Horizon-scanning the ostrich industry with bibliometric indicators

Timothy Pittaway*
Department of Agriculture and Game Management, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa. E-mail: Timothy.Pittaway@nmmu.ac.za

Pieter van Niekerk
Department of Agriculture and Game Management, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa. E-mail: pieter.vanniekerk@nmmu.ac.za

* Corresponding author

Abstract

Predictive modelling of publication activity is a new concept in agricultural analysis. This study illustrates the potential for identifying and quantifying patterns from the past 42 years of ostrich husbandry research. Studies on ostriches have been influenced considerably by production challenges, opportunities and threats, and owing to the large size of and variations in data and the monopolistic nature of the ostrich industry, there has been limited progress in classifying these research-front trends. Identifying these fronts will facilitate understanding of likely future developments and enable forecasting and the appropriate allocation of resources. This paper presents one of the first exploratory studies to capture and visualise temporal changes in ostrich research activity in the Google Scholar database. The results suggest that there are associations between the deregulation of the industry and increased emerging ostrich research, between the time of political isolation of South Africa and substantially increased global research on ostrich reproduction by 1996, and between epidemiological research and the outbreak of the avian influenza virus between 2004 and 2006. *This indicative study contributes to the development of a novel prototype for the prediction of future developments and trends in agricultural research.*

Key words: *Struthio camelus*; Google Scholar; text-mining; research fronts; avian influenza

1. Introduction

The domestication of the ostrich in 1865 was a significant development for South African ostrich farming (Mellett 1985) and, by the early 1900s, ostrich feathers were one of the most important economic commodities, ranked fourth on the list of South African exports following wool, diamonds and gold (Schutte 2008). The feathers, which were highly prized for high-fashion clothing, appealed mainly to affluent buyers (Stein 2007). But this booming feather industry collapsed during 1913 and it was only at the end of the Second World War that the ostrich industry recovered (NAMC 2003). In the mid-1970s the first ostrich-leather tannery was built in the Oudtshoorn area of South Africa (Hoffman 2008). The tanning of ostrich leather grew into a strong and high-quality, focused industry up until the mid-1990s. During this period the industry procured, processed and exported ostrich leather successfully through a single-channel marketing system (NAMC 2011). There also was an increasing demand for ostrich meat at the beginning of 1990 due to the low-fat, low-cholesterol and high-protein characteristics of the red meat. The ostrich industry in South Africa progressed into a multi-processing business producing three main commodities, namely leather, meat and feathers. Other by-products were ostrich eggs and fat. This industry became export-focused, with meat sold predominantly to Europe and leather sold to Europe, the

United States and Japan. The contribution of the ostrich industry towards export earnings and the balance of payments was estimated at R1,2 billion per annum, representing approximately 5% of total South African agricultural exports (Brand & Jordaan 2011).

The ostrich single-marketing channel provided a protected market presence that allowed the core production area around Oudtshoorn to dominate domestic farming and the global ostrich product arena. The regulated quantity of ostrich skins from South Africa enabled the agencies to 'corner the market'. This marketing channel co-ordinated the quantity of ostrich skins sold to the leather market, which facilitated control of the supply and demand of a Veblen product for European fashion brands. In 1993 the single-marketing channel was deregulated and the monopoly lost control of the ostrich markets and domestic supply, which resulted in an increased number of South African farmers entering the industry, with the effect that ostrich stock numbers doubled and exceeded the demand for skins (Wessels 2003). The range of activities related to ostrich farming spread from the Oudtshoorn area into the rest of the Western Cape, the Free State, Gauteng, Limpopo, Mpumalanga and the Eastern Cape (Greibenow 2003), which also resulted in wider participation from other countries, producers, exporters and agents. Generic material was freely exported and eggs, chicks and breeders were widely sold to international buyers. However, the South African ostrich industry still maintained its global lead due to the highly developed value chain, ranging from farming to slaughtering, to deboning and processing, and finally to marketing (Greibenow 2003). But, according to NAMC figures from 2012, there has been a steady decline in the South African ostrich slaughter figures since 2001, due to the economic climate, farmers not investing in new stock, reduced promotional support, protectionist policies and avian influenza outbreaks (Du Plessis 2013).

