

L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE ET ACTIFS CONCURRENTIELS CRITIQUES: DISCUSSION DE L'ESSENTIALITÉ DES DONNÉES

***Documents de travail GREDEG
GREDEG Working Papers Series***

FRÉDÉRIC MARTY

GREDEG WP No. 2024-12

<https://ideas.repec.org/s/gre/wpaper.html>

Les opinions exprimées dans la série des **Documents de travail GREDEG** sont celles des auteurs et ne reflètent pas nécessairement celles de l'institution. Les documents n'ont pas été soumis à un rapport formel et sont donc inclus dans cette série pour obtenir des commentaires et encourager la discussion. Les droits sur les documents appartiennent aux auteurs.

*The views expressed in the **GREDEG Working Paper Series** are those of the author(s) and do not necessarily reflect those of the institution. The Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate. Copyright belongs to the author(s).*

L'intelligence artificielle générative et actifs concurrentiels critiques : discussion de l'essentialité des données ¹

Frédéric Marty

CNRS – GREDEG – Université Côte d'Azur

GREDEG Working Paper No. 2024-12

Résumé

Le développement de l'intelligence artificielle générative fait l'objet d'une attention particulière de la part des autorités de concurrence. Ses impacts peuvent être déterminants en ce qu'elle peut aussi bien rebattre les cartes du jeu concurrentiel, c'est-à-dire affaiblir les positions de force des grandes firmes pivot des grands écosystèmes numériques actuels, que donner lieu à une nouvelle consolidation, en leur permettant d'étendre leur contrôle à cette technologie d'usage général qui est appelée à exercer un rôle déterminant dans la structuration de notre économie.

Le ressort des initiatives des régulateurs de la concurrence tient à la crainte que le contrôle de certaines ressources essentielles conduise à étendre la puissance économique de ces acteurs vers ce nouveau marché. Les autorités de concurrence feraient dès lors face aux mêmes enjeux que ceux induits par les situations de dominance et de verrouillage des écosystèmes actuels : difficultés dans la définition et dans la mise en œuvre de remèdes concurrentiels effectifs ou encore nécessité d'instaurer des réglementations spécifiques pour prévenir les dommages concurrentiels.

Mots clés

IA générative, avantage lié aux données, écosystèmes numériques, Big Techs

Codes JEL

K21, L12, L13, L41

¹ Ce document de travail est préparé dans le cadre d'une participation au séminaire du CERCRID (CNRS - Université de Saint Etienne), *Données massives et libre concurrence*, organisé le 8 avril 2024 par Mouna Mouncif-Moungache et Vincent Giovannini. Ce document doit beaucoup aux échanges avec de nombreux collègues notamment dans le cadre de la conférence de la Chaire Numérique et Concurrence de l'Université Paris Panthéon Assas du 14 novembre 2023, *IA générative et concurrence*, du workshop organisé par la FIDES aux Orres le 15 mars 2024 (*Les 4^{èmes} entretiens de Saint Sauveur*), de la préparation d'une réponse commune de l'AFEC à la consultation de l'Autorité de la concurrence relative à la saisine d'office pour avis portant sur le secteur de l'intelligence artificielle générative (22 mars 2024) et, enfin, au sein de l'équipe de recherche du GREDEG et de son équipe CCorsé, avec notamment Ekaterina Prytkova, Patrice Bougette et Simone Vannuccini. Toute erreur ou imprécision demeure de ma pleine responsabilité.

Abstract

Competition authorities are paying particular attention to the development of generative artificial intelligence. Its impact can be decisive in that it can both reshuffle the cards of the competitive game, i.e. weaken the positions of strength of the major firms at the heart of today's major digital ecosystems, and give rise to new consolidation, by enabling them to extend their control over this general-purpose technology which is destined to play a decisive role in the structuring of our economy.

The driving force behind the initiatives of competition regulators is the fear that control of certain essential resources will lead to the economic power of these players being extended to this new market. Competition authorities would face the same challenges as those arising from dominance and foreclosure in current ecosystems: difficulties in defining and implementing effective competitive remedies and the need to introduce specific to prevent competitive damages.

Keywords

Generative AI, data-based advantage, digital ecosystems, Big Techs

JEL codes

K21, L12, L13, L41

I – Les enjeux concurrentiels de l’IA générative la donnée est-elle un actif essentiel ?

Il n’est de meilleure entrée sur les enjeux concurrentiels liés au développement de l’intelligence artificielle (ci-après I.A.) générative que le rapport publié en novembre 2023 par l’autorité portugaise de la concurrence dont nous reproduisons *infra* l’un des graphes de synthèse.

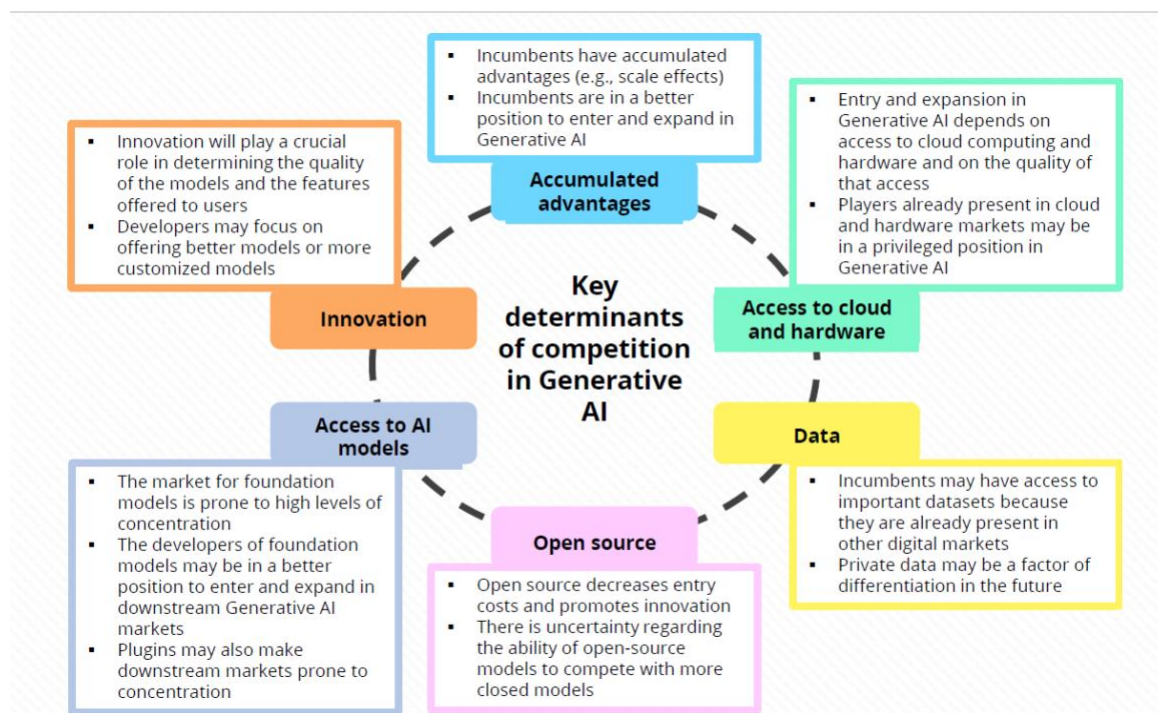


Figure 1 : les déterminants clés de la concurrence dans le secteur de l'IA générative (source Autorité portugaise de la concurrence, 2023)

Celle-ci nous permet en effet de cartographier les différents risques concurrentiels qui caractérisent le développement de l’IA générative². Ces risques peuvent notamment tenir à des situations de verrouillage quant à l’accès à certaines ressources clés que sont notamment les données. Les données ne sont pas les seuls actifs critiques susceptibles de fournir un avantage déterminant aux firmes dominantes de l’économie numérique actuelle. Le contrôle de ces actifs critiques peut les mettre en position d’influer significativement sur la dynamique concurrentielle de l’IA générative que cela soit pour forclure l’accès au marché des concurrents

² De façon très résumé l’IA générative se distingue de ses prédécesseurs en ce qu’elle est capable à partir d’un apprentissage sur de larges bases de données de générer à partir de questions (prompts) des contenus originaux pouvant prendre la forme de textes, de lignes de code, d’images, de sons ou de vidéos.

ou étendre leur position dominante à ce nouveau marché par stratégie de levier anticoncurrentiel³.

Dans les deux cas des pratiques d'auto-favoritisme peuvent être engagées pour favoriser ses propres services ou ceux d'un opérateur 'partenaire' au détriment des tiers (Marty, 2021). Un refus d'accès à ces ressources critiques ou encore un accès dans des conditions discriminatoires peut conduire à fausser la concurrence. Comme le montre la note de l'Autorité de la concurrence portugaise (2023), ces actifs ne tiennent pas aux seules données que nous abordons de façon privilégiée dans notre document de travail mais tiennent également à des actifs physiques tels les infrastructures et services de cloud ou encore des processeurs graphiques. Ces infrastructures constituent autant de goulets d'étranglement que peuvent instrumentaliser des acteurs déjà installés pour entraver le développement de nouveaux entrants. Ces opérateurs dominants de l'économie numérique peuvent également bénéficier d'avantages liés à leur position de contrôleurs d'accès aux grands écosystèmes numériques, de leur portefeuille de clients, de leurs capacités à intégrer de nouveaux services à des services existants ayant une très grande base d'utilisateurs....

Ainsi, les craintes affichées par de nombreuses autorités de concurrence ou par les parties prenantes s'inscrivent dans un contexte dans lequel ce qui pourrait apparaître comme une opportunité de disruption peut éventuellement se traduire par une nouvelle extension des positions dominantes vers des marchés adjacents au travers de la prise de contrôle directe ou indirecte des acteurs de ce marché, c'est-à-dire à l'extension des services rendus par les écosystèmes dominants au marché de l'IA générative⁴, voire à des phénomènes de double marginalisation par articulation de deux types d'opérateurs interdépendants et verticalement reliés.

³ Il convient de distinguer les algorithmes d'IA et notamment d'IA générative des algorithmes de recherche ou d'appariement propres à l'économie numérique traditionnelle. Un algorithme se caractérise par des coûts fixes élevés pour sa conception mais par des coûts marginaux extrêmement faibles pour son utilisation. A l'inverse le déploiement d'une IA générative et son utilisation est coûteuse et nécessite de reposer sur une infrastructure. En ce l'IA générative doit être envisagée dans un cadre théorique se rapprochant de celui des télécommunications. Elle appelle donc une analyse s'attachant bien plus aux complémentarités technologiques entre les différentes strates composant la filière (voir pour une analyse complète Vannuccini et Prytkova (2023)).

⁴ Voir notamment la réponse d'Open Markets à la consultation de la Commission européenne sur l'IA générative (lancée le 9 janvier 2024), rendue publique en mars 2024.

<https://www.openmarketsinstitute.org/publications/submission-ec-ai-concentration>

https://ec.europa.eu/commission/presscorner/detail/en/IP_24_85

Voir également la consultation publique initiée par l'Autorité de la Concurrence

https://ec.europa.eu/commission/presscorner/detail/en/IP_24_85

Cependant, si l'accent est souvent mis sur la nécessité de *prévenir une extension de la position dominante vers le domaine de l'IA générative*, plusieurs questions peuvent être posées. Existe-t-il un seul marché de l'IA générative ? Il convient en effet de distinguer un segment amont, celui des modèles de langage fondamentaux et un segment aval correspondant aux modèles adaptés à des segments spécifiques⁵.

Cette distinction conduit à se pencher sur deux types de concurrence : d'une part les relations de concurrence horizontale sur ces deux segments amont et aval et d'autre part les relations de concurrence verticale entre ces derniers. De la même façon, le narratif du verrouillage concurrentiel d'un marché naissant peut être discuté au regard de plusieurs hypothèses alternatives. Les acteurs dominants des modèles fondamentaux de langage ne sont-ils pas de nouveaux entrants ? Sont-ils à l'inverse dépendants des infrastructures des grands acteurs de la Tech voire liés contractuellement à ces derniers ? Retenir la première hypothèse conduit à l'instar de Manne et Auer (2024) à considérer que des entraves à des obligations asymétriques pesant sur ces derniers pourraient affaiblir la concurrence sur le marché de l'IA générative en affaiblissant de grands groupes qui pourraient potentiellement être *disruptés* par de nouveaux entrants. A l'inverse, faire sienne la seconde option peut à conduire à préconiser la mise en place de garde-fous visant à prévenir des risques de verrouillage que ceux-ci procèdent de stratégies d'auto-favoritisme ou de stratégie d'intégrations verticales que celles-ci prennent la forme d'acquisitions (croissance externe) ou de quasi-intégration par contrats au travers de partenariats⁶. Une autre dimension de la concurrence peut-être également considérée : elle tient à la capacité des acteurs aval qui développent des modèles fines-tunés mis à disposition au travers d'applications ou de plug-ins de dégager de la valeur sans que celle-ci puisse être captée par les firmes amont. A nouveau les questions de dépendance économique et technologique sont à considérer.

Notre développement va accorder une place importante à l'analyse de l'un des verrous ou l'une des voies de levier de comportement anticoncurrentiel : les données. A cette fin les deux prochaines sections vont successivement s'attacher à une discussion de l'importance des données comme actif crucial dans le développement des grandes firmes du numérique et ensuite

⁵ Il s'agit ici de modèles fine-tunés en d'autres termes ayant fait l'objet d'un réglage fin i.e. d'une adaptation à un corpus particulier de données.

⁶ L'appui sur les grands écosystèmes numériques peut s'expliquer par l'accès à des ressources (techniques, financières...) mais également par la possibilité de rattacher les services fournis par l'IA générative à des services existants (plugins) ou à des applications propres aux différents écosystèmes. Les partenariats peuvent cependant être analysés comme autant de solutions à disposition des opérateurs installés pour neutraliser des menaces concurrentielles potentielles.

à celle de la possibilité pour les grandes firmes actuelles d'instrumentaliser leur éventuel avantage lié aux données pour étendre leur dominance sur ce nouveau marché. Nous discutons dans une quatrième section les conclusions d'étape qu'il est possible d'en tirer. En d'autres termes, les données constituent-elles des actifs 'essentiels' en matière de marché numérique et, dans le cas de l'IA générative, l'avantage basé sur les données à disposition des Big Techs leur permet-il de conjurer la menace concurrentielle des nouveaux entrants ?

II - De l'existence d'un avantage lié à la donnée dans l'économie des plateformes et des agrégateurs

Considérer le cas des marchés numériques actuels fait d'autant plus sens pour saisir les enjeux du développement de l'IA générative que les arguments en faveur d'une régulation ex ante du secteur se fondent sur leur précédent. Les phénomènes d'ultra-dominance observés sont attribués aux effets de réseaux, aux économies d'échelle et d'envergure, aux stratégies d'acquisitions consolidantes mais également à l'exploitation d'un avantage basé sur les données (Belleflamme et Peitz, 2020). Cet avantage serait déterminant en ce qu'il permettrait une meilleure performance es algorithmes, qu'il s'agisse des algorithmes de prix, de recherche ou d'appariement. Cet avantage serait d'autant plus accru que des algorithmes prédictifs sont en jeu. Une boucle de rétroaction positive pourrait s'initier, faisant que les concurrents, notamment les nouveaux entrants, ne pourraient jamais dès lors devenir aussi efficaces que les opérateurs dominants. Le *tipping*, en d'autres termes le basculement vers une situation de dominance, significative et durable, serait alors inéluctable. S'il est accéléré par des pratiques anticoncurrentielles, les sanctions pourraient être considérées comme inefficaces ; s'il procède des seuls mérites, seules des initiatives réglementaires comme le Règlement sur les marchés numériques (DMA) sont envisageables. C'est contre ce scénario qui pourrait conduire à une 'monopolisation' des marchés émergents de l'IA générative que se construisent de nombreux appels à la régulation. Ces appels portent comme nous le verrons sur plusieurs actifs critiques, notamment les services de cloud, mais aussi les données qui constituent la dimension clé sur laquelle nous nous penchons.

