which is www.nla.gov.au. .

SGS: Right. I'll include a link to that in the show notes as well.

MP: Okay and down the bottom of the National Library's web page there is a logo, a PANDORA logo and anyone who is interested can go into the archive by following that link.

SGS: And whilst they're there, they can look through information about PANDORA and they can also have a look at the web sites that you've archived and made publicly available.

MP: Indeed, they can.

SGS: Much like, on your main website, the nla.gov.au one, they can also browse through your catalogue as well.

MP: They can; they can gain access to the Library's own catalogue, they can gain access to the Libraries Australia service which I mentioned earlier, which will give them access to 800 Australian libraries' catalogues and many other information sources and services that are available at the National Library-- including how to contact National Library staff if they want assistance in locating information; it's all there on that site.

SGS: Right. Well, Margaret, it's been fascinating talking to you. This is, I was, I've been planning to talk to someone from the National Library for quite a while and I am very glad that I got to talk to you because, being the Director of Digital Archiving you know about PANDORA; you have a very intimate knowledge of PANDORA and obviously, with my website being a part of the PANDORA archive, it's something that is close to my heart as well. So, Margaret, thank you very much for your time and thank you for coming on the program.

MP: Thank you, Samuel for the opportunity to talk about the Library and in particular, PANDORA.

SGS: You're welcome.
(Music)

Voiceover: You are listening to Samuel's Persiflage.

SGS: Alright. Well, we'll have another look at the Persiflage puzzle now. It was interesting last month the feedback I got about the Persiflage puzzle, most of it was good; lot of people seemed to think that it was a good way to keep people listening and engaged throughout the show. And I would have to agree with them because ultimately before we had the Persiflage puzzle in here, I thought the show was going pretty well but it really didn't seem to have much for purpose- it was just me sitting here talking to people a bit. But the show, sort of, has a little bit more of its own unique feel, I think; now because there is something there for you as a listener to do other than just listen. You can quite easily play along a bit if you wanted. It's just a bit of extra fun, I think that it adds into the show. Well, speaking of the puzzle, I think we should probably reveal a few more letters.

So, lets see are pulling out a few more letters, shall we. Well, we have got the letter 'R' for rabbit and its another one of those letters that just doesn't appear in the puzzle. This is becoming a bit of a problem I think, we can eliminate a lot of letters but it would also help to see what's in the puzzle and where it is. So, lets see what we can pull out this

time. And it's the letter 'B' for Bill—and well, that's in there, isn't it; it's the first letter of the first word, in fact. That's a 'B.' And we'll draw out another letter. It's the letter 'T' for Tom. And that's there, two of them- it's the first letter of the fourth word and the fifth letter of the fifth word. So, have a think about that.

(1:00:05)

As I said they were a fairly influential musical group. I won't say much more because if you think about it for long enough, you should get this one-.assuming you know your music so. We'll see how we go with that one. The answer will be revealed a little bit later on.

(Music)

SGS: It's odd news story time! It's one of the more interesting parts of the podcast, I suppose though I think its all interesting. Maybe you have a different opinion, but this is certainly one of my favorite parts of the podcast. Now you might recall last year there was a news story floating about, which was to do with self-heating cans of coffee. You'd walk into a store, you would buy this can of coffee and it had some sort of chemical concoction in it that heated the can of coffee. Well, they're being recalled. It seems that some of them didn't heat adequately and others, they just kept heating and they exploded and melted through the packaging. I really don't think, I would want to be holding one when it did that, but thankfully they've been recalled. So, whether there will be a fix for it or not I don't know.

This is a very unusual one-- a survey of close to 7000 women, aged 20 years or older, shows that snoring is most common among women in their 50s. And the heavier a woman is, the more likely she is snore. Now, it seems that this survey also thinks that the more you smoke or the wider your neck is, and it seems that there is even a link here to high blood pressure and stroke. The higher chance you have of that or the heavier you are or the wider your neck is- the more likely you are at snore. Now, this was all in aid of trying to work out why people get sleep aponea and how it can be fixed. Well, I think, they maybe on the track to solving that one.