From the domestication of the ostrich to the deregulation of the industry, the ostrich industry experienced rapid growth and changes in research, intellectual property, innovation and technology transfer. Ostrich production is typical of industrial agriculture, which combines techno-scientific methods with an intensive agro-processing value chain. Since the processing of the leather, meat, feathers and skins requires substantial technological know-how and advanced equipment, it has resulted in vast amounts of research and advisory literature emerging from private companies, co-operatives and industry organisations. Researchers conducting literature studies face an increasingly complex amount of literature, and the 'research landscape' is also in a continuous state of change. This article turns a vice into a virtue: it deals with the analysis and mapping of the large volumes of academic literature on ostrich husbandry found in the Google Scholar database (GS) by means of a word analysis of titles and keywords in abstracts to determine ostrich-related 'research foci'. It explores integrating the text-mining of Google Scholar with Themeriver^{TM} visualisation, which may serve as a point of departure for other applications in other agricultural research. The paper aims to answer the following two questions:

1. What have the main fronts been in ostrich husbandry research?
2. What external events may have contributed to stimulating the research fronts?

2. Implementation details

2.1 Google Scholar

GS is thought to provide increased access to grey literature, and not only to formally published peer-reviewed journal articles (Kesselman & Watstein 2005). This is the strongest attribute of GS, as it provides searches across all disciplines and from many sources, including theses, articles, books, abstracts, conference proceedings, professional societies, online repositories and universities. The search engine ranking algorithm grades documents by weighing each one according to where it

was published, by whom it was written as well as how often and how recently it has been cited in the literature (Google 2013). It has become an important tool for students with limited access to subscribing libraries or online databases, and especially for postgraduates in the poorer countries, even though it has been shown to have limitations, including lack of reliable advanced search functions, lack of controlled vocabulary and issues regarding scope of coverage (Shultz 2007).

2.2 Text-mining method

Miller (2005) defines text-mining as a process of imposing configurations on the literature so that relevant information can be extracted from it. According to Clark (2013), there is no universally accepted text-mining method, as "it is used by different communities for different purposes". What is common amongst researchers is that text-mining is a process of extracting knowledge from a large amount of unstructured data (Delen & Crossland 2008). According to Chen *et al*. (2012), text analytics has its roots in computational linguistics and information retrieval. Okerson (2013) found that forms of text analysis on a large scale started during the 1980s, when it was first applied to a vocabulary analysis of ancient Greek text entries in order to expand on meanings and to provide links to further information (Perseus Digital Library). The commercial sectors and government have been quick to embrace and deploy text-mining methods to analyse communication and literature data.

The biomedical sector has also been one of the main sectors to embrace text-mining to sharpen clinical and generic criteria to define disease categories (Rebholz-Schuhmann *et al*. 2012). In this case study, the purpose of text mining has been used to quantitatively analyse unstructured academic literature in order to define research fronts.

2.3 Research fronts

The literature sample was validated by assessing the content of research titles and abstracts that made reference to the scientific name of the ostrich, *Struthio camelus*. The keyword 'Struthio' was selected as the search keyword in order to filter out general or inappropriate research referring to ostriches in general. The single research front or 'research focus' was derived from coupling literature that had an association or topic similarity, and these topics were identified by using word-occurrence analysis in literature titles and abstracts. The clusters of topics were allocated to form the following research fronts in the study:

- epidemiology (Newcastle disease, avian influenza, inoculation)
- egg characteristics (colour, recognition, morphology, hatchability, fertility and incubation)
- physiological characteristics (physical functions, mechanics, locomotion, muscle movement, axial skeletal)
- reproduction (breeding, courtship behaviour, gender and group size)
- haematology (immunoglobulin, red cell, antibodies, electrolyte concentrations)
- feed (dry-matter intake, growth rate, nutrition)
- historical background (fossils and Palaeolithic sites)
- stomach (gastric, digestive tract, impaction, colon)
- hydration (thermoregulation and water consumption)
- meat (muscle make up and physical characteristics)

2.4 ThemeRiver™

The ThemeRiver™ visualisation software provides a visual image of variations in the research fronts over time in the form of a flowing 'river', which represents individual research fronts that are

assigned different colours for the length of the river. The visualisation provides the user with the ability to track individual fronts from left to right, with the changing 'river' width (volume) representing changes in the amount of research literature. The distinctive fronts provide the user with comparisons, trends or relationships between the fronts relative to time. The visualisation format provides a timeline below the river and markers for related external events. The spikes in the collection of research fronts can be attributed to the increased cumulative strength of fronts due to increased frequency of research literature. These changes in river width and spikes suggest speculation of influence from external events.