A - Des phénomènes de type *winner takes-all* reliés à des avantages en termes d'accès aux données.

La donnée au cœur du modèle économique des grands acteurs du Web 2.0. Elle peut en effet leur donner un avantage concurrentiel significatif en fonction de l'efficacité de la collecte, des traitements et de l'intégration dans les algorithmes. Cela peut permettre de consolider une position dominante sur le marché pertinent d'origine mais également de l'étendre par effet de levier vers des marchés adjacents. Ces avantages peuvent procéder à la fois des mérites (si la performance est indexée à la captation et au traitement de flux de données massives, diversifiées, renouvelées et vérifiables) mais également de jeux stratégiques autour des données (stratégies d'auto-favoritisme notamment).

- a) Le contrôle de données massives, renouvelées, variées et vérifiées est déterminant dans la performance des grandes entreprises du Web 2.0

Disposer de flux de données massives, diversifiées et de bonne qualité est l'une des variables déterminantes de l'avantage compétitif des grandes firmes en leur permettant d'offrir des services d'autant plus attractifs qu'ils permettent un meilleur ciblage, un meilleur appariement, de meilleures recommandations ou encore une meilleure évaluation de la performance ex post (taux de conversion etc...). Au travers des effets de réseaux, plus l'utilisation du service est massive, plus les flux de données seront élevés et donc meilleure sera l'attractivité du service.

Cependant, si la donnée permet d'initier pour la firme la plus performante un cercle vertueux, elle ne fait pas tout en elle-même. La performance des algorithmes et l'innovativité du service constituent les deux variables clés pour le démarrage. Historiquement, ce ne sont pas les firmes qui avaient le plus de données qui l'ont emporté. La qualité et la performance du service permettent de conquérir des parts de marché significatives dans le cadre de services naissants. Pour autant une question demeure : une firme entrante peut-elle proposer des services aussi performants que ceux d'une entreprise dominante à technologie équivalente ? En d'autres termes, un avantage en termes de données ne permet-il pas jusqu'à un certain point de bénéficier d'une barrière à l'entrée protectrice ?

L'accumulation de données n'est pas pour autant le seul élément de nature à permettre à une firme donnée de basculer vers la dominance et de la consolider. La performance n'est pas liée

qu'aux données massives mais au fait de disposer de flux de données sans cesse renouvelés, variés et de bonne qualité, c'est-à-dire vérifiés. La performance des algorithmes, notamment quand ils sont basés sur l'I.A., dépend étroitement de la capacité des firmes à canaliser ces différents flux, à traiter les données de façon adéquate (format, vérification) pour en assurer la qualité, à tester, expérimenter et développer de nouvelles solutions et enfin à les incorporer dans les services rendus. Pour reprendre les termes de Iansiti et Lakhmani (2020), cela revient à restructurer les process de la firme pour construire une *AI factory* reposant sur des pipe-lines de données, sur des tâches de développement des algorithmes prédictifs entraînés sur cette base, sur des plateformes d'expérimentation pour tester la validité des algorithmes et enfin sur une infrastructure physique d'infonuagique (*cloud*).

Le modèle de la collision concurrentielle issu des travaux de Iansiti et Lakhmani (2020) est particulièrement évocateur quant aux différentiels de capacités concurrentielles entre les firmes réorganisées autour de la donnée et les firmes traditionnelles. Les capacités d'apprentissage et de montée d'échelle des premières sont de nature à leur donner un avantage concurrentiel déterminant en fonction de la taille du marché (capacité de montée en échelle ou scalabilité). Quand les firmes traditionnelles dégagent de la valeur de façon croissante mais de moins en moins rapidement avec la taille du marché (i.e. selon une courbe parabolique), la valeur dégagée par les firmes de la donnée croît de façon exponentielle. En d'autres termes, la première est moins efficace pour des échelles de production faibles (du fait de l'importance des coûts fixes) mais devient de plus en plus efficace au fur et à mesure que le nombre d'utilisateurs croît (le coût marginal étant proche de zéro). A l'inverse, une firme traditionnelle voit ses coûts croître plus vite que ses revenus à partir d'une certaine échelle de production. Il s'agit de phénomènes de déséconomies d'échelle et d'inefficience X qui conduisent à une hausse croissante des coûts marginaux mais également à une capacité décroissante à collecter et à traiter l'information et à améliorer le service sur cette base ou à intégrer ces gains informationnels dans de nouveaux produits⁷.

Il convient de relever que tant l'avantage basé sur les données que le modèle de la collision reposent sur une concurrence par les mérites qui conduit la firme utilisant le mieux les données à évincer du marché ses concurrents moins efficaces. En ce, une telle éviction n'est pas

⁷ Iansiti et Lakhmani (2020, p.158) citent les cas d'Apple et Nokia ou encore d'Airbnb avec des groupes hôteliers tel Marriott.

anticoncurrentielle. Elle débouche cependant sur la constitution d'une position dominante qui peut s'avérer irréversible si hypothèse est faite d'un avantage irrémédiable lié aux données.

L'affaiblissement de la contestabilité de la position de marché ne résulte pas seulement de la captation de flux de données plus complets que ceux à disposition des concurrents pour accroître la performance des algorithmes, elle peut également procéder d'une plus forte capacité à maîtriser son environnement concurrentiel. Il s'agit ici de stratégie de détection avancée des opportunités et menaces concurrentielles reposant sur l'utilisation d'outils de prédiction du présent, c'est-à-dire de *now casting* (Marty et Warin, 2020a).

- b) L'avantage en termes de données peut également permettre d'étendre sa position dominante vers des marchés connexes

Il s'agit d'un second scénario d'éviction lié à un avantage reposant sur les données : l'acquisition ou la préservation de la position dominante ne se situe pas sur un seul et unique marché mais sur plusieurs marchés connexes.

Dans une telle configuration, les flux de données captées sur un marché permettent au travers de l'intégration des services, en termes concurrentiels de stratégies de ventes liées ou de ventes groupées, d'étendre sa position dominante sur des marchés connexes. Il s'agit en d'autres termes de stratégie d'enveloppement (Marty et Pillot, 2021). De telles stratégies ne sont pas inédites dans les marchés numériques. En effet, les stratégies de Microsoft dans les années 1990 ont été sanctionnées sur cette base. Cependant, le rôle de l'avantage lié aux données impose de considérer ces stratégies à nouveaux frais. Les stratégies en cause avaient été sanctionnées sur la base de pratiques pouvant aisément être qualifiées d'anticoncurrentielles qu'il s'agisse des ventes liées ou des entraves à l'interopérabilité des services. Le problème de l'avantage lié aux données est que la nature anticoncurrentielle de l'extension de la position dominante est plus difficile à établir. Plusieurs cas de figure sont à considérer.

Premièrement, l'acquisition des flux de données peut participer des seuls mérites. Une entreprise donnée qui doit une éventuelle position dominante à ses seuls mérites ne fait pas l'objet d'un principe de spécialité⁸. L'utilisation des données dans le cadre d'une activité sur

⁸ Une entreprise qui doit sa position dominante à des droits exclusifs ne peut sur le principe engager une stratégie de diversification en ce qu'elle pourrait s'appuyer sur ses « rentes de monopole » sur le marché initial pour « subventionner » la conquête d'une position dominante sur le marché visé (sur lequel la concurrence s'exerce). C'est le sens du principe de spécialité qui s'appliquait par exemple aux opérateurs du gaz et de l'électricité sur le marché français. Dans son avis 94-A-15 du 10 mai 1994, le Conseil de la concurrence avait établi que la diversification des activités des deux opérateurs considérés étaient conditionnés au titre de l'article 46 de la loi du 8 avril 1946 à l'ouverture à la concurrence de leurs activités historiques (Avis n° 94-A-15 du 10 mai 1994 relatif

un marché donné peut effectivement faciliter la conquête de marchés connexes mais n'est pas anticoncurrentielle en soi.

Deuxièmement, la conquête de marché adjacent peut procéder de stratégies plus discutables en termes de concurrence. Cela peut par exemple procéder de l'instrumentalisation par une firme dominante de sa place pivotale dans un écosystème pour accéder aux données générées ou collectées par ses complémentateurs, en d'autres termes par les entreprises utilisatrices de ses services d'intermédiation (Marty, 2019).

La procédure ouverte par la Commission européenne à l'automne 2020 contre Amazon⁹, qui s'est achevée par une procédure d'engagements en décembre 2022¹⁰, correspondait partiellement à ce schéma. Il était notamment fait grief à l'entreprise d'imposer des conditions de partage mais également de diffusion de l'information qui la plaçaient dans une position privilégiée par rapport à ses propres complémentateurs. L'opérateur de plateforme était dans le cas d'espèce également verticalement intégré. En d'autres termes, il pouvait vendre ses propres produits sur sa plateforme et ainsi agir comme un concurrent de ses propres complémentateurs (i.e. les vendeurs indépendants). Cependant, la situation de contrôleur d'accès dont jouit l'opérateur pivot (les vendeurs indépendants sont obligés d'utiliser ses services pour accéder à

à une demande d'avis sur les problèmes soulevés par la diversification des activités d'E.D. F et de G.D.F. au regard de la concurrence). Le cas des activités électriques et gazières en voie de libéralisation a donné lieu à d'autres problèmes liés à la capacité des opérateurs historiques à instrumentaliser leur position de marché spécifique pour jouir d'un avantage déterminant vis-à-vis de leurs concurrents. Ce parallèle est d'autant plus pertinent pour notre propos qu'il tient ici à la maîtrise des données. L'Autorité de la concurrence a eu à traiter deux cas dans lesquels des concurrents sur les marchés historiques des deux anciens titulaires de droits exclusifs avaient un avantage indu sur eux dans la mesure où eux seuls disposaient des données sur les clients demeurés aux tarifs de vente régulés. Ils jouissaient possiblement d'un avantage déterminant lié aux données : une meilleure capacité d'anticipation des clients qui étaient susceptibles de basculer vers des offres de marché. Ils avaient de ce fait un avantage de coûts sur leurs concurrents. Ils pouvaient en effet micro-cibler leurs investissements marketing et espérer un retour sur investissement plus favorable que celui de concurrents qui ne pouvaient être aussi précis dans leurs prospections commerciales. L'accès aux données était ici plus aisé à exiger que dans le cas des activités des plateformes numériques en ce que les données ne pouvaient être accessibles qu'à la firme titulaire de droits exclusifs. Il n'était pas pour autant aisé de les qualifier de facilités essentielles en ce qu'accéder à celles-ci n'étaient pas indispensable pour exercer une activité sur le marché de la fourniture d'électricité ou de gaz dans le cadre d'offres à prix de marché. Cependant, l'accès discriminatoire à des données auxquelles seuls les opérateurs historiques pouvaient accéder (pour une raison autre que les mérites) était effectivement de nature à leur donner un avantage « non-reproductible » par rapport à un concurrent. Pour le cas de l'électricité, il convient de se reporter à la décision n° 22-D-06 du 22 février 2022 relative à des pratiques mises en œuvre par la société EDF dans le secteur de l'électricité. Pour le cas du gaz, il convient de considérer la décision 14-MC-02 du 09 septembre 2014 relative à une demande de mesures conservatoires présentée par la société Direct Energie dans les secteurs du gaz et de l'électricité. Les pratiques en cause visaient ici des marchés pertinents provisoirement distincts (en l'espèce les marchés de fournitures à tarifs régulés et en offres de marchés). Elles peuvent également être mise en œuvre sur des marchés réellement distincts et donc faire écho à la question de la diversification. Le cas de l'utilisation des données liées aux droits exclusifs dans le cadre de la diversification des activités vers le photovoltaïque peut en constituer un exemple (Autorité de la concurrence, décision 13-D-20 du 17 décembre 2013, relative à des pratiques mises en œuvre par EDF dans le secteur des services destinés à la production d'électricité photovoltaïque).

⁹ https://ec.europa.eu/commission/presscorner/detail/fr/ip_20_2077

¹⁰ <https://competition-cases.ec.europa.eu/cases/AT.40462>

une part très importante de leurs clients) fait qu'il en mesure d'imposer des conditions déséquilibrées d'accès à l'information. L'accès asymétrique à certaines données clients faisait que chaque vendeur indépendant n'avait qu'une connaissance partielle de sa propre situation alors que la plateforme avait une information complète non seulement sur son activité de commercialisation mais également sur l'ensemble des activités de ses concurrents. Il ne s'agit pas alors d'utiliser un avantage lié aux données pour évincer ses concurrents d'un même marché pertinent mais pour étendre sa position dominante vers un marchés connexe.

Il est donc possible d'utiliser une position pivotale dans un écosystème numérique pour acquérir un avantage informationnel par rapport aux autres membres de l'écosystème qui sont à la fois des clients et des concurrents. Le processus que nous avons décrit dans le domaine des plateformes de commerce en ligne peut également être observé dans celui des services d'infonuagique. De façon plus générale, plus un écosystème numérique est incontournable pour les firmes utilisatrices, plus la firme pivotale est en mesure d'imposer des conditions d'accès lui permettant d'accéder à des données sur ces dernières, données susceptibles de lui offrir un avantage concurrentiel.

Troisièmement, la construction d'un avantage lié aux données peut se faire au détriment non pas des complémentateurs mais à celui des consommateurs du service. Il s'agit ici de disposer de plus de données que ses concurrents sur un versant de marché distinct, celui des services publicitaires par exemple. Ce faisant le service rendu sur le marché publicitaire sera plus valorisable que celui rendu par les concurrents en ce qu'il s'appuie sur des flux de données plus massifs, plus souvent actualisés, plus précis car provenant du croisement de plusieurs sources etc...

Ces différents points peuvent être approfondis selon deux angles.

Le premier angle tient à la caractérisation d'une nouvelle forme possible d'abus de position dominante, celui d'un abus d'exploitation prenant la forme d'une collecte excessive de données. La procédure ouverte par le Bundeskartellamt contre Facebook peut fournir un premier exemple d'une telle théorie du dommage (Kerber, 2022). Cette nouvelle approche transparaît également dans le nouveau communiqué de procédure de la Commission européenne relatif à la définition des marchés pertinents¹¹. Dans des marchés parfois caractérisés par des modèles de prix zéro, un abus d'exploitation passe par d'autres formes qu'une hausse des prix qui pourrait être saisie

¹¹ Commission européenne, *Communication de la Commission sur la définition du marché en cause aux fins du droit de la concurrence de l'Union*, C(2023) 6789 final, 8 février 2024.

au travers d'un test SSNIP. Elle peut se traduire par d'autres phénomènes comme une extraction excessive de données ou encore comme une dégradation de la qualité du service rendu (en termes de respect des données personnelles par exemple). Il s'agit alors de mettre en œuvre d'autres types de tests comme le SNDDQ.