Now, here's another story that seems to say, "What if you could turn your spouse's snoring in to nothing more than background noise. It may seem impossible especially if you've lived with it for years. However, some claim there's a way you can get that much needed sleep." It seems that using hypnosis, you can't stop a person from snoring by hypnotizing them and turning them into a chicken. But you can apparently make other people think that the snoring is a different sound. So, I suppose the classic tale of sleeping bliss and people's imaginations is being on a deserted island in the tropics with the waves crashing on to the beach. So, I suppose the theory is here that, the hypnotist will just make you think that your partner's snoring is, in fact, waves on a beach. Well, I don't know, would that work? The hypnotists certainly think it would, but I don't know; it seems a little bit strange to me.

Signals from mobile phones masts have been used to measure rainfall patterns in Israel. A team from the University of Tel Aviv analyzed information, routinely collected by mobile networks, to make their estimates. Writing in the journal called "Science", the researchers say their technique is more accurate than current methods used by the meteorological services. Similar work is being done by scientists in the UK. Now, it seems that what happens here is that when its raining, the mobile signals don't go through strongly enough.

So, the systems, the mobile phone carriers use, just increase the power of the signal automatically and decrease it when it doesn't need to be as strong, because its not raining and it seems that using this data, the scientists can work out how much rain there is.

(01:05:17)

Apparently, they can do something similar with GPS signals. So, that's an interesting, one, although one does have to wonder what happened to the humble weather radar. I can go on the Australian Bureau of Meteorology's website and see the latest radar images and animations of the last few radar images and see where the rain's coming from and how long it will probably be until it's here. So, I think I would rather use that, actually, because that, is that actually tells me where the rain is, not just how much there is-- which, of course, the radar tells me as well.

According to EE-Times, an Israeli Company has developed a personal video display device that looks like a simple pair of glasses. You can use these glasses with various sources, such as a portable media player or your cell phone. Well, why? Just thinking about this, if you are on a public transport for example and you're wearing headphones to listen to music or even a podcast, such as Samuel's Persiflage, that's not too bad, because if you have got to turn down low enough, you can hear your surroundings and you can still see your surroundings quite easily. If you have video glasses on and you're watching something- well, you are going to need the sound as well; you are going to have to have the earphones. So, not only you are eliminating most background noise, so eliminating the noise from your surroundings-- you are also eliminating the visuals of your surroundings. Not only can you not hear what's going on around you, you can't see what's going on around you. So, to start with I don't know how you are going your stop if you are on the bus or the train. And I certainly don't know how this is going to work in terms of personal safety, because you will occasionally get unscrupulous people on public transport or just in general day-to-day, out in public. And I don't think I'd really want to be blocking out my surroundings with those people around. I mean, some people would but not me, be interested to hear what your thoughts on that one.

And if you have got an unusual new story for me to take a look at, podcast@samuelgordonstewart.com is the E-mail address.

(Music)

Well, here's a quick thought for you for the month, "Why is it that we refer to an alarm clock as going to off, when in fact its turning on?"

(Music)

Oh! There's that music again, that must mean its time for us to go. Well, before we do that. Lets just see how you went with the Persiflage puzzle. Well, it was "Bill Haley and the Comets". Now, I have been doing a little bit of research and for some reason, some people seem to refer to them, as "Bill Haley and his Comets". There's two different names, so if that threw you off I apologise for that. I have always known them as "Bill Haley and the Comets" so hopefully you got it anyway.

That's the show for this month. I hope you enjoy it. I know I did. It's fairly redundant for me to say that, isn't it, because, of course, I am going to enjoy it. That's what I do; I enjoy doing the show. And if you have any feedback, well, there is the two options-- podcast@samuelgordonstewart.com or the spoken feedback which is quite simply, you can go to the Persiflage website and click on the button, which says, 'Send Spoken Feedback' or 'Send Audio Feedback.' And if you don't know where the website is and you are using iTunes, then you can actually find your way through to the website from the Samuel's Persiflage iTunes listings. Either way, I hope I can hear from you and get your feedback on the next month, but until then have a good month. I'll see you then. Tada.

(Music)

Total Duration: (01:10:36)