3. Data sources

The text-mining method was applied to the first 1 000 ranked literature sources on GS from 1967 to 2008. The literature represented an assortment of international researchers across all disciplines and from a variety of sources. The first 1 000 ranked literature pieces were then captured from the GS search and the titles and abstract were sorted into topics. After the research topics had been collated, they were refined by removing outliers and inappropriate research that had no direct association with the ostrich industry. These topics were then clustered into the main research fronts found in the ostrich industry. The fronts were then compared to when and how often they had been cited in other scholarly articles. A standard deviation was determined to show variation or dispersion from the average and to measure confidence, so that any extreme citations could be filtered out. The most active research range from 1967 to 2009 was then selected for visualisation on a Themeriver^{TM} graph to discern patterns or trends. Thus, the horizontal flow of thematic patterns over the period could be compared in the context of key ostrich industry milestones to facilitate possible relationships between research topics and corresponding external events.

4. Formation of prominent ostrich research fronts

The text-mining study was conducted on the ranked literature sources on GS and sorted into appropriate research fronts (Table 1). The frequency of literature topics in each research front, together with the number of citations, determined the hierarchy of research fronts. The results from the text-mining exercise indicated that the sample of literature was focused predominantly on epidemiology and egg production. This finding is not unexpected, as Highly Pathogenic Avian Influenza (HPAI) and velogenic Newcastle disease (ND) have been identified as the main diseases of poultry worldwide (Abolnik 2007). HPAI has been particularly disastrous for the ostrich industry, since about 320 000 ostriches were slaughtered worldwide from 2010 to 2011, of which about 211 000 were slaughtered in South Africa (DAFF 2012). In 2012, South Africa was not able to export ostrich meat to the EU, resulting in more than R1 billion in export losses (WCDA 2012). Despite substantial advances in incubator design and incubation techniques since the ostrich industry began in South Africa during the 1800s, problems with embryonic mortality during artificial incubation are still one of the main constraints to the development of the industry worldwide. Research conducted by Brand (2012) found that methods of incubation and chick rearing were among the most important limitations on the development of the ostrich industry, hence egg production has been a strong research front for the industry. However, even though there have been limitations, the growing demand for ostrich products has resulted in increased interest in ostrich reproduction beyond South Africa's borders.

Table 1: Hierarchy of research fronts in the literature on ostrich husbandry

Research fronts	Number of topics allocated to research fronts	Number of citations from other articles (mean)	Percent representation in final sample total (N = 578)
Epidemiology	131	23.76552	22.67 %
Eggs	106	21.12963	18.34 %
Physiology	85	16.38462	14.70 %
Reproduction	58	11.03175	10.03 %
Haematology	54	21.42857	9.34 %
Feed	46	20.02083	7.95 %
Historical background	29	25.40625	5.01 %
Stomach	27	14.07407	4.67 %
Hydration	22	31.11538	3.80 %
Meat	20	16.625	3.46 %

5. Visualisation and sample analysis using ThemeRiver™

The top ten research fronts in Table 1 were then captured in a ThemeRiver™ Graph across a time line from 1967 to 2009 (Figure 1). The increased stacking or bulges (increased frequency of research documents) of the river indicate the increased research front activity. According to Allwright *et al*. (1993), when South Africa experienced the first influenza outbreaks during 1981/1982 and 1986/1997, the river shows a slight increase in research activity on epidemiology after 1980.

