

Google's PageRank Algorithm: A Diagram of the Cognitive Capitalism and the Rentier of the Common Intellect

Matteo Pasquinelli

At the heart of [Google] is the PageRank algorithm that Brin and Page wrote while they were graduate students at Stanford in the 1990. They saw that every time a person with a Web site links to another site, he is expressing a judgment. He is declaring that he considers the other site important. They further realized that while every link on the Web contains a little bit of human intelligence, all the links combined contain a great deal of intelligence – far more, in fact, than any individual mind could possibly possess. Google's search engine mines that intelligence, link by link, and uses it to determine the importance of all the pages on the Web. The greater the number of links that lead to a site, the greater its value. As John Markoff puts it, Google's software "systematically exploits human knowledge and decisions about what is significant". Every time we write a link, or even click on one, we are feeding our intelligence into Google's system. We are making the machine a little smarter – and Brin, Page, and all of Google's shareholders a little richer.

— Nicholas Carr, *The Big Switch*¹

Abstract

The origin of Google's power and monopoly is to be traced to the invisible algorithm PageRank. The *diagram* of this technology is proposed here as the most fitting description of the *value machine* at the core of what is diversely called knowledge economy, attention economy or cognitive capitalism. This essay stresses the need of a political economy of the PageRank algorithm rather than expanding the dominant critique of Google's monopoly based on the Panopticon model and similar 'Big Brother' issues (dataveillance, privacy, political censorship). First and foremost Google's power is understood from the perspective of *value production* (in different forms: attention value, cognitive value, network value, etc.): the biopolitical consequences of its data monopoly come logically later.

This essay advances three main arguments in relation to the 'Google economy' by focusing respectively: value *production*, value *accumulation* and value *re-appropriation*. First, Google's Page Rank is introduced as the best implementation of the *diagram* of cognitive capitalism. This cognitive and economic diagram actually reverses the Panopticon diagram of Foucauldian lineage: it is not simply an apparatus of surveillance or control, but a machine to capture living time and living labour and to transform the *common intellect* into *network value*. Dataveillance is then made possible only thanks to a monopoly of data that are previously accumulated through the PageRank algorithm. Second, this model of *cognitive hegemony* needs a new theory of *cognitive rent* to be understood, as it is based on the exploitation of a new mediascape for the collective intelligence that is only apparently free and open. Google is defined as a *parasite* of the digital datascape as, on one hand, it provides benevolent free services but, on the other hand, it accumulates value through a pervasive platform of web advertisement (AdSense and Adwords). More importantly, Google establishes its own proprietary hierarchy of value for each node of the internet and becomes then the first systematic *global rentier* of the *common intellect*. Third, a political response can be conceptualised and organised only by reversing the chain of value production (blatantly: 'Reclaiming your page rank') instead of indulging in a nominal resistance to the 'digital Panopticon'.

1. Reversing the Panopticon: Google as a machinic parasite of the common intellect (or, *the production of network-value*).

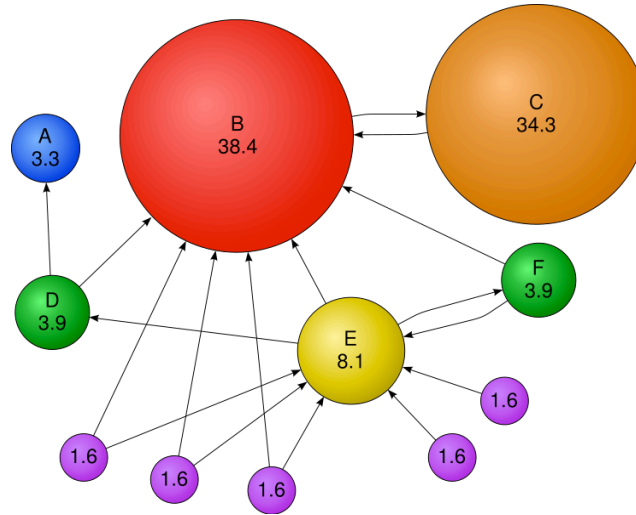
A large part of the recent critical studies about Google focuses only on the *imperial* nature of its monopoly, that is: dominant position, privacy issues, political censorship (see China) and global dataveillance.² Few are the studies about the *molecular* economic engine at the core of this dominion. Whereas many critical texts abuse of a Foucauldian jargon and indulge in the visualisation of a digital Panopticon to describe Google, more precisely its power should be traced back to the economic matrix that is drawn by the cabalistic formula of PageRank — the sophisticated algorithm that determines the importance of a webpage and its hierarchical position within the search engine results.³ As shown in the following paragraphs, PageRank mechanism is enough intuitive to understand but a ‘political economy’ of this apparatus is yet to come.