Le second angle peut tenir à l'instrumentalisation de la position financière des firmes dominantes au cœur des grands écosystèmes numériques pour acquérir des données au-delà de leur domaine d'activité.

Une première voie tient à la politique d'acquisition de firmes agissant dans des marchés distincts et générant des flux de données qui peuvent être stratégiques pour accroître la performance de leurs outils prédictifs, pour s'intégrer verticalement dans un secteur donné ou encore pour préempter des données qui font figure d'actifs qui auraient été déterminants dans la possibilité d'entrée de nouveaux concurrents. La prise en compte croissante des enjeux liés aux données dans le contrôle des concentrations en témoigne que cela soit en matière d'appréciation des risques concurrentiels ou en matière de définition des remèdes. La possibilité d'étendre sa position dominante d'un marché à l'autre à partir de la captation et de l'utilisation de données ou encore de sa consolidation au travers de la préemption de données indispensables aux concurrents a été importante dans de nombreux dossiers de concentrations. Nous pourrions citer les cas des remèdes imposés dans le cadre de l'acquisition de FitBit par Google¹² ou encore de la renonciation d'Amazon à acquérir iRobot¹³.

Une seconde voie d'utilisation de la puissance financière pour consolider un *data-based advantage* tient à l'achat direct de données auprès de tiers. Il peut s'agir de partenariats avec des opérateurs du marché des cartes bleues ou de chaînes de supermarché¹⁴ ou d'achats auprès de courtier en données. Dès 2014 la FTC avait relevé les risques liés à l'activité de ces derniers dans le domaine notamment de la protection des données personnelles (Marty, 2023b). Ils peuvent également induire des risques concurrentiels dès lors que les informations qu'ils détiennent sur les consommateurs pourraient être préemptées par les opérateurs dominants du monde de la Tech pour accroître leurs capacités prédictives et pour priver les entrants potentiels

¹² Voir la décision de la Commission du 17 décembre 2020 (COMP/M.9660). L'autorisation de l'opération de concentration a été conditionnée (pour limiter ses éventuels effets verticaux) à la constitution de silos de données.

¹³https://france.representation.ec.europa.eu/informations/declaration-de-mme-vestager-vice-presidente-executive-sur-lannonce-faite-par-amazon-et-irobot-de-2024-01-30_fr

¹⁴ Voir par exemple le brevet déposé par Google en 2012 : <https://patents.google.com/patent/US10255598B1/en>

de données qui pourraient leur permettre de compenser leur désavantage initial. Ce dernier est pour le moins paradoxal et suppose une analyse plus détaillée.

Les données commercialisées par certains opérateurs – même des fournisseurs d’antivirus et de services garantissant une meilleure confidentialité¹⁵ – peuvent être acquises par des courtiers en données. Ces derniers peuvent les traiter pour en augmenter la qualité et les croiser avec des données provenant d’autres sources pour en accroître la qualité et en améliorer la granularité. La commercialisation de ces données peut intéresser les grands opérateurs mais également les nouveaux entrants. Les seconds peuvent d’autant plus l’être que cela peut leur permettre de compenser leur désavantage initial. Cette possibilité même conduit à s’interroger sur l’opportunité même de considérer les données comme des barrières à l’entrée pour le web 2.0 lui-même.

B - La donnée était-elle la seule source du basculement vers des situations d’ultra dominance, de renforcement de celles-ci et d’extension à d’autres marchés ?

Il convient de considérer les facteurs qui peuvent conduire à des phénomènes de consolidation au-delà des données elles-mêmes avant de s’attacher à leur conformité avec les principes d’une concurrence par les mérites.

1) Des facteurs de consolidation hors des données

Les données sont-elles en effet des facilités essentielles ? Peuvent-elles être considérées comme des barrières à l’entrée ? Est-il en effet même possible de conquérir une position de force sur un marché numérique sans disposer initialement de données ou alors en acquérant des données disponibles auprès de tiers ? Qualifier les données de barrières à l’entrée ne va pas de soi et ce pour plusieurs raisons.

D’abord, elles peuvent être, comme vu *supra*, acquises auprès de tiers. Deuxièmement, elles ne sont pas purement ‘appropriables’ par un opérateur donné. En effet, une partie des données relatives à chaque consommateur peut être accessibles à différentes firmes dès lors qu’il ne s’engage pas dans un mono-hébergement absolu. Ensuite, si les données apportent un avantage à la firme capable de les collecter, de les traiter et de les valoriser pour entraîner les algorithmes,

¹⁵ Voir notamment le billet du Business Blog de la FTC en date du 9 janvier 2024 : “What goes on in the shadows: FTC action against data broker sheds light on unfair and deceptive sale of consumer location data”
<https://www.ftc.gov/business-guidance/blog/2024/01/what-goes-shadows-ftc-action-against-data-broker-sheds-light-unfair-deceptive-sale-consumer-location>

il n'est pas exclu qu'une entreprise puisse efficacement entrer sans disposer de données massives ou en disposer en propre. Une entreprise peut les acquérir par des moyens discutables en les récoltant sans autorisation sur Internet¹⁶ ; elle peut également comme vu *supra* les acquérir auprès de courtiers. Qui plus est des entrées ont été possibles et couronnées de succès sans disposer de données initiales (ni d'une vaste base d'utilisateurs, ni de l'intégration dans un écosystème) comme le montre le cas de Zoom à l'occasion du premier confinement de 2020. La supériorité du produit ou d'une stratégie de marché peut permettre de compenser l'absence initiale de données¹⁷.

Ces éléments peuvent conduire à nuancer l'importance de l'avantage lié aux données dans la croissance des grands opérateurs de l'économie numérique. Les phénomènes de basculement vers des situations d'ultra-dominance ne sont pas liés exclusivement à des avantages basés sur les données. Ils résultent de facteurs bien distincts.

Les premiers facteurs sont attribuables aux structures de production même avec de forts coûts fixes et des coûts marginaux proches de zéro¹⁸. Ils sont également attribuables aux économies d'envergure et pas seulement aux économies d'échelle. Ils peuvent de la même façon tenir aux effets de réseaux directs et indirects. En d'autres termes, la taille et la diversification mêmes des écosystèmes peut conduire mécaniquement à des effets de type *winner takes all*. Cette dynamique peut s'enclencher même sans avantage direct lié aux données et dans une certaine mesure sans supériorité technique absolue sur les concurrents (Pillot, 2009).

Les seconds facteurs conduisent à relâcher la première hypothèse selon laquelle la taille des opérateurs a des effets qui l'emportent sur l'innovativité. Le poids sur le marché dans des industries avec les caractéristiques techniques décrites ci-avant constitue un avantage de poids par rapport aux concurrents. Cependant l'histoire économique du secteur nous montre que de nouveaux entrants ont pu compenser ce handicap pour s'imposer contre des opérateurs installés. Il y a donc un élément clé à considérer qui tient à la capacité à innover et à orienter les apprentissages de la firme vers cette finalité.

¹⁶ Un contentieux a été engagé par X à l'encontre d'Open AI en janvier 2024 pour extraction de données sans autorisation tout comme le New York Times dans le cadre d'une action au civil (Case 1:23-cv-11195) enregistrée le 27 décembre 2023.

¹⁷ L'entrée de Google sur le marché des moteurs de recherche ou encore celle de TikTok sur celui des réseaux sociaux se sont faites sans base initiale d'utilisateurs comparables à celles des concurrents. La conquête du marché s'est faite dans les deux cas à partir de l'adéquation de l'algorithme aux besoins des utilisateurs.

¹⁸ Cette caractéristique doit être écartée pour l'IA générative du fait de ses coûts de fonctionnement.

Cette culture de l'innovation peut être à l'origine de disruptions significatives et permettre donc de compenser des désavantages en termes de position de marché et de contrôle de flux de données mais il n'est pas acquis qu'elle soit un apanage des nouveaux entrants et que les firmes déjà dominantes ne soient pas les mieux placées à cette aune. Ce facteur risque donc d'être bien moins un facteur de disruption qu'un facteur de consolidation. En effet, quelques traits saillants peuvent être dégagés.

Premièrement, toutes les entreprises ne sont pas égales dans la construction et dans le maintien d'une culture d'innovation. De surcroît, cette culture peut renforcer les différences de potentiels entre les firmes : plus une entreprise est innovante, plus elle a de chances d'innover encore dans le futur.

Deuxièmement, les entreprises dominantes du numérique ont des capacités d'absorption de l'innovation venant de tiers dont ne disposent pas les entreprises de l'économie traditionnelle. Les délais d'intégration des services innovants sont particulièrement courts dans la mesure où il n'y a pas de séparation entre les activités de recherche-développement et de production. Ensuite, le fonctionnement même des entreprises est orienté vers la captation et les développements ultérieurs des innovations potentielles. Ces caractéristiques font écho aux capacités financières et techniques des firmes déjà citées *supra*. Elles ont les capacités mais également les incitations internes pour investir dans la recherche-fondamentale, dans les talents¹⁹ etc... Elles ont des actionnaires dont la patience autorise des délais de récupération des investissements très longs et même des investissements dans des activités hors des marchés historiques dont les perspectives en termes de couple risque/retour ne sont pas attractifs. Elles ont également les capacités financières adéquates pour racheter les firmes innovantes.

Troisièmement, comme nous l'avons déjà relevé *supra*, les firmes du numérique ont des processus très courts de mise sur le marché des innovations et d'intégration de celles-ci à leurs services. A ce titre les gains potentiels liés aux rachats de firmes innovantes dans une perspective d'acquisition consolidante sont bien supérieurs et bien plus rapides que ceux qu'ils peuvent l'être dans des secteurs d'activité traditionnels. Cela est par exemple particulièrement le cas comparé au secteur de la pharmacie. Le parallèle est intéressant en ce que la notion d'acquisition tueuse a été forgée dans ce domaine et qu'au-delà même du débat entre finalités tueuses ou

¹⁹ Des stratégies de préemption de compétences rares et d'augmentation des coûts des rivaux peuvent être prises en considération. Il convient cependant de noter que se faire concurrence pour attirer les talents dans un secteur en tension n'a rien d'anticoncurrentiel et procède d'une concurrence par les mérites. Cependant, il peut ne pas en aller ainsi dès lors qu'un jeu sur les clauses de non concurrence est mis en œuvre.

consolidatrices des acquisitions, il présente l'intérêt d'ouvrir sur la question de l'accompagnement de l'accès des innovations au marché. Dans le domaine des biotechs, le modèle économique des start-ups peut être d'être racheté avant même d'accéder au marché. Les raisons en tiennent aux contraintes réglementaires de mise des produits sur le marché, aux difficultés propres du passage à la production. Non seulement ces contraintes sont moindres dans le domaine du numérique mais les entreprises dominantes du secteur se caractérisent par des capacités particulièrement élevées en termes d'investissements et de scalabilité. Dernier point enfin, l'utilité ou encore la possibilité de valorisation d'une innovation va étroitement dépendre de son intégration dans un ensemble de services rendus par un écosystème donné.

2) Des avantages qu'il est possible d'interroger au point de vue concurrentiel

L'ensemble de ces facteurs fait que des variables autres que les données peuvent expliquer le succès sur les marchés numériques... mais également que les opérateurs les plus puissants sont les mieux placés pour s'imposer. Cependant nous n'avons considéré ici que les facteurs tenant aux mérites. Il est également possible de considérer que l'avantage des contrôleurs d'accès peut également résulter de stratégies d'entraves dans l'accès des tiers aux données ou encore de distorsions dans la présentation des informations.

Commençons par le second point. La capacité des firmes contrôleuses d'accès à verrouiller ou à biaiser l'accès aux informations des consommateurs peut leur permettre à pousser certains services vers les consommateurs ou à entraver leur accès à d'autres. Le jeu sur les architectures de choix biaisées (*dark patterns*) tel qu'il a été illustré par Ezrachi et Stucke (2020) montre de façon heuristique comment un opérateur pivot peut pousser ses propres innovations au détriment de celles de ses complémenteurs même si elles sont moins intéressantes pour le consommateur en jouant sur l'information qui lui est transmise. Le jeu sur les données disponibles peut entraver l'adoption de services concurrents.

Il peut également entraver leur développement même dès lors que l'accès des concurrents à des données générées par le fonctionnement de l'écosystème est refusé, voire dégradé. En d'autres termes, des stratégies d'obstruction anticoncurrentielles peuvent être mises en œuvre en matière d'accès aux données. Elles peuvent passer par des refus d'accès « absolus », ou par des refus d'accès « relatifs » qui peuvent porter sur des entraves tenant à aux difficultés de portabilité des données ou à des contraintes techniques (ou financières) en matière d'interopérabilité. Une autre pratique anticoncurrentielle possible pourrait tenir à des discriminations en matière d'accès aux données que cela soit en matière tarifaire ou technique. La firme dominante

pourrait, que cela soit sur son marché pertinent d'origine ou sur un marché connexe, favoriser ses propres activités au détriment de celles de ses concurrents au travers d'un meilleur accès aux données ou d'un accès moins coûteux. Dans de telles situations, les différences de traitement entre les services internes et ceux des concurrents conduisent à des pratiques d'auto-favoritisme se traduisant sinon par l'éviction des rivaux du moins par l'augmentation de leurs coûts. L'Autorité de la concurrence a considéré ces questions dans le cadre de son enquête sectorielle sur le *cloud*.

Dans le domaine de l'application des règles de concurrence, la prise en considération de l'accès aux données prend une place croissante tant dans le domaine de la sanction des pratiques anticoncurrentielles que dans celui du contrôle des concentrations. Dans le domaine de l'antitrust, des remèdes en termes d'accès aux données peuvent être prononcés ou rendus obligatoires, lesquels peuvent aller jusqu'à des partages de données. Dans le second domaine, celui des concentrations, une attention est par exemple portée sur l'effet de levier que pourrait constituer la prise de contrôle des données de la cible. Celles-ci seront-elles exclusivement disponibles pour la firme acquéreuse ? Cette dernière pourra-t-elle gagner un avantage déterminant sur ses concurrents par la mise en commun des données, etc... ? Les cas des opérations Google / FiBit et Amazon / iRobot témoignent de ces problématiques. Dans un cas la concentration a été acceptée sous conditions de mise en place de silos de données ; dans l'autre les deux parties ont préféré retirer leur projet. De façon plus structurelle encore les nouvelles lignes directrices en matière de contrôle des opérations de concentration publiées par la FTC et le DoJ en décembre 2023 considèrent qu'il convient d'évaluer les effets possibles d'une concentration au regard des capacités de porter préjudice à la concurrence en refusant à des firmes rivales l'accès à des données détenues par la firme acquise. Il va de même pour les *merger assessment guidelines* de la CMA britannique dans le §7.19^e (CMA, 2021). Des acquisitions peuvent porter préjudice à la concurrence dès lors qu'elles permettraient d'accéder à des données susceptibles d'informer l'acquéreur sur les activités et les décisions de firmes partenaires de l'entité acquises, firmes qui seraient déjà ses concurrentes ou qui pourraient le devenir au travers de diversifications futures²⁰. Au-delà même du domaine des règles de concurrence strictement défini, le Règlement sur les Marchés Numériques (DMA), fait peser

²⁰ Un exemple dans le domaine de la réalité virtuelle peut être apporté avec le cas de la procédure qui avait été engagée par la FTC dans le cadre de l'acquisition par Meta de l'application Within (voir la plainte additionnelle de la FTC sur ce projet de concentration déposée en 2022 (n°605837)). Notons que l'acquisition a été finalement autorisée en février 2023 (<https://www.ftc.gov/legal-library/browse/cases-proceedings/221-0040-metazuckerbergwithin-matter>).

des obligations sur les contrôleurs d'accès en termes d'interopérabilité des services et de portabilité.