The abrupt changes in the river are indicated when economic sanctions against South African were instituted in 1985. This change suggests increased research efforts in order to maintain or increase ostrich production during South Africa's isolation. This trend was disrupted by more focus being directed towards research into haematological and egg-production topics in 1987. In 1991, the first avian influenza outbreak affected the main ostrich farming area near Oudtshoorn and, from this point onwards, the graph shows the start of a persistent epidemiological research front.

One of the most distinctive changes that can be noted in Figure 1 happened in 1993, which indicates an increase in research that may be attributed to the process of deregulation of the marketing of ostrich products. This trend may be evidence of growing interest in learning more about ostrich reproduction and finding ways to increase production in other parts of the world. In this regard, the increased number of studies on ostrich eggs may have contributed to the 62% increase in ostrich production outside of South Africa by 1996. If the international ostrich production trend increases and production outside South Africa persists, it could adversely affect the South African industry, which currently accounts for 80% of ostrich products worldwide. This trend may have been stimulated originally during the South African political unrest and isolation from 1984 on. There also was an increase in the volume of epidemiological research literature from 1993 due to significant disease outbreaks, namely Newcastle disease (1993) and avian influenza outbreaks (1995, 2004 and 2006), and this research front suggests that a concerted effort was made in epidemiological research topics from 1991 to 2009. According to Alexander (2000; 2007) there were only four other recorded avian influenza outbreaks outside South African borders, namely in Italy (2002), the Netherlands (2003) and Zimbabwe (1995, 2005). The height of the spikes in the 2004 and 2006 fronts may indicate an association with the first HPAI outbreaks that occurred in South Africa. The outbreak in the Eastern Cape province of South Africa in 2004 resulted in the euthanasia of 30 000 ostriches, which represented about 40% of the ostriches in the province (Alexander 2007). The HPAI outbreak near Mossel Bay in 2006 resulted in the euthanasia of 8 000 ostriches (SAOBC 2006).

As discussed, ostrich production has been centred mainly in South Africa, but the sample for the ThemeRiver™ graph derived from the Google Scholar database was not confined to South African research literature. The research articles used in this case study were derived from a variety of internationally cited sources. Recent studies on Africa's contribution to the worldwide research show that agricultural research intensity is in decline compared to global competitors (Tijssen 2007; Liebenberg 2013). Therefore, the trend of an increase in ostrich husbandry research from 1967 to 2009 does not reflect the African trend in research outputs in 'mainstream' science.