If the *biopolitical* dimension of Google is widely debated (and often articulated in the above-mentioned post-structuralist jargon), what is missing is a *bioeconomic* analysis to explain how Google extracts value from our life and transform the *common intellect* into *network value* and wealth. Besides true concerns, there is an abuse of a Foucauldian paradigm that highlights only one side of the problem, as Google’s power is not given as a metaphysical being but it is originated from its technological platform and business model. As Paolo Virno puts it, to really understand biopolitics we should begin from the potentiality of our living bodies and from labour power itself: biopolitical structures come later as an apparatus of capture of this potentiality.⁴ The metaphor of the Panopticon must be reversed: Google is not simply an apparatus of dataveillance from above but an apparatus of value production from below. Specifically, Google produces and accumulates value through the PageRank algorithm and by rendering the collective knowledge into a proprietary scale of values — this is the core question. The political economy of Google starts from the political economy of PageRank.

The first description of Google’s PageRank was presented by Sergey Brin and Lawrence Page in their 1998 paper “The Anatomy of a Large-Scale Hypertextual Web Search Engine”.⁵ The PageRank algorithm introduced a revolutionary break in the Information Retrieval technologies and in the search engine panorama of the late 90s: for the first time the apparently flat data ocean of the internet was shaped by Google in dynamic hierarchies according to the visibility and importance of each website. The ranking of a web page is quite intuitive to understand: this value is determined by the number and quality of incoming links. Particularly, a link coming from a node with a high rank has more value than a link coming from a node with a low rank.

While in the late ‘90s search engines like Yahoo were still hand-indexing the web and organising it according to the tree structure typical of encyclopaedic knowledge, Google provided a formula to trace a semantic value across a dynamic and chaotic hypertext. PageRank started to describe webpages according to their popularity and the search engine returned a hierarchy of results according to their rank. Apart from Yahoo’s trees and Google’s rankings, there are other techniques of Information Retrieval and new ones will be developed in the future.⁶ The software and the mathematical formula of the PageRank algorithm are in itself a highly complex construct

accessible only to professional mathematicians: here its understanding is kept at an accessible level to attempt a first political analysis of this apparatus.



Pic. 1 – A visual example of rank value calculated by PageRank.
Source: <http://en.wikipedia.org/wiki/PageRank>⁷

The PageRank diagram depicted above has no resemblance with the centralised structure of the Panopticon described by Foucault in *Discipline and Punishment*.⁸ The liquid and hypertextual nature of the web (and more generally of the noosphere) demands a further illustration. A diagram of cognitive capitalism can be intuitively traced if — within the structure of a hypertext — each symmetrical link is replaced by an asymmetrical vector of energy, data, attention or value. What PageRank unveils and measures is precisely this asymmetrical constitution of any hypertext and network.

The source of inspiration for PageRank was the academic citation system. The ‘value’ of an academic publication is notoriously calculated in a very mathematical way according to the number of citations that an article receives from other articles. Consequently, the general rank of an academic journal is the sum of all the incoming citations received by its articles. As Brin and Page explain:

Academic citation literature has been applied to the web, largely by counting citations or backlinks to a given page. This gives some approximation of a page's importance or quality. PageRank extends this idea by not counting links from all pages equally, and by normalizing by the number of links on a page.⁹

This bookish genealogy of PageRank should not be underestimated. A similar way to describe value can be applied to any *cognitive object* and it is *native* also to the ‘society of the spectacle’ and its wild economy of brands. In a spectacular regime the value of a commodity is produced mainly by a condensation of attention and collective desire driven by mass media and advertisement. From academic publications to commercial brands and the

internet ranking itself equivalent processes of condensation of value can be assumed. As the digital colonisation gave an *online* presence to any *offline* entity, this matrix of social and value relations migrated online and it became digitally traceable and measurable by search engines. PageRank specifically describes the attention value of any object to a such extent that it has become the most important source of visibility and authority even outside the digital sphere. Eventually PageRank gives a formula of value accumulation that is hegemonic and compatible across different media domains: an effective diagram to describe the *attention economy* and the *cognitive economy* in general.

