



BRUNELLA LONGO



London, 14 March 2012

Intellectual Property Office
21 Bloomsbury Street
London WC1B 3HF

copyrightconsultation@ipo.gov.uk

**re: Consultation on proposals to change the UK's copyright system
(Reference: 2011-004) - Response by Brunella Longo**

Dear Sir / Madam,

thank you for the given opportunity to express my views on the above matter and to further understand the current positions of the main stakeholders we have heard during the debate about 'Text and data mining for research' you held in London Yesterday.

Please find the attached document as a way of providing the IPO with my response to the Consultation.

Yours Faithfully

Brunella Longo



CONSULTATION ON PROPOSALS TO CHANGE THE UK'S COPYRIGHT SYSTEM (REFERENCE: 2011-004)

Response by Brunella Longo

1. The respondent

I am a self-employed consultant in the information management sector, author of books and articles as an independent scholar. I have reviewed copyright laws and assurance policies for a number of projects on behalf of public and private customers and professional associations¹.

I am an associate member of the British Computer Society and the Chartered Institute of Librarians and Information Professionals and a qualified IT project manager (Prince2 Practitioner).

I attended the meeting organised by the IPO in London on 13 March 2012 in order to discuss proposals about a new exception to copyright for text and data mining purposes, where I shared most of the following views.

2. About the scope and method of the Consultation

Although I agree with the Government intention to develop a Digital Copyright Exchange in the UK and with the general aim of this consultation, I am skeptical about the possibility that all the actors that could be impacted by such a new system have the time, the experience and the understanding of copyright principles and law needed to answer more than 100 very technical questions.

Nevertheless the Consultation's questionnaire is itself a source of relevant evidence in that it confirms that all the highlighted problems, including those apparently newest because of technological advances, have already been studied, dissected and mostly unsolved through the "exceptions" regime.

3. About the merit of the Consultation

The copyright system needs a radical reform in order to overcome the danger that ineffective, partial and costly adjustments to the "exceptions" regime further erode fundamental principles and universal benefits of copyright by means of exploiting apparent special needs or apparent interests of minorities.

Authors are often unrepresented in Today's copyright system whereas large organisations tend to profit from "commoditization" of authors and researchers' time and expertise either through traditional or to new "open access" channels.

¹ International Online Information Meeting, Sessione Italiana. *L'e-Publishing in Italia nel 2001: a che punto la riforma dell'editoria?* Londra, 5 dicembre 2001. "AIDA informazioni", 20 (2002), n. 1, p. 45-59.



If one considers ownership, resources and provenience of Today's research outcomes one can easily see that a "non commercial research" sector does not exist anymore in many countries including the UK, being the world of universities and research institutions, charities and foundations, publishers and open access actors extensively intertwined with private and commercial organisations.

Such intertwined public, private and not for profit relationships are established through various instruments that include financial agreements (private equities, venture capitals) and commercial and policy deals (patronages, advertising and sponsorship agreements, media, social media and boards exchanges) and extend to the management of supply chain via procurement and outsourcing practices.

Financial, commercial and operational agreements are mostly managed within programmes and projects that tend to be articulated in networked (and often globalised or internationalised) portfolio organisations. These in turn employ flexible staff, temporary workers, part time freelance collaborators, interim managers, consultants, volunteers often without any formal recognition of moral and economic rights due to authorship. Some aspects of the process (teamwork, shared authorships, shared computer platforms and repository) or the same binding publishing requirements / licenses are used to overcome copyright.

In these configurations, I see commercial benefits are sought at the beginning of the research lifecycle as business justifications of both short term endeavours and long term investments, as a sustainable way to assure the needed economic resources to research and publishing programmes and projects.

Thanks to technological developments, this scenario is opened to all sort of legitimate as well as abusive private contributions, including reuses of original contents stored via cloud computing platforms (or "big data" repositories) for advertising reasons or in order to gain competitive advantages (obtained through real time crawling, surveillance systems, desktop virtualisation and other forms of remote control of computer or smartphone equipments).

I believe that we need a radical new way of managing intellectual property rights in all the old and new manifestations, together with new governance practices, so that original creation can be strengthened and rewarded in the new digital marketplaces, reuses can be compensated in a regulated way and abuses can be identified and punished more effectively.

I consider large scale, massive developments of micro-payments and micro-financing formulas for digital authorship or digital reproductions as the only way to encourage protection of a fundamental human right as well as research and commercial interests at present. I also see positive externalities on the side of the development of computer and network security tools and behaviours.

Another instrument to create new formulas and original ways to manage copyright in digital environment consists in promoting and protecting human beings as authors, designers, inventors, creators, contributors or aggregators of digital contents together with software or algorithmic procedures that are



programmed to search, extract, sort, categorize and publish contents either in static or dynamic ways.

In that I would welcome new legislative measures that foster commercial and service level agreements, the design of new product and services and the enforcement of existing copyright and contract law in the sector of text and data mining applications, instead of the introduction of a new exception.

With regard to the technicalities of **question 77 - Text and data mining for research** - these should in my opinion be reframed:

- a) in the context of the European directive on Databases (*sui generis right*) that provides enough clarifications on what may or can be considered object of protection;
- b) taking into account evidences already existing about uses, researchers' behavioural patterns and technological opportunities applied to large repositories of contents, beyond the fascination of data mining applications to the financial sector or bayesian crowdsourcing of user-generated contents.

I personally researched the above issues since 1992². Twenty years ago the conclusion was that unless one knows or imagines what might have been already said, researched, published, dismissed, censored, amended and / or reviewed because of human decision making processes, there are not such many possibilities of serendipitous discoveries, no matter how large are the databases or how long is the "long tail" (number of sources considered) or how powerful is the mathematic assumption of the software - simply because the researcher does not know how to instruct the software to discover what might be interesting.

Following internet and search engine developments and social networking and statistical advances of the last ten years, it seems to me that innumerable contexts confirm that a Pareto law exists governing the distribution of original items in a given universe of data or instances.

Expressed with the terms of a librarian (Bradford law), such universal rule says that 80% of relevant contents are always found in the 20% of available sources.

The only way to make money from knowledge assets, either for cancer research or for dog clothes patents, seems to me to start recognizing them as possible assets especially when they are created by non dominant actors, even before they become discoverable.

Copyright principles offer this possibility.

² Gian Babbeo e le banche dati fulltext: appunti sul metodo della ricerca in linea, in "Biblioteche Oggi", 11 (1993), n. 2.