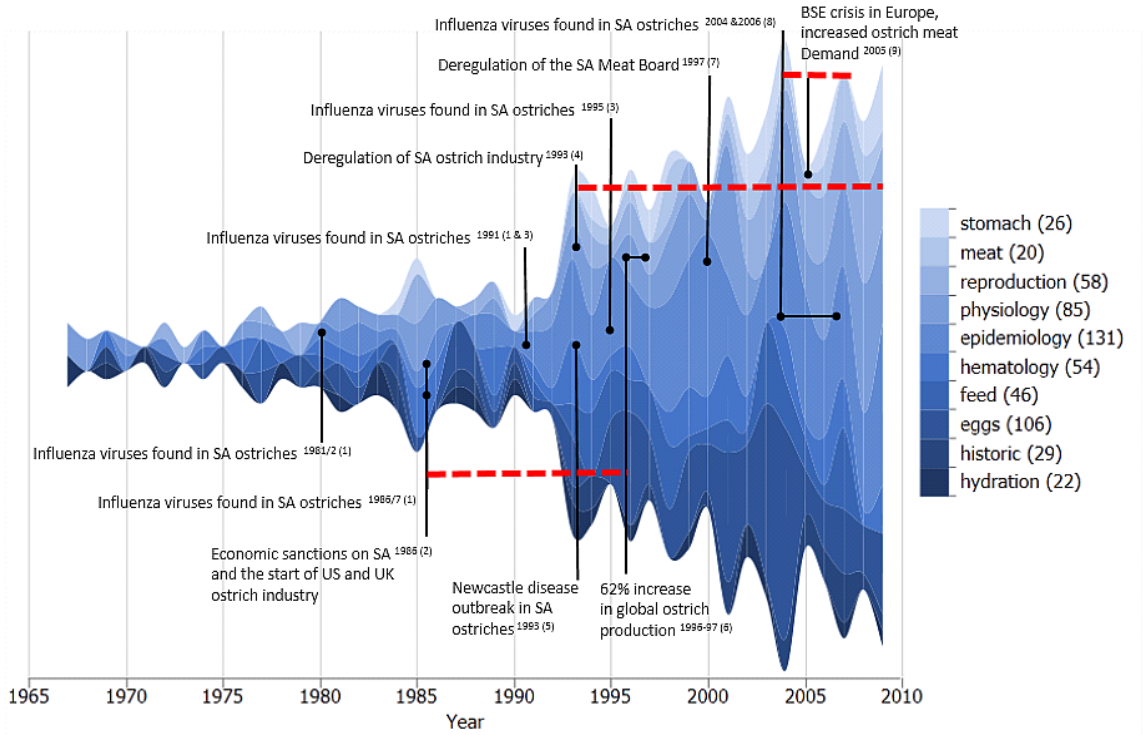

Figure 1: ThemeRiver™ graph of thematic variations between ostrich research documents and external events from 1967 to 2009 (the panel on the right indicates the research front colours and the frequency of literature topics in each research front in brackets, as presented in Table 1).

External event references in Figure 1:
[1] Allwright *et al*. (1993).
[2] Shanawany (1995).
[3] Alexander (2000).
[4] Cooper (2006).
[5] Pfitzer *et al*. (2000).
[6] Cooper (2001).
[7] NAMC (2001).
[8] Abolnik *et al*. (2010).
[9] Hoffman (2008).

6. Conclusion

This exploratory study may contribute to mapping researched topics and theoretical concepts. Owing to the copious amounts of agricultural information available, the fusion of the text-mining of Google Scholar and ThemeRiverTM visualisation could be used as an indicative tool to view vast amounts of research literature. Users are able to visualise thematic changes and obtain a macro-view of variations over time across a large collection of research topics, which may enable the user to identify persistent research fronts. It might make it possible to question contexts, examine the

current and emerging literature trends, and also assist towards a chronological display of the progression of researched themes. As a predictive model it has reveal responses in epidemiology research to disease outbreaks, which today represent one of the main threats to the South African ostrich industry. The usefulness of this prototype demonstrates the ability to reveal comparisons, trends or relationships between research fronts and speculate on influences. This type of study may also provide direction in areas where conservation, environmental or climatic discourse is in a constant state of change, with new or unexplored themes emerging continually.

References

Abolnik C, 2007. Molecular epidemiology of Newcastle disease and avian influenza in South Africa. Doctoral dissertation, University of Pretoria, Pretoria, South Africa.

Abolnik C, Gerdes GH, Sinclair M, Ganzevoort BW, Kitching JP, Burger CE & Olivier AJ, 2010. Phylogenetic analysis of influenza A viruses (H6N8, H1N8, H4N2, H9N2, H10N7) isolated from wild birds, ducks, and ostriches in South Africa from 2007 to 2009. Avian Diseases 54(s1): 313–22.

Alexander DJ, 2000. A review of avian influenza in different bird species. Veterinary Microbiology 74(1): 3–13.

Alexander DJ, 2007. Summary of avian influenza activity in Europe, Asia, Africa, and Australasia, 2002-2006. Avian Diseases 51(s1): 161–6.

Allwright DM, Burger WP, Geyer A & Terblanche AW, 1993. Isolation of an influenza A virus from ostriches (*Struthio camelus*). Avian Pathology 22(1): 59–65.

Brand T & Jordaan J, 2011. The contribution of the South African ostrich industry to the national economy. Applied Animal Husbandry & Rural Development 4: 1–7.

Brand Z, 2012. Studies on embryonic development and hatchability of ostrich eggs. Doctoral dissertation, Stellenbosch University: Stellenbosch, South Africa.

Chen H, Chiang RH & Storey VC, 2012. Business intelligence and analytics: From big data to big impact. MIS Quarterly 36(4): 1165–88.

Clark J, 2013. Text mining and scholarly publishing. Publishing Research Consortium [Online], February, 20(1): 3.