The notion of *attention economy* is useful to describe how (part of) the value of a commodity is produced today via a media-driven accumulation of social desire and intelligence.¹⁰ Regarding the constitution of this value, other schools of thought may refer to *cultural capital* (Pierre Bourdieu),¹¹ *collective symbolic capital* (David Harvey)¹² or *general intellect* (especially in the tradition of post-Operaismo, with a more cognitive spin). Before the internet this process was described as a generic collective drive – after the internet, the structure of the network relations around a given object can be easily traced and measured. PageRank is the first mathematical formula to calculate the *attention value* of each node in a complex network and the general *attention capital* of the whole network. What is the nature of the value that is measured by PageRank? More interestingly each link and vector of attention is not simply an instinctive gesture but a concretion of intelligence and often a conscious act. If it is fashionable to describe the network society as a conurbation of *desiring flows*, however those flows are dense of knowledge and belong also to the activity of a common intelligence.

In the introducing quote of this article, Nicholas Carr described very well how Google's Page Rank works, how it feeds on our collective intelligence and how value is produced and accumulated starting from this common intellect. PageRank establishes so its own attention economy, but a great part of this attention capital is more precisely built on intellect capital, as each link represents a concretion of intelligence. In this sense Google is a parasitic apparatus of capture of the value produced by the common intelligence.¹³

2. PageRank: a diagram of cognitive capitalism (or, the network value).

Can network theory exist without a notion of *network value* — a notion of value specific to the network ecosystem and economy? Sounding the cognitive density of the internet, PageRank unveils precisely a mechanism that is responsible for setting a *rank value* for each node of the web. This *rank value* set by Google is unofficially recognized as the currency of the global attention economy and crucially influences the online visibility of individuals and companies and subsequently their prestige and business. This *attention value* is then transformed in monetary value in different ways. If the PageRank algorithm occupies the inner core of Google's hegemonic matrix, its revenues are coming from the advertisement platform Adwords that exploits this dominant position (99% of revenues are derived from advertisement according to 2008 Annual Report).¹⁴ The PageRank algorithm plus gigantic data centres (running 24-hour and constantly indexing the web) provide a monopolistic position for Google advertisement channels.

The way through which Google generates value deserves a more attentive analysis, as contrary to traditional mass media Google does not produce any content by itself. Specifically, Google captures millions of websites and users through its advertisements syndication program AdSense. Google's AdSense provides a light infrastructure for advertising that infiltrates each interstice of the web as a subtle and mono-dimensional parasite, extracting profit without producing any content. Money enters the cycle in Adwords and are then distributed through AdSense to single bloggers or web companies. Within the economy of the internet, both the traffic of a website and the redistribution of value is today extensively governed by PageRank. PageRank is at the core of the attention economy of the internet as well as at the core of a *general economy of prestige* that affect many other domains controlled directly or indirectly by Google (take for instance academia and Google Scholar, music industries and Youtube, etc. — many, for instance, are the cases of a symbiosis between the internet and the show business).

What PageRank identifies and measure is *network value* in a very numeric form. If a commodity is described traditionally by *use-value* and *exchange-value*, *network-value* is a further layer attached to the previous ones to describe its 'social' relations. This term is ambiguous for many as it might simply point to a 'value of networks' (like in Benkler's much-celebrated 'wealth of networks').¹⁵ To be more precise, a new notion of *network surplus-value* should be advanced and articulated here.¹⁶ Indeed, PageRank produces what Deleuze and Guattari described as a *machinic surplus-value* referring to the surplus-value accumulated through the cybernetic domain, that is the transformation of a *surplus-value of code* in a *surplus-value of flux*.¹⁷ Through PageRank, Google has not simply conquered a dominant position in the storage of web indexes, but also the monopoly of the production of this *network value*.

The diagram of PageRank underlines an important aspect about the relation between two nodes of a network. This relation is never purely symmetrical yet asymmetrical: each link features indeed a one-way direction like an arrow, each link represents an exchange of desire, attention and knowledge that is never symmetrical. This relation is never binary and equal, but actually *ternary*, as there is always a third node influencing it and then an accumulation of value absorbed to another direction. A network is never flat and horizontal. The digital *ontology* is always influenced by external values and material networks, by the analogue world of labour and life (that is the influence of the bio-political and bio-economic fields). A network is never symmetrical and homogenous, it is a topological surface rippled in molecular vortices. Between the *vertical* hierarchies of traditional knowledge and the so-much celebrated *horizontal* networks of today's knowledge production, this *vortical* dimension shows how the two axes are always connected and how dynamic hierarchies keep on following us also onto the digital realm. Google's PageRank installed itself precisely on this movement that shapes the collective sphere of knowledge and the internet in molecular vortices of value.