Au final, si les données ont joué un rôle important dans la conquête et à la dominance de certains marchés numériques par les opérateurs principaux du Web 2.0, il apparaît non seulement qu'elles ne constituent pas à elles seules une condition sine qua non de développement sur le marché et qu'un ensemble de mesures correctives visant ces mêmes opérateurs critiques se traduit par des garanties additionnelles en termes de partage de données, de portabilité des données ou encore d'interopérabilité des services qui sont susceptibles de favoriser le développement des concurrents en réduisant les barrières à l'entrée. Le développement de l'IA générative est-il appelé à reproduire les mêmes phénomènes que ceux connus dans le Web 2.0 ou les spécificités de cette technologie, conjuguée aux effets des mesures détaillées supra, va-t-il faire qu'une concurrence pourra efficacement se développer malgré des positions de départ différentes en matière de contrôle des données ?

III. Dans le domaine de l'IA, et plus spécifiquement de l'IA générative, la question de l'essentialité de l'accès aux données

Avant de s'attacher à l'existence éventuelle d'une barrière à l'entrée ou à l'expansion dans le domaine de l'IA générative via les données contrôlées par les Big Techs, il convient de se pencher sur le rôle des données dans le développement des modèles de langages. Il convient à cette fin de reprendre la distinction faite en introduction entre les modèles de langages fondamentaux et les modèles développés sur leur base pour répondre à des besoins spécifiques (*fine tuning* ou *knowledge distillation*). Ces traitements aval visent à aligner les « réponses » avec les attentes des utilisateurs. Il s'agit également au travers du fine tuning de réaliser des entraînements complémentaires en laissant une partie du modèle en l'état. La partie du modèle qui n'est pas « gelée » reçoit pour sa part des données complémentaires. Cela permet à partir d'un modèle général de faire des développements aval avec des données et des entraînements spécifiques pour permettre une meilleure performance dans un cadre d'utilisation plus précis. Ainsi, de façon plus générale, des développements additionnels sont apportés à des modèles généraux visent à répondre de façon plus précise à des besoins plus spécifiques en utilisant des données originales tout en nécessitant des capacités de calcul moindres. Ces développements aval peuvent prendre la forme de plug-ins ou de services spécifiques rendus aux utilisateurs finaux. Ces modèles spécifiques peuvent être développés sur la base de modèles de langages fondamentaux en open-source. L'open source est une ressource de nature à réduire les barrières

à l'entrée tant en matière de modèles de langage que de données disponibles. Il pourrait rendre possible le développement d'une concurrence aval.

Nous considérons cette hypothèse dans une première sous-section (A) avant de considérer certaines limites qui pourraient être de nature à consolider les positions dominantes actuelles (B).

A - La disruption des opérateurs du Web 2.0 et la possibilité d'avoir de développer des modèles spécifiques à partir de modèles de langages fondamentaux

De nombreux facteurs peuvent être de nature à réduire les barrières à l'entrée. Cela peut concerner des données en accès libre, les règles d'accès aux services d'informatique en nuage, etc... Pour autant, les règles propres au développement de l'IA générative peuvent rendre ces actifs moins nécessaires que par le passé.

1) Les nouveaux entrants bénéficient déjà des obligations pesant sur les opérateurs dominants du web 2.0

Comme nous l'avons vu ci-avant, les nouveaux entrants peuvent bénéficier des remèdes concurrentiels précédemment imposés ou des régulations mises en place. Cela porte non seulement sur l'accès aux données (portabilité) mais également sur l'accès aux services cloud. Ces remèdes apparaissent comme des conditions essentielles pour garantir la libre concurrence de ce marché émergents et pour prévenir les risques de verrouillage par les opérateurs dominants sur des segments de marché connexes.

La prévention de ce risque est leitmotiv des régulateurs de la concurrence : ne pas se laisser surprendre par la stratégie des Big Techs dans le cadre de cette nouvelle vague d'innovations²¹. Il s'agit d'éviter de réitérer les erreurs des décennies précédentes et de prendre en considération un risque additionnel au regard du caractère critique de l'IA qui est une technologie d'usage général (Acemoglu and Lensman, 2023). En effet, dans la première décennie des années 2000, une application trop précautionneuse des règles de concurrence aurait laissé la possibilité à

²¹ Cette crainte était notamment exprimée dans le rapport Furman de 2019 : comment éviter que les entreprises déjà dominantes sur un segment de marché donné soient en mesure d'instrumentaliser l'avantage lié aux données pour consolider leurs positions sur le marché considéré mais également de l'étendre à des marchés connexes par effet de levier (Furman et al., 2019, p.4).

certaines firmes de « monopoliser » les marchés émergents²² au travers de la prise de contrôle de complémentaires stratégiques, de la captation de flux de données etc.... Les remèdes concurrentiels ex post dans le cadre de la sanction des pratiques anticoncurrentielles ou ex ante dans le cadre du contrôle des concentrations n'en corrigent ou n'en préviennent que très mal les effets. Les régulateurs pourraient donc être tentés d'intervenir pour prévenir le fait que les opérateurs dominants d'aujourd'hui soient ceux de demain. Ainsi les nouveaux entrants peuvent bénéficier des remèdes concurrentiels mis en œuvre et des réglementations mises en place notamment en termes d'interopérabilité et de portabilité des données.

Comme le relèvent Manne et Auer (2024), les entreprises en position de pointe dans le domaine de l'IA générative ne sont pas les grandes entreprises de plateformes qui pourtant bénéficiaient d'un avantage significatif en matière d'accès aux données. Non seulement des start-ups peuvent naturellement être plus agiles pour innover que des grands groupes mais de telles entrées dans un domaine pour lequel l'avantage lié est théoriquement central peut conduire à relativiser l'hypothèse d'un *data-based advantage*.

Premièrement, il n'y a pas une relation monotone entre la masse de donnée à disposition et la performance. Il existe des rendements décroissants en matière de contrôle de données (Tucker, 2019). Il suffit donc d'avoir suffisamment de données pour être performant et non pas de contrôler le plus de données possible²³. Deuxièmement, les progrès réalisés en matière de développement logiciel peuvent abaisser le seuil de données nécessaires (Schrepel et Pentland, 2023). Troisièmement, il peut être possible d'utiliser comme substituts des données synthétiques (Koperniak, 2017).

- 2) Un avantage concurrentiel répondant moins à la masse de données qu'à la qualité et à la spécificité des données utilisées

Ainsi, plusieurs hypothèses peuvent venir nuancer les avantages liés aux données dans le cas de l'IA générative (voir notamment Manne et Auer, 2024). Les trois premières hypothèses peuvent concerner la concurrence sur le marché des modèles fondamentaux et les deux dernières sur le marché des modèles spécialisés. Il est tout d'abord possible de considérer que

²² Nous utilisons le terme monopoliser dans le sens du Sherman Act, c'est-à-dire dans le sens acquérir, maintenir et étendre leurs positions dominantes sur une autre base que celles des purs mérites.

²³ En matière de modèles de langages, les rendements croissants peuvent apparaître plus certainement pour les modèles de langage fondamentaux que pour les modèles fine-tunés (Schrepel, 2023).

c'est la qualité du modèle qui fait la différence bien plus qu'un avantage éventuel lié aux données. En outre, les données d'entraînement peuvent être fournies par des bases publiques. Enfin, il n'existe pas d'avantage lié à la quantité même de données. Il existe un seuil à partir duquel les performances ne s'améliorent plus. Les deux dernières hypothèses concernent la distinction entre langages fondamentaux et langages fine-tunés. Les seconds peuvent être développés à partir des premiers sans nécessiter autant de données et d'infrastructures pour le développement. Ils peuvent en outre permettre de dégager une valeur additionnelle dès lors que les données en question sont spécifiques et présentent une valeur particulière pour les utilisateurs.

Un certain nombre de paramètres peuvent en effet conduire à nuancer l'importance des barrières à l'entrée liée aux données²⁴. Il est possible de développer des modèles fine-tunés sur la base de modèles de langages fondamentaux déjà entraînés (et il est possible de le faire à partir d'équipements moins spécifiques et moins coûteux). Ces nouvelles versions peuvent créer d'autant plus de valeur ajoutée que les données sur lesquelles elles sont développées sont spécialisées (dans le cadre d'un champ particulier) et qu'elles ne sont pas contrôlées par les Big Techs. Au-delà de ces données spécifiques, le déploiement de modèles peut également reposer sur des données disponibles librement (voir The Pile²⁵ ou encore GSM8K²⁶) ou encore sur des données synthétiques. De la même façon, plus un algorithme ou un processus d'apprentissage est performant, moins son besoin en données sera important. Comme le note Manne et Auer (2024) : “[...] the real challenge to create cutting edge AI is not so much in collecting data, but rather in creating innovative AI-training processes and architectures”. Enfin, l'accumulation des données en elle-même ne crée pas de valeur, même si elle contribue à une meilleure capacité prédictive. C'est la capacité de la firme à développer des algorithmes qui lui permettent de transformer un éventuel avantage dans ce domaine à un avantage de marché (Tucker, 2020).

Ces dimensions sont en effet essentielles pour saisir les paramètres de la concurrence dans le domaine de l'IA générative. Trois plans doivent en effet être distingués. Le premier est le plan horizontal amont : il concerne la concurrence entre développeurs de langages fondamentaux. Le deuxième est le plan horizontal aval : il porte sur la concurrence entre développeurs de

²⁴ Au sein des données, il convient de distinguer les données utilisées dans le cadre du développement des modèles et celles qui sont utilisées dans le cadre de leur déploiement (Gans, 2024). Les premières sont déjà produites. Elles sont nécessaires avant l'entrée sur le marché. Les secondes sont générées pendant le déploiement au travers des réponses aux questions posées par les usagers. Elles sont spécifiques et peuvent l'amélioration des qualités prédictives des algorithmes (Hagiu et Wright, 2023).

²⁵ <https://pile.eleuther.ai/>

²⁶ <https://paperswithcode.com/dataset/gsm8k>

modèles fine tunés. Le troisième niveau est le plan vertical. Il est particulièrement important pour notre propos. Il traite des relations verticales entre les grands opérateurs de l'économie numérique et les entreprises développant des modèles de langages fondamentaux et celles développant des modèles de fines tunés.

La valeur des résultats d'une IA générative dans un domaine d'application précis va dépendre étroitement de la qualité et de l'unicité des bases de données proposées. C'est par exemple le cas dans le domaine financier. Des entreprises qui sont en mesure de développer des modèles fine-tunés dans de tels domaines peuvent avoir un avantage déterminant même vis-à-vis des Big Techs.

Les firmes qui se développent sur le marché de l'IA générative peuvent à terme devenir des concurrentes des Big Tech traditionnelles et donc contribuer, bien plus que l'action des régulateurs de la concurrence, à éroder leur puissance de marché. Elles peuvent tout d'abord se substituer à certains de leurs services clés, comme la recherche en ligne par exemple. Elles peuvent ensuite s'intégrer verticalement pour se doter de leurs infrastructures propres (en termes de cloud et de processeurs) voire développer des solutions technologiques propriétaires²⁷. Elles peuvent enfin devenir elles-mêmes des pivots d'écosystèmes. Au-delà des firmes intervenant dans le secteur du développement des modèles fondamentaux, d'autres firmes peuvent mettre à profit des avantages spécifiques (par exemple en matière de contrôle des données) pour conquérir des niches de marché particulièrement profitables.

B - La persistance de goulets d'étranglement et de possibilités d'exercer des pouvoirs de marché abusifs

Ce scénario est-il pour autant assuré ? Cela peut se discuter selon deux angles. Le premier angle relève des données elles-mêmes. Ont-elles perdu leur statut d'actifs « quasi-essentiels » ? Le second angle relève de la concurrence elle-même. N'existe-t-il pas d'autres barrières à l'entrée susceptibles de mettre en question l'hypothèse de disruption ?

1) De la persistance des barrières à l'entrée liées aux données

Il convient de discuter l'hypothèse de la réduction des barrières à l'entrée liées aux données avec l'IA générative. Deux questions peuvent en effet être posées. Premièrement, peut-on les

²⁷ Il est possible de dresser ici un continuum entre les approches en termes d'échelle des investissements jusqu'à des notions de co-développement à la von Hippel (1986).

obtenir par moissonnage (*scraping*)²⁸ ou par des sources de données publiques et le cas échéant ces ressources sont-elles d'une qualité équivalente à celles à disposition des Big Techs ? Deuxièmement, même si on a des données spécifiques peut-on partir *from the scratch*, à partir de données disponibles en *open source* ou a-t-on besoin des données de ces opérateurs ?

Même si des données sont disponibles en open source, un enjeu demeure : celui de leur qualité. De la même façon, plus les données recherchées sont spécifiques à un domaine précis, plus leurs coûts d'acquisition sera élevé. De la même façon la qualité initiale des données peut beaucoup jouer sur les coûts de traitements. Enfin, il convient de distinguer entre les données d'entraînement et les données de contrôle. Les secondes sont particulièrement importantes dans la mesure où les modèles nécessitent de longues expérimentations. Celles-ci peuvent constituer des actifs plus critiques que les données d'entraînement.

La disponibilité des données constitue un avantage en termes concurrentiels en ce qu'elle favorise une relative concurrence à égalité des armes. Cependant, certaines difficultés peuvent être relevées. Premièrement, les données disponibles en open source, tout comme celles qui peuvent être disponibles du fait d'obligations relevant de remèdes concurrentiels (accès aux données...) peuvent poser problèmes en matière de compatibilité avec les règles de protection des données personnelles. Deuxièmement, des données de mauvaise qualité peuvent donner lieu à des résultats particulièrement problématiques en termes collectifs qu'il s'agisse de risques de manipulations (via les données) ou de risques de résultats de très mauvaise qualité (hallucinations numériques) résultant d'un contrôle et d'une curation insuffisants²⁹.

Il convient également de noter que les avantages techniques à disposition des grands groupes portent également sur leurs capacités à traiter efficacement les données. Ce travail sur les données suppose des compétences particulières à la fois en termes techniques et légaux mais également en termes de coûts. Les grands groupes ont une pratique développée et efficace d'externalisation des tâches qui peut les mettre en mesure d'avoir un très net avantage sur leurs concurrents en matière de qualité, de vitesse et de coûts³⁰.

²⁸ Certaines Big Techs peuvent pour des raisons potentiellement légitimes restreindre les capacités des concurrents à collecter des données sur leur écosystème. Un impact direct peut alors tenir à la préemption de ces données au détriment de tiers voulant entraîner ou superviser leurs propres algorithmes.

²⁹ Le problème revient ici à l'adage *garbage in / garbage out*.

³⁰ L'avantage en matière d'externalisation peut tenir non seulement à la qualité des procédures internes mises en œuvre (division des tâches, consolidation des résultats) mais également à l'exercice d'un pouvoir de marché aval. La capacité de négociation sur les prix n'est pas la même selon qu'il s'agisse d'un nouvel entrant ou d'une Big Tech.