Cooper RG, 2001. Ostrich (*Struthio camelus* var. domesticus) skin and leather: A review focused on southern Africa. World's Poultry Science Journal 57(02): 157–78.

Cooper RG, 2006. Global ostrich industry: Current state and future possible developments. Animal Production and Animal Science Worldwide, 293–8.

DAFF, 2012. Standard for the requirements, registration, and maintenance of registration and official control of ostrich compartments in South Africa. VPN/04/2012-01 (Revision 6.0). Pretoria: National Directorate Animal Health, Department of Agriculture, Forestry and Fisheries.

Delen D & Crossland MD, 2008. Seeding the survey and analysis of research literature with text mining. Expert Systems with Applications 34(3): 1707–20.

Du Plessis, C. 2013. The ostrich industry. Presentation to South African Ostrich Business Chamber, 23 May 2013. Oudtshoorn: Agricultural Business Chamber.

Google, 2013. About Google Scholar. Available at http://www.google.com/intl/en/scholar/about.html (Accessed 1 August 2013).

Greibenow F, 2003. Dream and reality: SA ostrich. Oudtshoorn: Design Studio.

Hoffman LC, 2008. Value adding and processing of ratite meat: A review. Animal Production Science 48(10): 1270–5.

Kesselman M & Watstein SB, 2005. Google Scholar™ and libraries: Point/counterpoint. Reference Services Review 33(4): 380–7.

Liebenberg F, 2013. South African agricultural production, productivity and research performance in the 20th century. Doctoral dissertation, University of Pretoria, Pretoria, South Africa.

Mellett FD, 1985. The ostrich as meat animal: Anatomical and muscle characteristics. Doctoral dissertation, Stellenbosch University, Stellenbosch, South Africa.

Miller TW, 2005. Data and text mining: A business applications approach. New Jersey: Pearson/ Prentice Hall.

NAMC (National Agricultural Marketing Council), 2001. Report on the investigation into the effect of deregulation on the red meat industry. Pretoria: NAMC.

NAMC (National Agricultural Marketing Council), 2003. Report on the investigation into the effects of deregulation on the South African ostrich industry. Pretoria: NAMC.

NAMC (National Agricultural Marketing Council), 2011. A profile of the South African ostrich market value chain. Pretoria: NAMC.

Okerson A, 2013. Text and data mining: A librarian overview. International Federation of Library Associations, World Library and Information Congress, 17–23 August, Singapore.

Pfitzer S, Verwoerd DJ, Gerdes GH, Labuschagne AE, Erasmus A, Manvell RJ & Grund C, 2000. Newcastle disease and avian influenza A virus in wild waterfowl in South Africa. Avian Diseases 655–60.

Rebholz-Schuhmann D, Oellrich A & Hoehndorf R, 2012. Text-mining solutions for biomedical research: Enabling integrative biology. Nature Reviews Genetics 13(12): 829–39.

SAOBC (South African Ostrich Business Chamber), 2006. The South African ostrich industry. Available at http://www.ostrichsa.co.za/ostriches.php?action=get_article&id=17 (Accessed 16 March 2013).

Schutte S, 2008. Development of value added ostrich (*Struthio camelus*) meat products. Doctoral dissertation, Stellenbosch University, Stellenbosch, South Africa.

Shanawany MM, 1995. Recent developments in ostrich farming. World Animal Review 83: 3–8.

Shultz M, 2007. Comparing test searches in PubMed and Google Scholar. Journal of the Medical Library Association 95(4): 442–53. doi:10.3163/1536- 5050.95.4.442

Stein SA, 2007. Falling into feathers: Jews and the trans-Atlantic ostrich feather trade. The Journal of Modern History 79: 772–812.

Tijssen RJ, 2007. Africa's contribution to the worldwide research literature: New analytical perspectives, trends, and performance indicators. Scientometrics 71(2): 303–27.

WCDA (Western Cape Department of Agriculture), 2012. Annual report. Information on predetermined objectives for 2012. Cape Town: Western Cape Government.

Wessels PL, 2003. An analysis of the potential for the marketing of ostrich meat in SA. Unpublished research report, Port Elizabeth Technikon, Port Elizabeth, South Africa.