Putting together the semantic topology of PageRank, the *vortical* accumulation of value affecting networks and the notion of *machinic surplus-value* in a single theoretical object, we can start to sketch a new diagram of the

(aside from the PageRank patent!). So which sort of cognitive rent is embodied by Google? After reversing the Panopticon model, it is necessary also to reverse the common interpretations regarding network economy and the production of network value.

A new understanding and model of rent has been recently advanced within the post-Operaismo debate. Carlo Vercellone and Antonio Negri, Christian Marazzi and many others have rediscovered the dimension of rent as a crucial node in the shift from industrial capitalism to cognitive capitalism.²² What reminded so much of the *Ancient Regime* and its feudal landowners, relict of a pre-industrial economy, today reincarnates itself under the forms of financial and cognitive rent.

In classical economic theory, rent is distinguished from profit. Rent is the *parasitic* income an owner can earn just by possessing an asset and it is traditionally associated with land property. Profit, on the other hand, is meant to be *productive* and it is associated with the power of capital to generate and extract surplus-value (from commodities and the workforce). Yet Vercellone criticises the idea of a “good productive capitalism” by highlighting the *becoming rent of profit* as the characteristic trait of current knowledge and financial economy.²³ Vercellone, accordingly, provides a slogan for cognitive capitalism: “rent is the new profit”. Accordingly, Google can be described as a global rentier that is exploiting the new lands of the internet with no need for strict enclosures and no need to produce content too. In this picture, Google appears as pure rent on the *meta* dimension of information that is accumulated through the digital networks. Google does not possess the information of the internet but the fastest diagram to access and measure the collective intelligence that has produced it.

The rent form is a more suitable model to describe the exploitation of the common intellect and the *common* itself (if profit and wage are more related to an individual dimension and rent to a more collective and social dimension of production). Such a new theory of rent is useful to escape the impasse of the so-called *new media criticism* that is still incapable to identify the axes of production and exploitation along the digital domain. A taxonomy of the new forms of rent and new business models is necessary and urgent. For Negri and Vercellone themselves, for example, the central axis of contemporary valorisation is the “expropriation of the common through the rent”. According to them (as well as many others), this explains the ongoing pressure for a stronger intellectual property regime: copyright is one of the strategic evolutions of rent to expropriate the cultural commons and reintroduce artificial scarcity. Speculation then is directed toward intellectual property, forcing artificial costs on cognitive goods that can paradoxically be reproduced or copied virtually for free. However, the composite case of intellectual property must be further illuminated, as rent may not necessarily arise simply from knowledge enclosures, but also from the exploitation of cognitive spaces that are completely new and virgin, as Google shows in relation to the internet. The PageRank diagram seems to suggest a sort of *differential rent* along dynamic spaces that would deserve a further investigation.²⁴

5. Conclusion: 'Reclaim your page rank' (or, *the reappropriation of network value*).

A consistent political response to Google's neo-dominion should be based on an alternative ranking system able to undermine the monopoly of attention economy and also the accumulation of value controlled by Google. Can such a monopolistic production of network value be reversed in some way? A first option would be to imagine a collective voluntary hand-made indexing of the web based on an open protocol (a sort of Wikipedia of network relations described under the FOAF ontology).²⁵ However, Google cannot be challenged on the scale of its computing power: such a competition would be quite silly and primitive. On the other hand, the idea of an *open source page rank* algorithm would not address the issue of value accumulation and monopoly. By the way, the idea of an OpenRank algorithm has been rapidly abandoned.²⁶ The fatal attraction of the masses for Google seems to rely more on its mystical power to set a *spectacular value* for anything and anybody than on the precision of its results. Rumours say that PageRank will be replaced soon by TrustRank, another algorithm developed by Stanford University and Yahoo researchers to separate useful webpages from spam and establish a sort of community trust or a new *cybernetic social pact* across the internet.²⁷ In such a scenario the everyday life and production of social networks will be integrated in an even deeper way.