Des règles en termes de partage de données peuvent également avoir des effets anticoncurrentiels comme en témoignent par exemple les multiples contentieux liés aux collusions algorithmiques au travers de schémas en étoile, comme le montre le cas Real Page aux Etats-Unis³¹. Notons par ailleurs que la littérature sur l'IA générative souligne des risques de possibles collusions mais celles-ci peuvent tout à la fois procéder de données d'entraînement identiques comme dans les configurations mentionnées *supra* que des échanges avec les applications elles-mêmes. (Motwani et al., 2024).

2) De la persistance des barrières à l'entrée liée aux stratégies et aux structures de marché

Il convient de considérer la concurrence à deux niveaux. Le premier correspond aux modèles de langages fondamentaux, le second à celui des modèles dérivés.

Le secteur des modèles fondamentaux peut se caractériser rapidement par des phénomènes de forte concentration au travers de caractéristiques techniques le rendant particulièrement sujet à des phénomènes de *tipping*. L'Autorité de la concurrence portugaise (2023) permet de les caractériser de façon exhaustive. Ces caractéristiques tiennent premièrement aux effets d'échelle, d'envergure et aux *switching costs* qui résultent de l'entraînement des modèles. Ces économies d'échelle et d'envergure sont d'autant plus significatives que ces modèles fondamentaux peuvent être réutilisés pour de nombreux modèles aval fine-tunés. Deuxièmement, le développement des modèles et le contrôle des résultats mettent en jeu des effets d'apprentissages techniques et organisationnels qui donnent un avantage déterminant aux firmes qui descendent le plus rapidement la courbe d'expérience. Troisièmement, les réglementations qui pourraient s'appliquer aux modèles d'IA pourraient induire (en ce qu'elles sont symétriques et non asymétriques) des coûts de conformité³² dont le poids relatif serait moindre pour les grands opérateurs, leur donnant ainsi un avantage concurrentiel au travers notamment de l'augmentation des coûts des plus petits rivaux. Les entreprises dominantes peuvent dans ce cadre plaider pour un renforcement des réglementations visant à la protection des données utilisateurs, dans la mesure où cela a pour effet d'entraver le développement de solutions concurrentes.

³¹ <https://mlexmarketinsight.com/news/comment/comment-realpage-litigation-over-rent-setting-cartel-highlights-tacit-collusion-ai-enforcement>

³² L'autorité portugaise de concurrence (2023) cite notamment l'évaluation de Renda et al. (2022) laquelle évalue les coûts liés à la conformité à 17% des investissements initiaux.

Deux voies de discussion peuvent être envisagées. La première tient à la possibilité d'un développement autonome d'un marché aval de l'IA générative. La seconde tient à l'existence de goulets d'étranglement. Il convient de discuter l'hypothèse d'une construction de marché avec un segment aval (fine-tuning et applications) pouvant se développer de façon libre et non biaisée par rapport aux écosystèmes amont. Il s'agit donc de s'interroger sur le réel degré d'intégration verticale. ? Le modèle des partenariats peut se traduire par plusieurs dynamiques peu désirables qu'il s'agisse de quasi-intégrations verticales par contrat ou de phénomènes de dépendance technique ou économique conduisant à des logiques d'inféodation (Kenney et al., 2019). Nous développons donc ci-après une discussion sur les goulets d'étranglements allant au-delà même des modèles fondamentaux et des données sur lesquels ils ont été développés.

a) Les goulets d'étranglement

Il est possible de mettre en évidence de tels verrous en aval (avec l'accès aux applications utilisées par les clients finaux) et en amont avec l'accès aux infrastructures de stockage et de computation.

Canaux d'accès au marché (dimension aval)

Les grands écosystèmes numériques contrôlent un ensemble de ressources qui peut être déterminant dans l'accès au marché de l'IA générative et qui peut leur permettre de 'ré-équilibrer' leurs relations avec les entreprises à la pointe du développement de ces solutions et à partir de là être en mesure de consolider encore leur avantage concurrentiel malgré la disruption technologique à l'œuvre. Quelles sont ces ressources ? Il s'agit successivement des capacités commerciales d'accès aux clients, du contrôle d'infrastructures déterminantes comme les services cloud et enfin des capacités humaines et financières. Les Big Techs sont a priori particulièrement bien placées pour étendre leurs positions dominantes que cela soit en termes de concurrence par les mérites³³ ou de capacités de fausser la concurrence à leur profit.

L'accès aux services basés sur l'IA générative se fait essentiellement via des services proposés par de grands écosystèmes numériques, notamment des applications qui sont distribuées sur des magasins d'applications ou directement dans des services intégrés à chaque écosystème. En

³³ Elles jouissent d'avantages en termes de capacités techniques (concurrence sur le marché des talents), d'infrastructures techniques (infonuagiques), de capacités organisationnelles (firmes fondées sur la collecte, le traitement et la valorisation des données) et de capacités de négociation avec les tiers qu'il s'agisse en amont de contrats avec les fabricants de GPUs ou en aval d'externalisation de certaines tâches à des tiers dans des conditions de coûts particulièrement attractives du fait de l'existence d'un pouvoir d'oligopsonne étroit (voir Amazon Mechanical Turk par exemple).

outre, l'amélioration continue des résultats de l'IA générative va dépendre étroitement de la qualité et de la rapidité des boucles de rétroactions avec les usagers. Les Big Techs qui contrôlent la distribution des services peuvent de ce fait collecter et analyser des données déterminantes qui sont indispensables aux développeurs d'IA générative. Cela justifie d'autant plus des partenariats étroits.

Cependant, qu'il s'agisse de développer ses propres services d'IA générative ou d'héberger sur ses services ceux d'une entreprise verticalement liée, les risques de verrouillage des utilisateurs, d'orientation vers des services propres etc.. qui ont été mis en exergue dans le domaine des services numériques 'traditionnels' (Ezrachi et Stucke, 2020) sont ici à nouveau en jeu. Les modèles d'IA générative, qu'il s'agisse d'applications basées sur des modèles de langages fondamentaux ou des modèles fine-tunés, ont besoin d'accéder aux écosystèmes des Big Techs pour pouvoir accéder efficacement au marché. Cet accès est paramètre important en matière de concurrence horizontale sur ces deux marchés aval et amont en ce qu'il permet de rendre les services aisément accessibles aux utilisateurs mais également d'augmenter les données disponibles non pas seulement les données d'entraînements mais celles liées au déploiement des systèmes d'IA (cas des input-data). Cet accès peut également avoir un impact important dans l'éventuelle concurrence entre ces deux niveaux horizontaux en ce qu'il peut permettre aux développeurs de langages fondamentaux d'accéder aux données disponibles dans l'écosystème, lesquelles peuvent être susceptibles de permettre de contourner l'exclusivité des données à disposition des développeurs de modèles fine tunés.

Services d'infonuagique (dimension amont)

L'essentialité des Big Techs pour les développeurs d'IA générative tient également à l'accès aux services cloud à la fois pour le développement des modèles mais également pour leur fonctionnement. Cette essentialité joue encore comme un facteur d'intégration verticale (du moins en termes de partenariat) et ouvre également à des craintes de distorsions de concurrence aval, que celles-ci soient au détriment des développeurs ou dans le cadre de discriminations entre développeurs³⁴.

Il s'agit ici de considérer également des pratiques de nature unilatérale. Elles posent le problème des abus que pourrait mettre en œuvre un opérateur jouissant d'une position dominante vis-à-vis de ses concurrents et consommateurs. Par rapport aux consommateurs (ou aux utilisateurs

³⁴ Sur les avantages liés au contrôle des infrastructures (à la fois en matière de *hardware* mais également de *middleware*), voir notamment Jacobides et al. (2021).

de ses services), un premier type d'abus peut consister en un abus d'exploitation. La position dominante peut être instrumentalisée pour pratiquer des prix excessifs (*mark-up*) ou pour imposer des conditions de transaction inéquitables.

Le cas des services cloud, comme l'a illustré l'AdIC dans son enquête sectorielle de juin 2023, illustre une telle pratique³⁵. La fixation du niveau des frais de sortie du cloud (*egress fees*) peut conduire les firmes pratiquant du multi-hébergement à supporter des coûts rendant cette stratégie non viable financièrement³⁶. De surcroît, il n'y a pas qu'une finalité d'extraction du surplus mais également une finalité de coercition quant aux choix des utilisateurs, laquelle a un impact sur la concurrence.

Cela conduit au deuxième type d'abus, l'abus d'éviction. En l'espèce, celui-ci peut prendre deux formes. La première forme correspond aux stratégies de levier : il s'agit ici d'étendre sa position dominante vers un marché non encore dominé en instrumentalisant sa dominance sur le marché initial. La seconde forme correspond à une forclusion. Le marché de l'IA générative peut donner lieu à de tels risques. Les opérateurs majeurs de l'économie numérique peuvent instrumentaliser leurs positions de force respectives pour étendre leurs positions dominantes vers ce marché. Cela peut se faire au travers du contrôle des données, de capacités computationnelles et d'infrastructures de cloud ou encore via le contrôle de l'accès à leur écosystème.

Les positions de partenaires essentiels que peuvent avoir les grands opérateurs tiennent non seulement à la capacité à disposer de flux données massifs et diversifiés mais également au contrôle d'infrastructures et de capacités de calcul à la fois coûteuses et difficiles d'accès pour les nouveaux entrants du fait des engagements des fabricants vis-à-vis de leurs principaux utilisateurs déjà présents sur le marché. Le contexte de tensions sur l'offre de processeurs graphiques est donc de nature à renforcer leur essentialité. L'autorité portugaise de concurrence (2023) cite notamment l'exemple de Chat GPT4 qui a nécessité pour son développement une architecture technique de 285 000 CPUs et de 10 000 GPUs. Il convient d'insister sur le fait que les barrières à l'entrée (ou à l'expansion) en la matière tiennent non seulement aux coûts des investissements nécessaires mais également à la disponibilité des processeurs. Dans ce cadre, les stratégies d'intégration verticales doivent être considérées comme participant d'une logique de réduction de la dépendance amont. Elles doivent être considérées non seulement

³⁵ Autorité de la concurrence, avis n°23-A-08 du 29 juin 2023 portant sur le fonctionnement concurrentiel de l'informatique en nuage.

³⁶ Pour une synthèse de l'approche économique de ces questions, voir Gans et al. (2023).

sous le prisme de la dépendance des développeurs de modèles fondamentaux par rapport aux fabricants de processeurs mais également sous celui de l'équilibre entre Big Techs et ces mêmes fabricants. Le secteur des processeurs soulève également des préoccupations dans le domaine concurrentiel comme en témoignent la visite d'entreprises de l'Autorité de la concurrence en septembre 2023 dans le secteur des GPUs³⁷ ou encore l'investigation qui aurait été lancée par la Commission en octobre sur NVidia³⁸.

Les Big Techs peuvent également instrumentaliser ces positions pour verrouiller l'accès à des concurrents sur le marché aval de l'IA générative. Cela peut notamment être le cas pour des modèles dédiés construits sur la base de modèles de langages fondamentaux. En l'espèce, l'accès à ces modèles fondamentaux pourraient être considérés sous l'angle de la théorie des facilités essentielles. Se poseraient alors les questions d'examen du caractère objectif d'un refus d'accès, des conditions d'accès raisonnables et non discriminatoires si celui-ci devait être imposé en enfin de l'évaluation des effets nets sur les incitations³⁹.

Une dernière pratique abusive pourrait être prise en considération, celle de l'abus de dépendance économique. Celle-ci n'existe pas en droit de la concurrence de l'UE, elle pourrait se concevoir comme un abus d'exploitation sous la menace d'un abus d'éviction. Dans quelle mesure la position de verrouillage au profit du développeur d'un modèle fondamental ou d'un contrôleur de ressources essentielles peut-elle être instrumentalisée par ce dernier pour extraire une part indue du surplus ou pour consolider sa dominance. La notion de *kill zone* dans les écosystèmes numériques peut illustrer ces risques (Kamepalli et al., 2021).

Il serait possible d'interroger les verrous au développement de l'IA générative en termes d'essentialité : l'accès à certaines ressources est-il indispensable pour accéder au marché et un accès dans des conditions 'non discriminatoires' est-il indispensable pour une concurrence par les mérites sur le marché ?

Premièrement, comme nous l'avons vu *supra*, l'IA générative se caractérise par un certain nombre d'actifs qui sont sinon des actifs essentiels au sens du droit de la concurrence⁴⁰ du moins des verrous d'accès au marché (des *bottlenecks*) : au-delà des données et des infrastructures

³⁷ <https://www.autoritedelaconcurrence.fr/fr/communiqués-de-presse/le-rapporteur-général-de-l'autorité-de-la-concurrence-indique-qu'une-operation>

³⁸ <https://www.bloomberg.com/news/articles/2023-09-29/eu-begins-early-stage-probe-into-ai-chip-market-abuses-that-nvidia-dominates#xj4y7vzkg>

³⁹ Notons que l'existence de données disponibles en open source joue contre cet argument.

⁴⁰ Voir FTC (juin 2023) : *Generative AI Raises Competition Concerns*. <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns>

physiques que nous avons déjà abordées, nous pourrions citer les droits de propriété intellectuelles ou encore les talents⁴¹.

Deuxièmement, s'il est possible que les Big Tech actuelles (qui contrôlent ces actifs) ne soient pas obligatoirement les vainqueurs d'une course à l'IA générative par ce qu'il existe un marché des langages fine-tunés sur la base des langages fondamentaux, il n'en demeure pas moins que la concurrence sur le marché aval va dépendre des conditions d'accès à ces ressources amont. La capacité de certains acteurs à verrouiller va dépendre également de l'ouverture ou non des langages fondamentaux (choix de *l'open source*⁴²). Il s'agit donc de veiller à ce que la possibilité d'avoir une compétition 'aval' sur le segment des modèles 'personnalisés' (Carugati, 2023) ne soit pas entravée par les stratégies de verrouillage et / ou de levier des Big Techs.

Il convient cependant d'évaluer l'essentialité de l'accès tant en matière de droit de la concurrence que de dynamique industrielle. En effet, des innovations permettent de réduire le coût de développement des modèles de langages fondamentaux (LoRa de MS en 2021) mais également la « quantité » de données nécessaire (Schrepel, 2023). Pour les modèles fine-tunés, l'avantage tient plus à l'accès à des données uniques⁴³. De la même façon la dépendance amont va étroitement dépendre des possibilités techniques d'entraîner des modèles de langages sur des machines moins performantes et donc moins coûteuses. Encore une fois les progrès techniques pourraient être de nature à réduire l'importance des goulets d'étranglement⁴⁴.

b) Les stratégies d'intégration verticales « 3.0 » - Des acquisitions consolidantes aux partenariats

Nous avons vu supra que les marchés numériques en général se caractérisent par de fortes vagues d'acquisition. Cela est non seulement le fruit des capacités financières des opérateurs majeurs du secteur mais également d'un modèle de développement de nombreuses start-ups qui

⁴¹ Des stratégies de verrouillage concurrentiels peuvent être mises en place sur le marché du travail (clauses de non concurrence, no-poaching provisions, stratégies de préemption ou d'augmentation des coûts des rivaux...). Pour une synthèse, voir Posner (2023).