The battle against the accumulation of data operated by PageRank reminds the social struggles against the traditional forms of monopoly and accumulation of capitals. PageRank is to the internet, as primitive accumulation and rent are to early capitalism.²⁸ If we refer to Marx's *general intellect*, we should imagine also an *original accumulation of knowledge* at the source of the digital economy. Anyhow, a critique of the present mode of networking cannot be established simply on the predictable narrative of the good networks against the evil monopolies. A political response can be imagined only if the nature of the molecular *dispositif* that produces the *network value* is understood. PageRank and Google cannot be easily made more democratic. On the other side, interestingly, also the new fashionable schools of peer-to-peer cooperation and internet-based "social production" will fail to represent a decent political proposal until they address the issue of production and accumulation of *network surplus-value*.

¹ Nicholas Carr, *The Big Switch: Rewiring the World, from Edison to Google* (New York: W.W. Norton, 2008).

² See: Michael Zimmer, "The Gaze of the Perfect Search Engine: Google as an Infrastructure of Dataveillance", in A. Spink and M. Zimmer, *Web Search: Information Science and Knowledge Management*, vol. 14, Berlin-Heidelberg: Springer, 2008. And: Benoît Dupont, "Hacking the Panopticon: Distributed Online surveillance and Resistance", in: Mathieu Deflem (ed.), *Surveillance and Governance: Sociology of Crime Law and Deviance*, vol. 10, Emerald, Bingley, 2008. See also the critique advanced by the president of the *Bibliothèque nationale de France*: Jean-Noël Jeanneney, *Quand Google défie l'Europe : plaidoyer pour un sursaut*, Paris: Mille et une Nuits, 2005.

³ 'PageRank' (without space) is a trademark and patent owned by Google Inc.

⁴ Paolo Virno, *A Grammar of the Multitude* (New York: Semiotexte 2004), pp. 81-84: "In my opinion, to comprehend the rational core of the term "bio-politics," we should begin with a different concept, a much more complicated concept from a philosophical standpoint: that of labor-power. [...]What does "labor-power" mean? It means *potential* to produce. Potential, that is to say, aptitude, capacity, *dynamis*. Generic, undetermined potential: where one particular type of labor or another has not been designated, but *any* kind of labor is taking place, be it the manufacturing of a car door, or the harvesting of pears, the babble of someone calling in to a phone "party-line," or the work of a proofreader. [...] Having said this, we still have to address a crucial question: why is life, as such, managed and controlled? The answer is absolutely clear: because it acts as the substratum of a mere faculty, labor-power, which has taken on the consistency of a commodity. [...] One should not believe, then, that bio-politics includes within itself, as its own distinct articulation, the management of labor-power. On the contrary: bio-politics is merely an effect, a reverberation, or, in fact, one articulation of that primary fact—both historical and philosophical—which consists of the commerce of potential as potential."

⁵ Sergey Brin and Lawrence Page. 'The Anatomy of a Large-Scale Hypertextual Web Search Engine', 1998. <http://infolab.stanford.edu/~backrub/google.html>

⁶ See: Amy N. Langville and Carl D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, 2006. And: Michael W. Berry and Murray Browne, *Understanding Search Engines: Mathematical Modeling and Text Retrieval* (Philadelphia: Society for Industrial and Applied Mathematics, 1999).

⁷ Wikipedia comments this image in this way: "Mathematical PageRanks (out of 100) for a simple network (PageRanks reported by Google are rescaled logarithmically). Page C has a higher PageRank than Page E, even though it has fewer links to it: the link it has is much higher valued. A web surfer who chooses a random link on every page (but with 15% likelihood jumps to a random page on the whole web) is going to be on Page E for 8.1% of the time. (The 15% likelihood of jumping to an arbitrary page corresponds to a damping factor of 85%.) Without damping, all web surfers would eventually end up on Pages A, B, or C, and all other pages would have PageRank zero. Page A is assumed to link to all pages in the web, because it has no outgoing links." Source: <http://en.wikipedia.org/wiki/PageRank>, 12 March 2009.

⁸ Michel Foucault, *Surveiller et punir. Naissance de la prison* (Paris, Gallimard, 1975); trans.: *Discipline and Punish: The birth of the prison* (New York: Pantheon, 1977).

⁹ Sergey Brin and Lawrence Page. 'The Anatomy of a Large-Scale Hypertextual Web Search Engine', 1998. <http://infolab.stanford.edu/~backrub/google.html>

¹⁰ Herbert Simon, "Designing Organizations for an Information-Rich World", in M. Greenberger (ed.), *Computers, Communication, and the Public Interest*, Baltimore: Johns

Hopkins Press, 1971. See also: T. Davenport and J. Beck, *The Attention Economy: Understanding the New Currency of Business*, Harvard Business School Press, 2001.

¹¹ Cfr. Pierre Bourdieu, "The Forms of Capital", in: J.G. Richardson, *Handbook for Theory and Research for the Sociology of Education*, Westport, CT: Greenwood Press, 1986.

¹² Cfr. David Harvey, "The art of rent: globalisation and the commodification of culture", chapter in: *Spaces of Capital*, New York: Routledge, 2001.

¹³ This model of capturing the network capital of each node and user can be applied generically to the whole internet and, for instance, to social networks as well (like Facebook and Myspace, where the most important number is precisely the figure of 'friends' showed on each personal page).

¹⁴ "AdWords is Google's flagship advertising product and main source of revenue (\$16.4 billion in 2007). AdWords offers pay-per-click (PPC) advertising, and site-targeted advertising for both text and banner ads. The AdWords program includes local, national, and international distribution. Google's text advertisements are short, consisting of one title line and two content text lines." [Source: Wikipedia entry 'Adwords', 11 March 2009]

¹⁵ Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven: Yale University Press, 2006).

¹⁶ This network-value should be distinguished from the traditional definition: Metcalfe's law of "network value" states that the value of a telecommunications network is proportional to the square of the number of connected users of the system (n^2).

¹⁷ Gilles Deleuze and Félix Guattari, *L'Anti-Oedipe. Capitalisme et schizophrénie* (Paris: Minuit, 1972). Translation: *Anti-Oedipus: Capitalism and Schizophrenia* (Minneapolis: University of Minnesota Press, 1983)

¹⁸ For further models of knowledge economy see: Enzo Rullani, *Economia della conoscenza: Creatività e valore nel capitalismo delle reti*, Milano: Carocci, 2004.

¹⁹ Originally, the precise 1998 definition adopted by the Creative Industries Task Force set up by Tony Blair stated: "Those industries that have their origin in individual creativity, skill and talent and which have a potential for wealth and job creation through the generation and exploitation of intellectual property."

²⁰ See: Richard Florida, *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life* (New York: Basic Books, 2002)

²¹ See: Paolo Virno, *A Grammar of the Multitude. For an Analysis of Contemporary Forms of Life* (New York: Semiotexte, 2004).

²² Antonio Negri e Carlo Vercellone, "Il rapporto capitale/lavoro nel capitalismo cognitivo", in *Posse*, "La classe a venire", Nov. 2007. Web: www.posseweb.net/spip.php?article17

²³ As Vercellone explains in a previous study: "According to a widespread opinion in Marxian theory that stems from Ricardo's political economy, rent is a pre-capitalist inheritance and an obstacle to the progressive movement of capital's accumulation. On this premise, real, pure, and efficient capitalism is capitalism with no rent." In: Carlo Vercellone, "La nuova articolazione salario, rendita, profitto nel capitalismo cognitivo", in *Posse*, "Potere Precario", 2006; trans. by Arianna Bove, "The new articulation of wages, rent and profit in cognitive capitalism", Web: www.generation-online.org/c/fc_rent2.htm

²⁴ See my taxonomy of rent in: Matteo Pasquinelli, *Animal Spirits: A Bestiary of the Commons* (Rotterdam: NAI Publishers / Institute of Network Cultures, 2008).

²⁵ FOAF (an acronym of Friend of a Friend) is a machine-readable ontology describing persons, their activities and their relations to other people and objects. Anyone can use FOAF to describe him or herself. FOAF allows groups of people to describe social networks without the need for a centralised database. See: www.foaf-project.org

²⁶ See: www.openrank.org

²⁷ Gyöngyi, Zoltán; Hector Garcia-Molina, Jan Pedersen. "Combating Web Spam with TrustRank", in: *Proceedings of the International Conference on Very Large Data Bases*, 2004

²⁸ The so-called *original accumulation* or *primitive accumulation* is still at the core of the present history of capitalism as Sandro Mezzadra warns us in: Sandro Mezzadra, *La condizione postcoloniale. Storia e politica nel presente globale*, Verona: Ombre Corte, 2008.