⁴² Le choix d'un modèle open source n'est pas obligatoirement bienveillant (dans le sens de *benevolent*) pour un acteur en position de pivot. Il peut faire de sa technologie un standard, générer des innovations complémentaires et alimenter la demande pour ses propres services (van der Vlist et al., 2024).

⁴³ Ainsi, les logiques économiques des ecosystem foundation models et des personal foundations models peuvent différer, les premiers seraient plus susceptibles de conduire à des phénomènes de *tipping*. Une concurrence par la qualité pourrait subsister sur les marchés aval : « Personal foundation models can thus be easily trained to assist each individual in a space where quality matters » (Schrepel, 2023, p.13).

⁴⁴ Il convient de relever que l'accès aux infrastructures contrôlées par les Big Techs ne se limite pas à un accès à des capacités computationnelles et de stockage. Il porte également sur la mise à disposition de services liés. La dépendance à une infrastructure donnée est renforcée par l'utilisation de ces outils spécifiques.

doivent s'adosser à un grand groupe porter leurs innovations vers le marché en disposant de leurs compétences techniques, de leurs portefeuilles de clients, de données etc... Un tel point commun avec le secteur des biotechs soulève des questions similaires quant à proportion d'acquisitions tueuses comme mentionné ci-avant. Si la littérature académique dans le secteur des Big Techs (Gautier et Lamesch (2021) et Ivaldi et al. (2023)) discute l'ampleur du phénomène, il convient d'envisager également d'autres scénarii dont celui de l'acquisition consolidante. L'acquisition consolidante a un double effet : pérenniser la position dominante sur le marché initial et l'étendre éventuellement à un autre.

Pourrait-on étendre ce scénario à l'IA générative ? De telles acquisitions ont lieu dans le domaine de l'IA générative. La prise de contrôle des entités développant des modèles fondamentaux par les 'Big Techs' est une illustration de ces risques sur les structures de marché. Dans ce cadre, la participation de MS dans Open AI peut être un premier angle d'analyse en termes de concentration verticale. Cependant la notion de prise de contrôle est à discuter au regard du statut légal d'Open AI qui est depuis 2019 une entité à profit plafonné. En 2015, il s'agissait d'une entité à but non lucratif). Pour autant, l'intérêt de la participation additionnelle prise par MS ne tient pas à une rémunération (plafonnée à cent fois la mise) mais au contrôle sur la technologie). Il n'existe pas d'exclusivité à proprement parler (et donc de logique de ventes groupées) dans la mesure où il est possible d'accéder à ChatGPT sans utiliser les services de MS, cependant, comme nous le verrons infra le partenariat peut être envisagé sous l'angle d'une « intégration verticale par contrat » même si elle n'est que partielle. Premièrement, comme nous le verrons ci-après, OpenAI utilise exclusivement les services cloud de MS. Deuxièmement, le partenariat n'exclut pas une influence dans le management de la firme, la crise au sein d'OpenAI dans le courant de l'automne 2023 et l'intervention de MS qui a contribué à ré-installer son CEO peut en constituer un indice. Troisièmement, les services d'OpenAI sont accessibles au travers de nombreuses applications de l'écosystème MS, à l'instar de GitHub ou de Copilot⁴⁵. Doit-on craindre le même scénario que celui des années passées ? Doit-on aller jusqu'à une limitation des capacités de diversifications (Khan, 2019) ou encore jusqu'aux propositions de mergers ban ⁴⁶?

⁴⁵ De la même façon, même s'il ne s'agit pas de *bundling* à proprement parler, il a pu être relevé que certaines des fonctionnalités de GitHub ou CoPilot sont accessibles au travers d'une souscription aux services cloud de MS. https://www.theinformation.com/articles/microsofts-ai-coding-product-becomes-weapon-in-battle-with-aws?utm_source=ti_app

⁴⁶ <https://www.reuters.com/technology/us-senator-wants-ban-big-tech-buying-anything-ever-again-2021-04-12/>

Les problèmes de concentrations ne se limitent pas aux Big Techs⁴⁷. Un second angle peut concerner les concentrations ‘horizontales’ sur le marché des processeurs, dès lors que l’on considère que l’accès à ces derniers est un moyen de verrouiller l’accès au marché pour les firmes aval ou du moins de se placer comme un partenaire incontournable vis-à-vis d’elles. Dans ce cadre le contrôle des opérations de concentrations entre les producteurs de GPUs peut s’avérer une entrée pertinente⁴⁸.

De façon plus générale encore se pose la question de la dépendance vis-à-vis des opérateurs de cloud et donc des Big Techs. Les services de cloud ne répondent pas à la catégorie des facilités essentielles en ce que des services concurrents et substituables sont effectivement présents sur le marché. Ils n’en ont pas moins deux caractéristiques qui peuvent poser problèmes en matière de concurrence. Premièrement, si la structure de l’offre amont n’est pas monopolistique, elle est étroitement oligopolistique. Deuxièmement, les stratégies multi-clouds voire les migrations d’un cloud à l’autre peuvent être techniquement difficiles et risqués ainsi que financièrement coûteux. Un verrouillage vertical peut se faire d’autant plus possible que les développeurs de langages fondamentaux ont tous noué des partenariats avec les grands exploitants de cloud⁴⁹, lesquels peuvent mettre en jeu des contrats d’exclusivité.

Les liens verticaux ne constituent pas en eux-mêmes une situation préjudiciable en termes de concurrence. Des coopérations entre firmes dans des activités coûteuses en investissements et particulièrement risquées sont nécessaires et une coordination à la fois stratégique et technique peut être nécessaire en matière d’investissements. Il s’agit de reconnaître que ces complémentarités nécessitent une forme de coordination à la fois en matière de choix de standards mais également de temporalité de déploiement des investissements (Amendola et al., 2010). Cependant, cette situation peut avoir un revers en termes de concurrence : elle réduit les incitations de l’une et l’autre des firmes engagées dans cette coopération verticale de chercher à s’intégrer pour mettre en cause la position de l’autre. Coopérer avec un concurrent naissant peut être une manière mutuellement profitable de le neutraliser voire de le faire contribuer au développement futur de l’écosystème.

⁴⁷ La distinction entre concentration verticale et concentration horizontale n’est d’ailleurs guère évidente dans la mesure où les IA génératives sont intégrées dans les moteurs de recherche.

⁴⁸ Voir pour un cas de concentration verticale entre un concepteur de GPUs et un producteur, l’action lancée en décembre 2021 par la FTC dans l’opération entre NVidia et ARM - <https://www.ftc.gov/news-events/news/press-releases/2021/12/ftc-sues-block-40-billion-semiconductor-chip-merger>

⁴⁹ Au-delà des liens entre MS et OpenAI, l’Autorité de la concurrence portugaise (2023) mentionne dans son rapport qu’Hugging Face utilise les services d’Amazon et qu’Anthropic utilise le cloud de Google notamment ses clusters de processeurs GPUs et TSUs (tensor processing units). Ces derniers étant des processeurs spécialement conçus pour accélérer l’entraînement des modèles d’IA.

De la même façon, il convient de considérer les dynamiques concurrentielles à l'œuvre en matière verticale entre producteurs de processeurs, entreprises de cloud et développeurs de langages fondamentaux. La séparation des trois maillons dans la chaîne de valeur peut s'avérer de moins en moins nettes avec des stratégies de remontées vers l'amont (Amazon) ou de descentes vers le cloud (NVidia). Cependant, l'essentialité du cloud peut être discutée au regard de deux voies de contournement. D'une part, le développement de modèles fine-tunés peut se faire en utilisant des infrastructures moins exigeantes et d'autre part, le déploiement lui-même peut ne pas obligatoirement reposer sur des services cloud, il peut reposer sur des technologies de type *edge*⁵⁰ ne nécessitant pas une connexion permanente et s'avérant moins coûteux en ressources (Vuruma et al., 2024).

Les distorsions de concurrence sur le marché aval

Il convient également de considérer avant de conclure un dernier plan de la concurrence. Celui de la concurrence sur le segment aval, celui des modèles fine-tunés. Les entreprises actives sur ce segment peuvent faire face à des conditions de concurrence faussées que cela soit en matière d'accès aux modèles de fondations ou à certaines ressources contrôlées par des entreprises verticalement liées. La question sous-jacente est celle de la disponibilité de ressources adéquates en open source. Cela a été le cas pour certains modèles en open source mais ces derniers même ouverts peuvent présenter des restrictions en matière d'adjonction de services⁵¹, notamment pour ce qui est des applications commerciales (Autorité de la concurrence portugaise, 2023, p.29). De la même façon, l'accès à ces modèles par le moyen d'API est techniquement relié à des services de cloud.

Des problèmes concurrentiels somme toute classiques peuvent donc être appréhendés. Un premier tient à de possibles stratégies d'auto-favoritisme. Ces stratégies de levier peuvent se traduire par des pratiques de ventes liées ou d'accès discriminatoire à certaines ressources. Parmi ces conditions biaisées, il est possible de craindre des distorsions dans l'accès aux

⁵⁰ L'edge computing correspond à une informatique en périphérie de réseau. Le stockage et le traitement informatiques ne sont pas exécutés sur le réseau mais en local. Les communications montantes et descendantes ne se font que de façon ponctuelle.

⁵¹ La distinction entre modèles ouverts et modèles fermés doit être discutée. Il existe un continuum en matière de degrés d'ouverture pouvant conduire à de réelles restrictions en matière de développement de modèles dérivés. Ces restrictions peuvent se traduire par des impossibilités techniques en termes de reproductibilité, d'accès aux données d'entraînement, au code source utilisé pour ce même entraînement etc... Parfois seulement des modèles simplifiés sont disponibles en *open source* et l'accès au modèle dans son entièreté passe par une API spécifique comme le montre le cas du modèle Mistral Large.

<https://mistral.ai/fr/news/mistral-large/>

modèles fondamentaux ou encore des accès discriminatoires aux données nécessaires⁵². Ce dernier point est indubitablement à discuter à l'aune des dimensions abordées dans ce document de travail. Certes, des données sont disponibles en accès ouvert, certes des données peuvent faire l'objet d'obligations de partage sous l'effet des régulations mais des restrictions peuvent être envisagées. Cela peut être le cas pour des données spécifiques utilisées pour le monitoring des modèles ou encore pour des données utiles commercialement dont disposent les firmes pivotales des écosystèmes numériques grâce à la diversification de leurs activités. Pour autant, les firmes aval peuvent bénéficier d'accès à des données spécifiques qui peuvent être génératrices de valeur⁵³. La question finale est donc reliée à un phénomène de dépendance technologique et économique. Les firmes amont peuvent-elles à partir de leurs ressources 'essentielles' capturer la valeur ajoutée créée par ces nouveaux acteurs ou risquent-elles d'être disruptées ?

L'ensemble des coûts et des difficultés d'accès aux ressources citées *supra* pourrait permettre aux Big Techs de prendre directement ou indirectement le contrôle des activités de compléments, même si ces derniers disposent de données leur permettant de développer des modèles personnalisés à forte valeur ajoutée. Ces mêmes firmes peuvent accéder aux données au travers de stratégies d'intégration verticale ou de partenariat, voire les contourner au travers de stratégies de diversification de leurs activités vers les secteurs de la santé, des secteurs financiers etc... En outre, l'idée selon laquelle les nouveaux entrants peuvent disrupter les opérateurs installés en acquérant des bases de données originales ou en utilisant des données synthétiques peut être discutée à l'aune des capacités d'investissement des opérateurs installés eux-mêmes⁵⁴.

⁵² Il est également possible de considérer les risques d'accès biaisés aux capacités computationnelles. Si un service cloud devient 'incontournable' pour un utilisateur donné (par exemple si les restrictions d'interopérabilité et de portabilité des données sont significatives ou encore si les *egress fees* aux stratégies *multiclouds* sont significatifs), alors il serait possible de réduire la menace concurrentielle qu'il fait peser en amont en entravant son accès aux ressources cloud en matière tarifaire et ou en matière technique.

⁵³ Une Big Tech dont les activités seraient diversifiées vers des secteurs donnant accès à des données à forte valeur ajoutée (données de santé, données juridiques) pourrait acquérir une position de force sur le marché des modèles fine-tunés. Elle pourrait profiter de ses autres actifs et de ses capacités techniques. Elle pourrait également au travers de ses ressources financières nouer des partenariats exclusifs avec des producteurs de données particulièrement intéressantes dans le cadre de la concurrence sur ce marché aval (groupes de presse, éditeurs scientifiques, ...).

⁵⁴ Ces acteurs peuvent avoir des difficultés en termes financiers pour s'aligner sur les propositions des développeurs de langages fondamentaux pour négocier l'accès à des données complémentaires ou substituables à celles qu'ils contrôlent déjà. Cette concurrence peut être également biaisée par l'adossement de certains de ces développeurs de modèles fondamentaux avec les Big Techs. En outre ces mêmes Big Techs peuvent être en mesure de fausser la concurrence en induisant des distorsions dans l'accès aux capacités computationnelles au profit des firmes avec lesquelles ont des partenariats. Cet accès est d'autant plus déterminant qu'il ne concerne pas seulement

Cette question peut être considérée non seulement à l'aune des relations verticales entre Big Techs et leaders de l'IA générative mais également dans une perspective aval. La capacité d'investissement de certains de ces opérateurs (notamment au travers de leurs financements venant des Big Techs) ne leur permet-elle pas d'acquérir des bases de données particulièrement riches (presse, éditeurs scientifiques etc...) leur permettant de se doter d'un avantage déterminant sur leurs compétiteurs horizontaux ? En d'autres termes, le modèle connu pour les écosystèmes numériques ne risque-t-il pas d'être reproduit en ce qui concerne cette nouvelle technologie ? Les risques seraient d'autant plus élevés, si l'on considère par exemple les données, que les contrats conclus le seraient à titre exclusif ? Le verrouillage concurrentiel – ou la création de positions dominantes aval que l'on pourrait craindre inexpugnables – pourrait être d'autant plus fort si ces acteurs bénéficiaient au titre de leurs partenariats avec les grands écosystèmes amont d'un accès privilégié aux infrastructures cloud.

Il convient également de relever que les phénomènes de dépendance verticale sur lesquels nous avons insisté (entre les Big Techs qui contrôlent des actifs essentiels amont et les développeurs de modèles fondamentaux) peuvent être reproduits en aval entre ces mêmes développeurs et ceux de modèles modifiés. Les mêmes stratégies de verrouillage et de levier peuvent être appréhendés dans le marché des applications et les mêmes logiques de dépendance économique et de discrimination peuvent être à l'œuvre. Les développeurs de langages fondamentaux peuvent fonctionner comme des pivots d'un écosystème numérique. Par exemple, Open Markets dans sa réponse à la consultation mise en place par la Commission européenne⁵⁵ montre qu'OpenAI a lancé un magasin d'applications permettant le développement de versions spécifiques de ChatGPT. Ainsi, les problématiques de contrôleurs d'accès peuvent être aisément retrouvées à ce niveau. Notons par ailleurs que le développement de processeurs spécifiques au gré d'une intégration amont pourrait rendre encore plus essentiel un partenariat des développeurs de modèles fine-tunés avec les développeurs de modèles de langages fondamentaux et accroître d'autant les risques de verrouillage et de forclusion décrits *supra*.

Le phénomène de platformisation propre à l'IA générative qui la rapproche du fonctionnement général des écosystèmes numériques. L'intégration entre le développement, la mise en œuvre et les retours d'apprentissage suppose l'insertion dans de tels écosystèmes tout comme l'impose le recours à des infrastructures techniques, qu'elles soient physiques (cloud) ou qu'elles

l'entraînement des modèles mais également leur déploiement. Ces différences d'accès peuvent concerner des paramètres relatifs à la qualité mais également aux coûts.

⁵⁵ https://ec.europa.eu/commission/presscorner/detail/en/IP_24_85

procèdent de questions de standards d'interopérabilité. Les économies d'échelle et d'envergures propres à l'IA générative, les effets d'apprentissage, les besoins financiers, mais également les effets de réseaux et les questions d'accès privilégié à telle ou telle base de données ou à des infrastructures de services cloud vont conduire possiblement non seulement à des phénomènes de concentration sur le segment aval mais également à des formes d'intégration verticale allant possiblement à l'encontre d'une hypothèse disruptive. La dépendance verticale peut prendre la forme d'une influence sur les décisions stratégiques des firmes développant les modèles fondamentaux mais également sur les décisions techniques, lesquelles peuvent avoir des effets particulièrement structurants en termes d'écosystèmes et renforcer les craintes d'une structuration en silos autour de *moligopoles* (Petit, 2020) certes en concurrence sur de nombreux marchés présents et futurs mais pouvant tenir leurs usagers et leurs complémentaires dans un mono-hébergement de plus en plus contraint.

IV. Discussion

Avant de conclure sur les outils réglementaires qui pourraient être nécessaires pour prévenir une prise de contrôle du marché de l'IA générative par les Big Techs, il convient de montrer que les avantages dont disposent celles-ci ne se limitent pas au contrôle d'actifs qu'il s'agisse des infrastructures physiques ou des données. Ces avantages tiennent aux capacités organisationnelles mêmes de ces firmes qui les rendent particulièrement capables à absorber les innovations développées dans leur environnement et à les amener efficacement au marché. Le concept de monopoles intellectuels (Rikap, 2021) recouvre en fait le développement par ces firmes de *Corporate Innovation Systems* (CIS) particulièrement efficaces.

A. Disruption au travers de l'IA générative ou consolidation de monopoles intellectuels ?

La question de l'IA générative permet *in fine* de réinterroger la dynamique concurrentielle dans les marchés de haute technologie. Doit-on craindre une confiscation des dynamiques d'innovation potentiellement disruptrices par les opérateurs installés ou au contraire peut-on compter sur les turbulences concurrentielles et technologiques pour rebattre périodiquement les cartes ?

La première approche conduit à la mise en œuvre d’une approche particulièrement précautionneuse *ex ante* et intrusive *ex post*. Elle doit viser par des règles à rendre la concurrence sur les marchés futurs libre, loyale et praticable ou à se doter de remèdes concurrentiels aptes à rétablir les conditions d’une concurrence à égalité des armes.

La seconde approche repose quant à elle sur une hypothèse d’imprévisibilité de la dynamique concurrentielle (Petit et Teece, 2021). Elle propose plus de favoriser l’innovation pour répondre aux problèmes actuels de dominance que de contraindre cette même innovation dès lors qu’elle serait le fait de l’un des opérateurs dominants actuels. L’idée sous-jacente est que les Big Tech sont en concurrence les unes avec les autres (Petit, 2020) et qu’elles sont peut-être des challengers sur certains marchés, notamment ceux de l’IA générative (Manne et Auer, 2024). Considérer dans le domaine de l’IA générative des dimensions reliées à l’avantage en termes de données ou au contrôle des infrastructures computationnelles comme quasi ‘essentiels’ peut conduire dans cette acception à une analyse exagérant l’importance des barrières à l’entrée (Schrepel, 2024).

Le caractère de barrière à l’entrée du segment des infrastructures peut être discuté dans cette perspective. Il est possible d’investir pour s’intégrer vers l’amont à l’instar d’OpenAI par exemple ou de le contourner au travers de solutions algorithmiques ou d’*edge computing*. De la même façon, le caractère essentiel des données peut être discuté comme nous l’avons vu dans ce document de travail. Certaines données pour développer des langages fondamentaux sont accessibles librement. Des entrants peuvent avoir besoin de moins en moins de données grâce aux innovations logicielles ou au travers du recours à des données synthétiques.

Cependant, cette approche d’une concurrence dynamique (Petit et al., 2024) qui préconise une réglementation flexible et à la main légère en faisant la part belle à la capacité des innovations pour remettre en cause les positions dominantes⁵⁶ et les capacités d’étendre celles-ci à des marchés connexes, fait l’objet de contestations (McLean, 2024, Stucke and Ezrachi, 2024). En effet, l’hypothèse d’une turbulence concurrentielle suffisante pour résoudre d’elle-même les questions de concentrations du pouvoir économique peut être vue comme participant d’un « techno-conservatisme » qui reviendrait à réhabiliter le laissez-faire au nom d’un caractère autorégulateur des marchés qui reposerait sur la dynamique de l’innovation.

Il convient de discuter cette hypothèse. D’une part, il ne s’agit pas réellement d’un laissez-faire dans la mesure où cette approche suppose une intervention antitrust pour prévenir des pratiques

⁵⁶ Les prémisses de cette approche sont notamment à rechercher chez Posner (2001).

qui peuvent s'avérer dommageables à l'innovation (Petit et al., 2024). Elle se rattache ce faisant à la prise en considération d'éléments tels les *innovation markets* dans le cadre du contrôle des concentrations⁵⁷. D'autre part, l'hypothèse que l'innovation peut provenir des entreprises les mieux installées conduit à revisiter les débats économiques sur le jeu des incitations et des capacités à investir (Aghion et al., 2005) et sur les capacités des opérateurs installés à faire obstacle au développement d'innovations venant de tiers (Marty et Warin, 2020b) ou à les capter à leur profit au travers notamment de prises de contrôle, quelles soient capitalistiques ou contractuelles.

Une autre dimension de ce débat tient à la capacité des Big Tech à maîtriser les risques de disruptions technologiques non seulement au travers de leurs capacités techniques et financières mais également au travers de leurs capacités à centraliser les informations et à se connecter au plus près des développements des connaissances scientifiques (Rikap, 2023). Nous avons vu avec Iansiti et Lakhmani (2020) que la supériorité des entreprises fondées sur la donnée tient à leur organisation technique qui leur permet d'intégrer les flux de données très rapidement dans leurs développements et dans leur cadre du déploiement même de leurs algorithmes pour avoir des boucles de corrections et d'amélioration très rapides. Il s'agit ici d'une dimension plus en amont. Il s'agit dans ce cas de figure de la caractérisation d'une nature de *monopole intellectuel* qui se caractérise par de fortes compétences en matière non seulement de détection mais également d'appropriation de l'innovation. En d'autres termes, les firmes en question développent des investissements leur permettant d'absorber les innovations venant de tiers mais également les connaissances produites par leur environnement (Cohen et Levinthal, 1990). Dans cette perspective, les entreprises dominantes de la Tech seraient d'autant plus susceptibles de maintenir et d'étendre leurs pouvoirs de marché qu'elles développent ces compétences. Les avantages liés aux données, au contrôle des infrastructures ou encore à la position de contrôleurs d'accès dans leurs écosystèmes respectifs ne sont pas les seuls qui peuvent expliquer la pérennité des positions de marché. L'ensemble de ces caractéristiques les met en mesure de « capturer » les connaissances développées par des tiers (qui sont souvent des complémentaires ou des partenaires en matière de recherche appliquée ou même fondamentale) et de les intégrer rapidement et efficacement dans leurs services.

⁵⁷ La notion fut introduite dans la littérature par Gilbert et Sunshine (1995) et connaît une nouvelle actualité dans les réflexions autour de la mise en œuvre des règles de concurrence dans le contexte de turbulences concurrentielles et technologiques (voir notamment OECD (2023)).

L'idée sous-jacente à l'hypothèse du « monopole intellectuel » est de considérer que les turbulences technologiques dans son environnement sont bien moins un risque de disruption qu'une ressource pour l'entreprise dominante. Les Big Techs seraient d'autant mieux placées pour en tirer profit qu'elles disposent de ressources essentielles pour permettre une rapide atteinte d'une échelle pertinente et que leurs propres investissements internes en matière d'innovation les rendent particulièrement efficaces en matière d'appropriation des connaissances venant de tiers et d'intégration rapide dans leurs services. Ces avantages sont d'autant plus adaptés au domaine de l'IA (Rikap et Lundvall, 2021).

En ce sens l'analyse de la position des Big Techs est indissociable de celle de leur centralité dans les écosystèmes numériques. Leur résilience tient en grande partie à leur capacité intrinsèque à développer des innovations mais également à absorber les connaissances créées au sein de leur écosystème et de les transformer en innovations aptes à accéder rapidement et efficacement au marché (Rikap et Lundvall, 2020). Dans le domaine de l'économie industrielle, les relations entre les firmes pivot et leurs complémentaires ont été particulièrement étudiées pour montrer de quelle façon les premières peuvent instrumentaliser les déséquilibres des relations contractuelles et leurs capacités à réguler leurs écosystèmes pour contrôler la dynamique concurrentielle à leur profit (voir par exemple Marty et Pillot, 2021). Cependant le cas de l'IA, cette centralité ne joue pas seulement vis-à-vis de firmes développeuses d'applications ou de vendeurs indépendants mais également vis-à-vis d'universités, de fondations etc.... Les firmes pivot accompagnent ces entités dans le cadre de leurs recherches en leur fournissant les ressources nécessaires à leur développement mais s'avèrent, du fait même de ce soutien, aptes à s'approprier au moins partiellement les innovations. La réussite de ce réseau d'innovation passe à la fois par la garantie de l'accès à des actifs physiques, des stratégies d'open source et le renforcement des capacités internes d'apprentissage organisationnel.

B. Quels enseignements provisoires ?

Considérant les logiques économiques et techniques et les risques concurrentiels associés, quels peuvent être les outils de prévention des risques les plus adaptés ?

Certaines des Big Techs contrôlent des infrastructures critiques – elles deviennent des goulets d'étranglement en amont notamment par le cloud. Elles ne sont cependant pas les seules à être

à jouer le rôle de verrous en amont. C'est le cas des fabricants de processeurs, voir le cas des GPUs (NVidia), mais cela peut aussi à terme être le cas des développeurs de modèles de langages fondamentaux. C'est donc un paysage concurrentiel plus complexe qu'à l'accoutumée qui se dessine mais avec néanmoins des constantes. Le marché de l'IA générative se comporte comme les autres marchés numériques. Il faut prendre en compte des actifs critiques comme les données mais aussi l'importance des effets de réseaux, des économies d'envergure et d'échelle et de la question de l'accès aux applications. Le paradoxe apparent est que les firmes leaders ne sont pas les Big Techs et qu'elles occupent peu à peu une place de pivots au sein d'écosystèmes permettant à des firmes complémentaires de développer des modèles fine-tunés. Ces firmes peuvent donc apparaître comme des rivaux potentiels des Big Techs. Si elles sont considérées comme dotées d'une puissance de marché, il s'agirait d'éviter d'entraver les Big Techs sur ce marché car leur entrée favoriserait la concurrence.

Se dessine cependant un paysage concurrentiel bien plus complexe dès lors que l'on considère la chaîne de valeur : il existe plusieurs goulets d'étranglement au sein de celle-ci et les tendances à l'intégration verticales peuvent être très fortes à la fois du fait des capacités techniques et financières des différents acteurs et des nécessités spécifiques de coordination d'investissements complémentaires. Au niveau le plus amont se situent les fabricants de processeurs. Ils sont essentiels aux Big Techs qui contrôlent les infrastructures cloud et sont en concurrence potentielles avec elles.

Les Big Techs peuvent être tentées de s'intégrer vers l'amont en produisant leurs propres processeurs⁵⁸, les fabricants de GPUs peuvent de leur côté être tentés de construire leur propre infrastructure cloud⁵⁹. En aval des Big Techs, se situent les développeurs de modèles de langages fondamentaux. Ils font figure de disrupteurs vis-à-vis des Big Techs en ce qu'ils peuvent développer des services concurrents à leurs cœurs de métiers et qu'ils sont au cœur de différents écosystèmes en aval dans le domaine de l'IA générative. Cependant, ils sont au minimum dans une situation de coopération avec ces derniers. En effet, ils ont besoin de ressources clés que les Big Techs détiennent que cela soit en matière de capacités computationnelles, de données, de capacités financières ou de ressources techniques. Il y a donc un certain lien de dépendance de l'aval vers l'amont. Cependant cette dépendance peut être vue comme bilatérale – au moins pour certaines Big techs – dans la mesure où les services d'IA

⁵⁸ C'est l'une des questions posées par l'Autorité de la concurrence dans sa consultation sur l'IA générative.

⁵⁹ La centralité de tel ou tel producteur de processeurs sur le marché tient non seulement à la performance de ses GPUs et à ses capacités de production, mais également à sa maîtrise du middleware.

proposés peuvent être intégrés dans leurs services. Ces liens bilatéraux ne sont pas pour autant symétriques. Les ‘partenariats’ conclus dans le domaine peuvent consacrer des intégrations par contrat ou conduire à des effets de verrouillage. Cependant, les développeurs de modèles de langages fondamentaux sont dans une situation bien supérieure à celles d’autres complémentaires dans les écosystèmes numériques en ce que leurs services sont difficiles à reproduire par les Big Techs. Un autre point est à relever : ces firmes peuvent être en mesure de s’intégrer verticalement notamment en développant leurs propres technologies de processeurs.

Finalement, il convient de considérer un dernier segment en aval : celui des modèles fine-tunés. Ici nous pouvons reprendre la même lecture. Nous sommes à la fois en présence d’acteurs qui peuvent bénéficier de la combinaison de modèles de langage fondamentaux en open source et de données propriétaires difficilement reproductibles. Ils peuvent gagner un avantage compétitif net au travers d’innovations algorithmiques et pourraient réduire leur dépendance aux services cloud en utilisant des processeurs moins élaborés pour l’entraînement. Cependant ce scénario doit être nuancé sur plusieurs points. Ils se situent toujours dans une logique de complémentaires vis-à-vis des services cloud des Big Techs et des modèles fondamentaux des grands acteurs de l’IA générative. Les risques concurrentiels décrits dans le cadre du DMA peuvent donc à nouveau être pris en compte. Ainsi nous sommes confrontés à une chaîne de valeur verticale caractérisées par des goulets d’étranglement et de fortes asymétries. Le risque de voir se développer une logique de silos ne peut être exclu, ainsi que celui d’observer des phénomènes de double marginalisation le long des chaînes de valeur de l’IA générative⁶⁰.

Des remèdes radicaux ont déjà pu être proposés en matière de concurrence dans le secteur numérique (Gal et Petit, 2021). Dans le domaine des données, ils tiennent à une confirmation des règles actuelles en matière de non-discrimination en matière d’accès voire de partage mais n’en posent pas de réels enjeux en termes de conformité, notamment en matière de protection des données personnelles. En fait, c’est plus au niveau des infrastructures que des demandes de remèdes assez intrusifs s’expriment. Ils peuvent prendre une forme d’exigence de non-discrimination dans l’accès à certaines ressources critiques (cloud, etc...) mais également la forme d’imposition de ‘principes de spécialité’ revenant à l’interdiction d’intégration entre fournisseurs de services cloud et développement de modèles fondamentaux.

⁶⁰ Sur la question de la double marginalisation, se reporter à Choné et al. (2023).

Une « régulation prudentielle » pourrait passer par l'application du DMA aux services cloud au titre des services de plateforme essentiels (ce qui est prévu par l'article 2 du DMA mais pas encore effectif dans la liste des services établie en septembre 2023) mais également aux grands modèles fondamentaux d'IA. Une telle régulation pourrait également passer par un élargissement du champ du contrôle des concentrations de l'UE aux partenariats et aux prises de participations même minoritaires⁶¹. Il pourrait également s'agir d'utiliser les ressources ouvertes par le DMA pour superviser spécifiquement les acquisitions réalisées par les contrôleurs d'accès et de développer spécifiquement l'outil des enquêtes de marché à la britannique dont la piste avait été esquissée dans le nouvel instrument concurrentiel de juin 2020 (Cartapanis et Marty, 2020). Il convient d'évoquer des remèdes qui relèvent de la politique industrielle dans un sens bien plus large. Dès lors que les services cloud et les langages de programmation apparaissent comme des infrastructures économiques essentielles, des options de cloud souverains⁶² voire de modèles de langage fondamentaux indépendants de ceux liés aux Big Techs ont pu être proposés⁶³.

Références

Acemoglu D. and Lensman T., (2023), “Regulating Transformative Technologies”, *NBER Working Paper*, n°31461, <https://www.nber.org/papers/w31461>

Aghion P., Bloom N., Blundell R., Griffith R. and Howitt P., (2005), “Competition and Innovation: An Inverted U Relationship”, *Quarterly Journal of Economics*, vol 120, pp.701 et s.

Amendola M., Bruno S. and Gaffard J.-L., (2010), “Hicks and Richardson on Industrial Change: Analysis and Policy”, in Gaffard J.-L. and Salies E., eds, *Innovation, Economic Growth and the Firm*, Edward Elgar.

⁶¹ Au sein de l'UE, l'Allemagne et l'Autriche se sont dotées d'un tel instrument. Il existe aussi sous d'autres formes aux Etats-Unis et au Royaume-Uni. De nombreuses réflexions sont en cours chez de nombreux Etats Membres de l'Union Européenne. Voir notamment van den Boom et al. (2023).

⁶² Voir Marty (2023a) sur ces questions.

⁶³ Voir notamment les annonces réalisées en avril 2023 par le gouvernement britannique. <https://www.gov.uk/government/news/initial-100-million-for-expert-taskforce-to-help-uk-build-and-adopt-next-generation-of-safe-ai>

- Autoridade da Concorrência, (2023), *Competition and Generative Artificial Intelligence*, Issue Paper, November, <https://www.concorrenca.pt/en/articles/adc-warns-competition-risks-generative-artificial-intelligence-sector>
- Belleflamme P. and Peitz M., (2020), “Platforms and Network Effects”, in Corchon L. and Marini M., eds, *Handbook of Game Theory and Industrial Organization*, Edward Elgar, <https://www.e-elgar.com/shop/gbp/handbook-of-game-theory-and-industrial-organization-volume-i-9781785363344.html>
- Cartapanis M. and Marty F., (2020), “Towards New Tools in Competition Law”, *Competition Forum: Law & Economics*, November, art. n° 0008, <https://www.competition-forum.com/>
- Carugati C., (2023), Competition in generative artificial intelligence foundation models, *Bruegel Working paper*, July. <https://www.bruegel.org/working-paper/competition-generative-artificial-intelligence-foundation-models>
- Choné P., Linnemer L. and Vergé T., (2023), “Double Marginalization, Market Foreclosure, and Vertical Integration”, *Journal of the European Economic Association*, <https://doi.org/10.1093/jeea/jvad069>
- Cohen W.M. and Levinthal D.A., (1990), “Absorptive Capacities: A New Perspective on Learning and Innovation”, *Administrative Sciences Quarterly*, 35(1), pp.128-152.
- Competition and Markets Authority, (2021), *Merger Assessment Guidelines*, CMA 129, March.
- Competition and Markets Authority, (2023), *AI Foundation Models Initial Report*, <https://www.gov.uk/government/publications/ai-foundation-models-initial-report>
- Ezrachi A. and Stucke M.E., (2020), “Digitalisation and its impact on innovation”, *R&I paper series*, 202/07, EU Commission.
- Iansiti M. and Lakhani K.R., (2000), *Competing in the Age of AI – Strategy and Leadership When Algorithms and Networks Run the World*, Harvard Business Review Press, Boston, Massachusetts, 267p.
- Furman J., Coyle D., Fletcher A., McAuley D. and Marsden P., (2019), *Unlocking Digital Competition*, <https://www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel>

- Gal M. and Petit N., (2021), “Radical Restorative Remedies in Digital Markets”, *Berkeley Technology Law Journal*, 36(2), pp.617-674.
- Gans J.S., (2024), “Market Power in Artificial Intelligence”, *NBER Working Paper*, n°32270, March.
- Gans J.S., Hervé M. and Masri M., (2023), *Economic analysis of proposed regulations of cloud services in Europe*, *European Competition Journal*, 19 (3), pp.522-568.
- Gautier A. and Lamesch J., (2024), “Mergers in the Digital Economy”, *Information Economics and Policy*, 54, March, <https://doi.org/10.1016/j.infoecopol.2020.100890>
- Gilbert R.J. and Sunshine S.C., (1995), “Incorporating Dynamic Efficiency Concerns in Merger Analysis: The Use of Innovation Markets”, *Antitrust Law Journal*, 63(2), pp.569-601.
- Hagi A. and Wright J., (2023), “Data-enabled learning, network effects and competitive advantage”, *Rand Journal of Economics*, 54(4), pp.638-667, <https://doi.org/10.1111/1756-2171.12453>
- Ivaldi M., Petit N. and Ünekbaş S., (2023), “Killer acquisitions: evidence from EC Merger Cases in Digital Industries”, *TSE Working Paper*, n°1420, March
- Jacobides M.G., Brusoni S. and Candelon F., (2021), “The Evolutionary Dynamics of the Artificial Intelligence Ecosystem”, *Strategy Science*, vol.6, pp.412-435.
- Kamepalli S.K., Rajan R.G. and Zingales L., (2021), “Kill Zone”, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3555915
- Kenney M., Rouvinen P., Seppälä T. and Zysman J., (2019), “Platforms and industrial change”, *Industry and Innovation*, 26(8), pp.871-879, <https://doi.org/10.1080/13662716.2019.1602514>
- Kerber W., (2022), “The German Facebook case: the law and economics of the relationship between competition and data protection law”, *European Journal of Law and Economics*, 54, pp.217-250.
- Khan L., (2019), “The Separation of Platform and Commerce”, *Columbia Law Review*, 119, pp.973 et s.
- Koperniak S., (2017), “Artificial Data Give the Same Results as Real-Data – Without Compromising Privacy”, MIT Press, March 3.

- Manne G.A. and Auer D., (2024), "From Data Myths to Data Reality: What Generative AI Can Tell Us About Competition Policy (and Vice Versa)", *CPI Antitrust Chronicle*, February.
- Marty F., (2019), "Plateformes de commerce en ligne et abus de position dominante : réflexions sur les possibilités d'abus d'exploitation et de dépendance économique", *Thémis (Revue Juridique Thémis de l'Université de Montréal - RJTUM)*, volume 53, pp. 73-104.
- Marty F., (2021), "Competition and Regulation Challenges in Digital Markets: How to Tackle the Issue of Self-Preferencing?", in Molan Gaban E. and Klein V., eds, *Concorrência e Inovações: Reflexões e Insights*, Instituto Brasileiro de Concorrência e Inovações, November, pp.109-141.
- Marty F., (2023a), « Les politiques publiques européennes en faveur d'un cloud souverain : fondements, modalités de mise en œuvre et évaluation critique », *GREDEG Working Paper*, n° 2023-03, mars.
- Marty F., (2023b), « Economie de la donnée : Ecosystèmes numériques, algorithmes et intelligence artificielle », in Bruguière J.-M., s.d., *L'émergence d'un droit des données*, Lefebvre-Daloz, collection Thèmes et Commentaires, Paris, pp.21-51, novembre.
- Marty F. and Pillot J., (2021), "Cooperation, dependence, and eviction: how platform-to-business competition relationships should be addressed in mobile telephony ecosystems", in Michal Gal and David Bosco eds. *Challenges to Assumptions in Competition Law*, Edward Elgar, April, pp. 2-22
- Marty F. et Warin T., (2020a), "Concurrence et innovation dans les écosystèmes numériques à l'ère de l'intelligence artificielle", *Concurrences*, 1-2020, pp.36-41.
- Marty F. and Warin T., (2020b), "Visa Acquiring Plaid: A Tartan over a Killer Acquisition? Reflections on the risks of harming competition through the acquisition of startups within digital ecosystems", *Cahier Scientifique du CIRANO*, 2020s-62, December.
- McLean A.P., (2024), "Innovation against change", *Journal of Antitrust Enforcement*, *forthcoming*
- Motawani S.R. and al., (2024), "Secret collusion among generative AI agents", arXiv:2402.07510v1, February.

- OECD, (2023), *The Role of Innovation in Competition Enforcement*, OECD Competition Policy Roundtable Background Note, www.oecd.org/daf/competition/the-role-of-innovation-in-competition-enforcement-2023.pdf
- Petit N., (2020), *Big Tech and the Digital Economy - The Mologopoly Scenario*, Oxford University Press.
- Petit N. and Teece D.J., (2021), “Innovating Big Tech Firms and Competition Policy: Favoring Dynamic against Static Competition”, *Industrial and Corporate Change*, 30, pp.1168 et s.
- Petit N., Heiden B. and Schrepel T., (2024), “Situating Dynamic Competition: An Evolution Beyond Chicago”, *DCI Working Paper*, n°8, January.
- Pillot J., (2009), « Comment s'établit un standard technologique : la stratégie de Sony sur le marché des lecteurs et supports numériques haute définition », *Innovations*, 2009/2, 30, pp.135-155.
- Posner E., (2023), “The New Labor Antitrust”, *University of Chicago Coase-Sandor Institute for Law & Economics Research Paper*, September. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4575258#
- Posner R., (2001), « Antitrust in the New Economy », *Antitrust Law Journal*, vol.68.
- Renda A. et al., (2022), *Study to support an impact assessment of regulatory requirements for Artificial Intelligence in Europe*, EU Commission, <https://op.europa.eu/en/publication-detail/-/publication/55538b70-a638-11eb-9585-01aa75ed71a1>
- Rikap C., (2021), *Capitalism, Power and Innovation. Intellectual Monopoly Capitalism Uncovered*, Routledge
- Rikap C., (2023), “Intellectual monopolies as a new pattern of innovation and technological regime”, *Industrial and Corporate Change*, <https://doi.org/10.1093/icc/dtad077>
- Rikap C. and Lundvall B-A., (2020), “Big Tech, Knowledge Predation and the Implications for Development”, *Innovation and Development*, 12(3), pp.389-416
- Rikap C. and Lundvall B-A., (2021), *The Digital Innovation Race: Conceptualizing the Emerging New World Order*, Palgrave Mc Millan.

- Schrepel T., (2023a), “Competition between AI Foundation Models: Dynamics and Policy Recommendations”, *MIT Connection Science Working Paper 1-2003*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4493900
- Schrepel T., (2023b), “Toward a working theory of ecosystems in antitrust law: the role of complexity science”, *Working paper – symposium – dynamics of generative AI*.
- Schrepel T. and Pentland A., (2023), “Competition Between AI Foundation Models: Dynamics and Policy Recommendations”, *MIT Connection Science Working Paper*, June.
- Stucke M.E. and Ezrachi A., (2024), “Innovation Misunderstood”, *American University Law Review*, vol.68
- Tucker C., (2019), « Digital Data, Platforms and the Usual [Antitrust] Suspects: Network Effects, Switching Costs, Essential Facilities”, *Review of Industrial Organization*, vol.54, pp.683 et s.
- Tucker C., (2020), “Digital Data as an Essential Facility: Control”, *CPI Antitrust Chronicle*, February.
- Vannuccini S. and Prytkova E., (2023), “Artificial Intelligence’s new clothes? A system technology perspective”, *Journal of Information Technology*, <https://doi.org/10.1177/02683962231197824>
- van den Boom J., Graef I., Monti G., Argenton C. et van Damme E., (2023), “Towards Market Investigation Tools in Competition Law: The Case of the Netherlands”, *Journal of European Competition Law and Practice*, 14(8), pp.553-564.
- van der Vlist F., Helmond A. and Ferrari F., (2024), “Big AI: Cloud Infrastructure Dependence and the Industrialisation of Artificial Intelligence”, *Big Data and Society*, January-March, pp.1-16.
- von Hippel E., (1986), “Lead users: a source of novel product concepts”, *Management Science*, 32(7), 791–805.
- Vuruma S.K.R., Margetts A., Su J., Faez A. and Srivastava B., (2024), “From Cloud to Edge: Rethinking Generative AI for Low-Resource Design Challenges”, <https://arxiv.org/abs/2402.12702v2>

DOCUMENTS DE TRAVAIL GREDEG PARUS EN 2024
GREDEG Working Papers Released in 2024

- 2024-01** DAVIDE ANTONIOLI, ALBERTO MARZUCCHI, FRANCESCO RENTOCCHINI & SIMONE VANNUCCINI
Robot Adoption and Product Innovation
- 2024-02** FRÉDÉRIC MARTY
Valorisation des droits audiovisuels du football et équilibre économique des clubs professionnels : impacts d'une concurrence croissante inter-sports et intra-sport pour la Ligue 1 de football
- 2024-03** MATHIEU CHEVRIER, BRICE CORGNET, ERIC GUERCI & JULIE ROSAZ
Algorithm Credulity: Human and Algorithmic Advice in Prediction Experiments
- 2024-04** MATHIEU CHEVRIER & VINCENT TEIXEIRA
Algorithm Control and Responsibility: Shifting Blame to the User?
- 2024-05** MAXIME MENUET
Natural Resources, Civil Conflicts, and Economic Growth
- 2024-06** HARALD HAGEMANN
Hayek's Austrian Theory of the Business Cycle
- 2024-07** RAMI KACEM, ABIR KHRIBICH & DAMIEN BAZIN
Investigating the Nonlinear Relationship between Social Development and Renewable Energy Consumption: A Nonlinear Autoregressive Distributed Lag (ARDL) Based Method
- 2024-08** ABIR KHRIBICH, RAMI KACEM & DAMIEN BAZIN
The Determinants of Renewable Energy Consumption: Which Factors are Most Important?
- 2024-09** ABIR KHRIBICH, RAMI KACEM & DAMIEN BAZIN
Assessing Technical Efficiency in Renewable Energy Consumption: A Stochastic Frontier Analysis with Scenario-Based Simulations
- 2024-10** GIANLUCA PALLANTE, MATTIA GUERINI, MAURO NAPOLETANO & ANDREA ROVENTINI
Robust-less-fragile: Tackling Systemic Risk and Financial Contagion in a Macro Agent-Based Model
- 2024-11** SANDYE GLORIA
Exploring the Foundations of Complexity Economics: Unveiling the Interplay of Ontological, Epistemological, Methodological, and Conceptual Aspects
- 2024-12** FRÉDÉRIC MARTY
L'Intelligence Artificielle générative et actifs concurrentiels critiques : discussion de l'essentialité des données